UNITED NATIONS
Office on Drugs and Crime

BULLETIN ON NARCOTICS

Volume LIV, Nos. 1 and 2, 2002

The science of drug abuse epidemiology

UNITED NATIONS OFFICE ON DRUGS AND CRIME
Vienna

BULLETIN ON NARCOTICS

Volume LIV, Nos. 1 and 2, 2002

The science of drug abuse epidemiology

UNITED NATIONS
New York, 2003

UNITED NATIONS PUBLICATION
Sales No. E.03.XI.17
ISBN 92-1-148170-8
ISSN 0007-523X

Sandeep Chawla, Editor,
Yolanda Luna, Editorial Assistant

United Nations Office on Drugs and Crime
Vienna International Centre
P.O. Box, 500
A-1400 Vienna, Austria

Telephone: +(43) (1) 26060-4654
Fax: +(43) (1) 26060-5866

The *Bulletin* is available on the World Wide Web at www.unodc.org

The Office for Drug Control and Crime Prevention became the United Nations Office on Drugs and Crime on 1 October 2002. The Office on Drugs and Crime includes the United Nations International Drug Control Programme (UNDCP).

PREFACE

The *Bulletin on Narcotics* is a United Nations journal that has been in continuous publication since 1949. It is printed in all six official languages of the United Nations—Arabic, Chinese, English, French, Russian and Spanish.

The *Bulletin* provides information on developments in drug control at the local, national, regional and international levels that would benefit the international community.

The present double issue of the *Bulletin* (vol. LIV, Nos. 1 and 2) is devoted to the science of drug abuse epidemiology. The next issue (vol. LV, Nos. 1 and 2) will deal with the practice of drug abuse epidemiology. The United Nations Office on Drugs and Crime wishes to thank Zili Sloboda, who acted as guest editor for these two volumes of the *Bulletin*. Particular thanks also go to Paul Griffiths and Rebecca McKetin, who planned these two volumes while they were working in the United Nations Office on Drugs and Crime (UNODC).

Editorial policy and guidelines for publication

Individuals and organizations are invited by the Editor to contribute articles to the *Bulletin* dealing with policies, approaches, measures and developments (theoretical and/or practical) relating to various aspects of the drug control effort. Of particular interest are the results of research, studies and practical experience that would provide useful information for policy makers, practitioners and experts, as well as the public at large.

All manuscripts submitted for publication in the *Bulletin* should constitute original and scholarly work that has not been published elsewhere or is not being submitted simultaneously for publication elsewhere. The work should be of relatively high professional calibre in order to meet the requirements of a United Nations technical publication. Contributors are kindly asked to exercise discretion in the content of manuscripts so as to exclude any critical judgement of a particular national or regional situation.

The preferred mode of transmission of manuscripts is the Word format. Each manuscript submitted should consist of an original hard copy and a 3.5" diskette, in Word for the text and Excel or Lotus for charts and tables, in any of the six official languages of the United Nations—Arabic, Chinese, English, French, Russian and Spanish. The manuscript should be accompanied by an abstract of approximately 200 words and by a complete set of references numbered in the order of their appearance in the text. The manuscript should be between 10 and 20 double-spaced typewritten pages, including tables, figures and references. Tables should be self-explanatory and should supplement, not duplicate, information provided in the text.

Manuscripts, together with brief curricula vitae of their authors, should be addressed to the Editor, *Bulletin on Narcotics*, United Nations Office on Drugs and Crime, Vienna International Centre, P.O. Box 500, A-1400 Vienna, Austria. A transmittal letter should designate one author as correspondent and include his or her complete address, telephone number, fax number and, if available, electronic mail (e-mail) address. Unpublished manuscripts will be returned to the authors; however, the United Nations cannot be held responsible for loss.

Reprints, purchases and subscriptions

All issues of the *Bulletin* (from vol. I, No. 1 (1949), to the present issue) are available on the home page of the United Nations Office on Drugs and Crime (www.unodc.org.).

The following special issues of the *Bulletin* are also available as United Nations publications:

1991

Involvement of intergovernmental and non-governmental organizations in matters of drug abuse control (vol. XLIII, No. 1)

1992

The role of law enforcement agencies in drug abuse control (vol. XLIV, No. 1)

The environmental impact of drug abuse (vol. XLIV, No. 2)

1993

Policy issues relating to drug abuse and the human immunodeficiency virus (HIV) (vol. XLV, No. 1)

Drug testing in the workplace (vol. XLV, No. 2)

1994

The family and drug abuse (vol. XLVI, No. 1)

General issue on drug abuse (vol. XLVI, No. 2)

1995

Special issue on gender and drug abuse (vol. XLVII, Nos. 1 and 2)

1996

Special issue on rapid assessment of drug abuse (vol. XLVIII, Nos. 1 and 2)

1997 and 1998

Double issue on cannabis: recent developments (vol. XLIX, Nos. 1 and 2, and Vol. L, Nos. 1 and 2)

1999

Occasional papers (vol. LI, Nos. 1 and 2)

2000

Economic and social costs of substance abuse (vol. LII, Nos. 1 and 2)

2001

Dynamic drug policy: understanding and controlling drug epidemics (vol. LIII, Nos. 1 and 2)

Correspondence regarding the purchase of copies of and subscriptions to the *Bulletin* should be addressed as follows:

For Asia, North America, Oceania and South America:

The Chief
Sales and Marketing Office in New York
United Nations Publications
United Nations Headquarters
New York NY 10017
United States of America

For Africa, Europe and the Middle East:

The Chief
Sales and Marketing Office in Geneva
United Nations Publications
United Nations Office in Geneva
Palais des Nations
CH-1211 Geneva 10
Switzerland

CONTENTS

Page

Preface iii

Drug abuse epidemiology: an overview
by Z. Sloboda 1

Estimating the prevalence of problematic drug use: a review of methods and their application
by M. Hickman, C. Taylor, A. Chatterjee, L. Degenhardt, M. Frischer, G. Hay, K. Tilling, L. Wiessing, P. Griffiths and R. McKetin 15

The role of dynamic modelling in drug abuse epidemiology
by C. Rossi 33

Drug use, drug abuse and heterogeneity
by M. D. Glantz and J. D. Colliver 45

The role of rapid assessment methods in drug use epidemiology
by C. Fitch, T. Rhodes, V. Hope, G. V. Stimson and A. Renton 61

Drug treatment data as an epidemiological indicator: methodological considerations and improved analyses
by M. Stauffacher 73

An ethno-epidemiological approach for the multi-site study of emerging drug abuse trends: the spread of methamphetamine in the United States of America
by A. Pach III and E. M. Gorman 87

Human immunodeficiency virus (HIV) infection rates and heroin trafficking: fearful symmetries
by C. Beyrer 103

The need for a global understanding of epidemiological data to inform human immunodeficiency virus (HIV) prevention among injecting drug users
by K. L. Dehne, M. Adelekan, A. Chatterjee and G. Weiler 117

An ethical framework for drug epidemiology: identifying the issues
by C. L. Fry and W. Hall 131

Epidemiology and policy: the post-war context
by V. Berridge 143

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

Drug abuse epidemiology: an overview*

Z. SLOBODA

Institute for Health and Social Policy, University of Akron, Akron, Ohio, United States of America

ABSTRACT

The field of drug abuse epidemiology is a relatively new one and recognition of drug abuse as a phenomenon amenable to study within the field of epidemiology is still forthcoming. Reviews of programmes for meetings of epidemiologic research societies rarely highlight drug abuse, even in association with the transmission of, human immunodeficiency virus (HIV), hepatitis C virus (HCV) or hepatitis B virus (HBV), despite the high rates of prevalence of those infections among drug-abusing populations. The "diagnosis" of drug dependence or addiction and identification of affected populations relies heavily on data concerning problems associated with drug abuse. Both the Diagnostic and Statistical Manual and the International Classification of Diseases, for example, use behavioural criteria to define abuse and dependence. However, most surveys of general and special populations do not include measures that reflect these criteria, but limit themselves to the type of drug used and the frequency of use within a specified time period. With the progress in the development of more biological and medical tools, improved measures of abuse that combine both behavioural and physical assessments will be available to study more carefully the various forms of drug abuse and dependence. The present article has two purposes. The first is to set a framework that describes the current state of the field of drug abuse epidemiology. The second is to provide a synthesis and overview of the articles in the present issue and in vol. LV, Nos. 1 and 2 (2003), of the Bulletin, *which represent work from several countries. The excitement of this emergent field is reflected in each of the articles. The study of drug abuse epidemiology is both challenging and rewarding. As more groups become aware of the global nature of this public health problem and its impact on the political, social and health situations of communities, countries and regions, we hope that drug abuse epidemiology will receive the attention it deserves by policy makers, researchers and, in particular, newly trained social scientists and epidemiologists.*

Keywords: drug use patterns; discontinuation; drug trends and social influences; risk and protective factors; research methods; surveillance; monitoring surveys.

*The author of this paper is a researcher from the United States of America, and much of the introductory statements refer to work conducted in her home country, with which she is most familiar.

Introduction

Epidemiology is a field that studies the distribution of health problems in populations to determine the nature of the health problems, the characteristics of the population affected with the health problems and the environmental conditions or circumstances that are uniquely associated with the health problems and the affected populations. Epidemiologists also describe the course or stages of health problems. They have developed statistical models that simulate the progress of health problems and can use the models to determine the impact of interventions intended to stem their progress. Such models exist for many infectious diseases, including the human immunodeficiency virus (HIV), but have not been applied well to drug abuse. Drug use is seen more as an individualistic behaviour and is therefore not readily conceived of from a population perspective.

One of the most important challenges to drug abuse epidemiologists is defining and measuring the phenomenon of study. In comparison to other health disciplines, the field relies more on behavioural than on physiological or biological criteria to differentiate the range of drug-using behaviours from initiation of use through to full dependence and abuse. Currently there are at least two dimensions of drug-using behaviours that are assessed: degree of use (i.e. frequency of use) and use associated with social and health problems.

There is general recognition of several phases and degrees of drug abuse, from initiation and early-use patterns to long-term chronic use. Onset of use for each drug of abuse, including tobacco, alcohol and, often, inhalants, is used as a measure of incidence of use. Age of onset is an important measure, as many studies have found a relationship between age of onset and subsequent drug abuse or drug dependency (using either the Diagnostic and Statistical Manual or the International Classification of Diseases) [1, 2]. However, substantial portions of those who initiate drug use discontinue such use. Estimates from school surveys from the United States of America show rates of discontinuation ranging from as high as 57 per cent for inhalants, for instance, to as low as 22 per cent for marijuana [3]. Our knowledge about those who continue use, become dependent and develop problems over the course of time is very limited. Preliminary studies indicate that those who discontinue use generally initiated use for social reasons, while those who continue to chronic use, either have a family history of substance use and/or have psychological problems [4]. The paper by Glantz and Colliver [5] in the present issue of the *Bulletin on Narcotics* addresses the need to develop better definitions and measures of drug-using behaviours in order to better understand this phenomenon and improve policies and prevention and treatment approaches to stem the behaviour and its associated social and health problems.

Recent brain-imaging studies are encouraging, as they present evidence of an association between drug abuse and neural/brain changes, some of which may be long-lasting [6-10], and may explain the altered and sometimes problematic behaviours observed with more chronic drug users. These changes have been observed for alcohol, tobacco, marijuana, cocaine, methylenedioxymethamphetamine (MDMA, commonly known as Ecstasy) and methamphetamines. In addition,

these studies have shown that craving for drugs, a key component of the abuse process, is the result of the stimulation of certain parts of the brain when triggers, such as seeing drug paraphernalia or persons administrating drugs, are introduced, even when drug abusers have been detoxified and have not used drugs for long periods of time. These studies suggest that, in the future, drug abuse epidemiologists may be able to develop more precise diagnostic tools that are behavioural measures of changing brain functioning in response to degree of drug use.

Influences on trends and patterns of drug abuse

A number of societal factors and individual vulnerabilities influence drug abuse patterns and trends within a geographic area and during a specific period of time. Unlike infectious diseases, drug abuse, and which drugs are used by whom, is influenced more by social rather than biological factors. However, once individuals begin to abuse drugs, their brains and bodies change, and biological and physiological factors become more dominant. The paper by Rossi in the present issue of the *Bulletin* presents drug-use progression using an infectious disease model as a framework [11]. The development and workings of such models are greatly limited by the availability of information related to the nature of drug abuse. At present, this information is not sufficient to develop more accurate models representing how drug use patterns spread within and across population groups.

The societal factors found to influence both the initiation and continued use of drugs include the public's tolerance of the use of drugs, measured by perceptions of negative social and health consequences of the use of these drugs, as well as the availability of drugs. For instance, researchers for the Monitoring the Future Study found that prior to an upturn or downturn of drug use among adolescents, there are changes in adolescents' perceptions regarding the harmfulness of drugs and the social acceptance of drug use [12, 13]. Indeed, the upturn in drug use among adolescents since 1992 may be related to their changing perceptions of the harmfulness of the use of drugs [14]. Until the early 1990s, during a period of declining drug use, teens' perceptions of harm associated with drug use were comparable to those of their parents' age group. Since 1992, when drug use began to increase, teens' perceptions became more like those of 18- to 25-year-olds, the age group showing the highest rates of drug use over time.

The availability of drugs and the ways in which drugs are marketed also influence which types of drugs are used and who uses them [15]. For instance, in the United States, methamphetamines were once produced by biker gangs or individuals with laboratories and were found mostly in the western region of the country. In the mid-1990s, Mexican drug traffickers began to produce methamphetamines and transport them along the marijuana routes throughout the country, changing reported rates of use in other regions and within subpopulations [16]. The paper by Pach and Gorman [17] in the present issue of the *Bulletin* discusses a multi-city study of these changing patterns of use.

The type, quality and price of drugs also influence use. In the early 1990s, heroin from South America, Mexico and Asia was more plentiful in the United

States than heroin from other parts of the world and was available in a purer form and at lower prices [18]. The increased purity and lower price meant that one did not have to inject the drug for a "high", but could snort it instead. This change in mode of administration made the use of heroin more attractive to populations such as suburban youth, who would not normally use heroin. However, past history of heroin use in the United States has shown that once dependent, it is more likely that one will inject [19]. Recent reports in both the United States and Europe show that even a drug that is usually smoked, such as crack-cocaine, can be dissolved in vinegar or lemon juice and injected for a more immediate effect [20, 21]. The movement from snorting to injecting puts the user at risk for a number of health problems. In the present issue of the *Bulletin*, Beyrer [22] discusses the association of changing trafficking patterns and the spread of HIV and hepatitis B and C in Asia. Even when preferred drugs become difficult to obtain because of effective interdiction or other circumstances, the changes in patterns of use can present problems for policy makers and service providers [23]. All of these patterns can have major health effects on the drug-using population that in turn will impact treatment and medical services.

Availability of drugs certainly can be seen to influence drug use patterns, but social and economic trends also appear to have an impact on who uses drugs and how they are used. Reports from countries in social and economic turmoil indicate growing rates of drug abuse among youngsters. Koshkina [24] describes the situation as it occurred in the Russian Federation.

The differential impact of these social influences results in the variation in drug use patterns observed among cities, countries and regions [25]. An understanding of the importance of the social factors that influence drug abuse trends should guide national and local policies regarding approaches to be taken to alter the devastating consequences of the problem [26, 27].

Individual or personal characteristics also influence drug-using behaviour patterns. A large number of factors have been found to explain why some people use drugs and others do not. They range from personality traits and other genetic and biological vulnerabilities to peer association to community problems [4]. We know that more people try drugs than actually abuse or are dependent on drugs. Although we still do not understand why some people move on to abuse and dependence, epidemiological studies that have followed young people over time do point to biological bases for significant increased drug involvement. It has been found that having parents or other relatives who have substance abuse problems or certain psychiatric illnesses increases one's risk for drug abuse or dependence [28-30]. Data from both people in treatment for drug abuse and people in treatment for mental illness indicate that many have both problems.

There is extensive literature that implicates a range of risk factors for drug use, including parent-child attachment, parental monitoring, peer association and school failure [31]. Researchers have made an effort to organize these in such a way as to generate hypotheses about the origins of drug use that could be targeted in prevention strategies [32, 33]. Research also shows that the more risk factors a person experiences in their youth, the more likely it is that drug abuse will occur [34].

Recent research into why many adolescents with such risk factors for drug use do not abuse drugs has shown that certain factors temper risks. These tempering or protective factors include family bonding and bonding to prosocial groups, behaviours and institutions [34]. There is also a growing focus on less well-articulated factors, such as resilience and positive assets [35-37]. However, all of these approaches to the issue of onset lack specificity for drug abuse.

In addition, epidemiologic studies have found a consistent sequencing of the use of substances, particularly among adolescents. Denise Kandel was the first to observe this sequencing in her research in the late 1970s [38], the results of which have been replicated in other longitudinal studies. The sequence goes from the use of alcohol and/or tobacco and marijuana to other drugs of abuse. The underlying mechanisms for this sequencing are not understood and the inevitability that everyone who smokes or drinks alcohol or even uses marijuana will move on to the next substance, is not total.

Our failure to specify risk is very much related to our failure to refine our diagnosis of drug use and abuse, [4]. More molecular epidemiologic research is needed linking neuroscience and behavioural science.

Epidemiologic methods specific to studying drug abuse

The stigmatized nature of drug abuse in most regions of the world has made it necessary for drug abuse epidemiologists to develop methods of sampling and targeting that are unique to the field. Over the past two decades, the integration of qualitative and quantitative or ethnographic approaches has provided significant information about drug-using populations that has been useful to both prevention efforts [14, 39] and treatment providers [40], as well as to policy makers [41]. The papers by Dehne and colleagues [42], Hickman and colleagues [43], Rossi [11], Degenhardt and colleagues [23], Fitch and colleagues [44] and Kemmesies [45] describe some of these approaches, including capture-recapture, multiplier methods, dynamic modelling, rapid assessment, snowball sampling, key informant interviews and focus groups.

The stigmatized nature of drug abuse suggests certain limitations in the usual epidemiologic approaches to data collection. It is generally recommended that several systems be established that collect data from both general and more at-risk populations to determine the extent and nature of drug-using behaviours in an area and to conduct special studies to address more focused research questions. The most widely used approaches to monitor drug abuse include use of existing data to determine general drug use patterns within the drug-using populations (surveillance) and surveys to estimate incidence and prevalence rates within a general population (monitoring). Both approaches are discussed below. They should be designed to be interconnected, each serving to guide the activities of the other. If they are well designed and comprehensive, surveillance systems will reflect the dynamic nature of drug-using behaviours in an area. New drugs being used, usual drugs used in new ways or new populations of users will be detected with a good surveillance system containing timely information from a number of sources. As

surveys of representative samples of general household or school populations provide rates of new and current users over time, consistency in sampling methods and data collection approaches is important to monitor trends in use. Information on new patterns of use detected through surveillance of drug use among drug-using populations must be integrated into monitoring surveys so that the spread of a new pattern can be documented and watched. Finally, key research questions that arise from either the surveillance system or monitoring surveys serve to form special population studies. The paper by Medina-Mora and her colleagues [46] discusses the importance of these systems for understanding drug abuse patterns in Mexico, while Hartnoll [47] discusses the establishment of a multi-systems approach in Europe.

Drug abuse surveillance systems

Surveillance research collects information over time from existing or archival data systems that reflect the consequences of drug abuse, such as drug abuse treatment, arrest reports for adults and adolescents, hospital emergency room visits, hospital discharge information, mortality information and infectious disease reports.

These systems tend to indicate new and existing drug abuse patterns within the drug abusing community and among new users, who may have some negative health effects. Surveillance epidemiologists look for reports of new types of drugs that are abused, new ways to administer drugs and changing characteristics of drug users. Any of these changes is a sign of new patterns of drug abuse that need to be studied further.

As the data used in surveillance systems reflect consequences and are not population-based, it is difficult to develop incidence or prevalence rates from them. However, one can review these data over time, draw conclusions from them about where new drug abuse patterns exist and follow how these patterns spread across geographic areas. The paper by Sloboda and Kozel [16] reports on what is probably the oldest ongoing drug abuse surveillance system, the Community Epidemiology Work Group. This group of experts representing 20 to 21 United States cities has met biennially since 1975. Over time, the CEWG has evolved common data elements across sites and a reporting format that provides descriptions of drug use patterns among varying populations within each geographic area. Both common, and unique patterns have been observed, and the movement of certain patterns across areas has been well documented. Other model surveillance systems have been or are being established in several countries and regions around the world. The papers by Stauffacher [48], Parry and colleagues [49], Douglas [50], Bless [51] and Koshkina [24] present the experiences of each group in the development of their systems. Several political and resource barriers confront these researchers [52]; however, perseverance and commitment has enabled several systems to be institutionalized [41].

Recognition of the global nature of drug abuse and the utility of surveillance systems brought a group of epidemiologists representing countries around the

world to form the International Epidemiology Work Group on Drug Abuse (IEWG). Until 1999, the Group met annually to discuss new trends in their countries or regions. The United Nations International Drug Control Programme's Global Assessment Programme on Drug Abuse (GAP) has continued the work initiated by IEWG. Both IEWG and GAP are working towards building comparable epidemiologic foundations in countries around the world. The paper by Griffiths and McKetin [53] provides an overview of this approach.

Monitoring surveys

Monitoring surveys are generally population-based, which means they collect data on people who live in a defined geographic area or who share similar characteristics. These surveys generally take "snapshots" of drug abuse in the defined population on a regular basis over long periods of time. Each time the survey is conducted the same measures are taken or questions asked in order to maintain comparability so that trends can be detected. In general, these surveys are either of persons who reside in households or other stable environments or of persons who represent specific populations, such as students, homeless groups, military or institutionalized patients.

Such surveys are expensive to conduct. Drug-using behaviours are usually rare events in a general population, and therefore large sample sizes are required in order to improve the precision of estimations of prevalence. Furthermore, drug abusers may be homeless or live in transient situations, often clustering in geographic areas, and therefore may be more likely to be missed if traditional sampling and interviewing techniques are applied. Until recently, many of these surveys required trained and skilled interviewers, as the survey methods used face-to-face interviews in combination with forms completed by the respondents. Research has shown that telephone and mailed surveys yield the lowest rates of drug use among household members. However, with current advances in computer technology, the surveys can be administered electronically [54, 55]. Some groups are more difficult to engage in surveys. For example, children and adolescents interviewed in the household when a parent/caregiver is present tend to report lower rates of drug use, even when techniques that imply privacy, such as computer-assisted interview procedures, are used.

Surveys of students conducted in the classroom using self-administered data collection instruments provide information regarding the incidence or initiation of the use of drugs and the prevalence rates for overall drug use. Age of onset, as well as the sequencing of substances used (e.g. tobacco, alcohol, inhalants and marijuana), assist prevention programmers in their design and targeting interventions.

Sampling issues, the timing of surveys and the drugs included in the surveys are challenges to both household and school surveys. Response rates vary by size of community, with metropolitan areas often having lower participation rates than suburban communities [56]. The lower population density of rural areas tends to

increase the expense, both of sampling residents of these areas and of conducting the survey with them. Surveys conducted during the day may bias who responds to the survey, with more stay-at-home mothers being available to respond and excluding those who are working. Attention must be given to these issues when conducting a survey.

School surveys conducted on one particular day may miss those who are absent from school because of their substance use for reasons such as associated illnesses or suspension or even expulsion. Furthermore, students who leave or drop out of school, and therefore do not always participate in the surveys, may be at highest risk for drug use.

The relationship between surveillance studies and monitoring surveys has not been well recognized. Yet this relationship is important to a more thorough understanding of existing patterns of drug use within a geographic area and identification of the characteristics of use and of the users. Examination of information from both sources has shown that it may take between one to two years for an emergent pattern of drug use found in a surveillance system to move from the endemic drug-using population into the general population, where it is reported in a survey. By establishing these systems, drug abuse epidemiologists can gain a better understanding of the problem within an area and can inform prevention and treatment practitioners about new issues of which they must be aware.

Special studies

Information from surveillance systems and surveys suggest other research questions or hypotheses that can be studied in more focused research. Both the papers by Pach and Gorman [17] and Dengehardt, Topp and Day [23] are good examples of studies that grew out of information first observed in surveillance systems. When the reports from the Community Epidemiology Work Group showed that methamphetamines were becoming available in new "virgin" areas and among new populations, a multi-city study was designed to learn more about how these drugs were being used, by whom and with what impact. Similarly, the Illicit Drug Reporting System in Australia allowed reports of decreased heroin availability to be examined more closely through a detailed analysis of data from interviews with heroin users and key informants and existing data which was routinely collected through the System. Medina-Mora and colleagues [46] also mention a number of special population studies conducted in Mexico relating to issues emerging from available surveillance information. Again, such studies have important policy implications.

Ethical issues

One cannot establish such information systems or conduct epidemiologic studies focusing on drug use without being mindful of ethical considerations within the context of the laws of each jurisdiction. In their paper, Fry and Hall [57] lay out

three such ethical considerations: *(a)* free and informed consent; *(b)* confidentiality, privacy and legal hazard; and *(c)* safety issues for those collecting data in the field. All epidemiologists conducting studies must be cognizant of the potential ethical pitfalls. The degree of concern, however, differs somewhat when dealing with a highly stigmatized health issue, such as drug abuse. Surveys of adolescents, in most cases, will require formal consent from parents or guardians. The conduct of interviews with intoxicated or otherwise impaired respondents also poses ethical problems concerning free consent and the reliability of reported information. Protection of human subjects and the information provided by them is paramount in any study; when the information collected concerns potentially illegal activities, however, additional efforts are required to ensure privacy and confidentiality.

Many studies that collect data in the field provide safeguards for those collecting the data. This is also true for studies of drug abusers. Sending interviewers in pairs to certain areas and providing cell phones and special training are examples of ways to prevent problems. Such procedures and processes add time and costs to any drug abuse study, and this is often not understood by funding organizations.

Finally, the drug abuse epidemiologist's work also faces challenges from those who do not understand either the phenomenon of drug abuse or the many well-designed, tested and accepted methods developed in the field. The reliability and validity of self-reported information on drug-using behaviours [58-60] and the value of qualitative studies are among the most pressing issues facing drug abuse epidemiologists. A range of other issues also confronts the drug abuse epidemiologist in particular when dealing with policy makers. The need to communicate findings from the research for policy decisions, while maintaining scientific rigor requires much patience and time. Some recommendations provided by Musto and Sloboda [27] are key to this process.

Conclusions

The field of drug abuse epidemiology is challenged by the very nature of the phenomenon being studied. The present issue of the *Bulletin* and Nos. 1 and 2 of vol. LV (2003) describe these challenges, discusses how drug abuse epidemiologists are addressing them and makes suggestions for the expansion of drug abuse epidemiologic studies around the world. The importance of collaboration and communication with policy makers is stressed and recommendations are made for easing this process so that it benefits all involved—the researcher, the policy maker and others who are directly and indirectly affected. Drug abuse is a global problem impacting not only individual lives but also whole communities. An understanding of the specific aspects of the problem, including which drugs are being used and by whom, will go a long way towards containing the problem. Drug abuse epidemiology is an emergent field that is rapidly accumulating methodological and scientific knowledge and, finally, the recognition it needs to grow.

References

1. J. C. Anthony and K. R. Petronis, "Early-onset drug use and risk of later drug problems", *Drug and Alcohol Dependence,* vol. 40, No. 1 (1995), pp. 9-15.
2. D. J. DeWit and others, "The influence of early and frequent use of marijuana on the risk of desistance and of progression to marijuana-related harm", *Preventive Medicine,* vol. 31, No. 5 (2000), pp. 455-464.
3. National Institute on Drug Abuse, *National Survey Results on Drug Use from the Monitoring the Future Study, 1975-1999, vol. I: Secondary School Students,* National Institutes of Health publication No. 98-4345, 1998.
4. M. D. Glantz and R. W. Pickens, "Vulnerability to drug abuse: introduction and overview", in *Vulnerability to Drug Abuse*, M. D. Glantz and R. W. Pickens, eds. (Washington, D.C., American Psychological Association, 1992), pp. 1-14.
5. M. Glantz and J. Colliver, "Drug use, drug abuse and heterogeneity", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 45-59.
6. N. D. Volkow and others, "Changes in brain glucose metabolism in cocaine dependence and withdrawal", *American Journal of Psychology*, vol. 48, 1991, pp. 621-626.
7. A. R. Childress and others, "Limbic activation during cue-induced cocaine craving", *Society for Neuroscience Abstracts*, vol. 21, No. 3 (1995), p. 1956.
8. J. Altman, "A biological view of drug abuse", *Molecular Medicine Today*, June 1996, pp. 237-241.
9. S. Grant and others, "Activation of memory circuits during cue-elicited cocaine craving", *Proceedings of the National Academy of Sciences*, vol. 93, 1996, pp. 12040-12045.
10. R. C. Taylor and others, "Tobacco craving: intensity-related effects of imagery scripts in drug abusers", *Experimental and Clinical Psychopharmacology*, vol. 8, No.1 (2000), pp. 75-87.
11. C. Rossi, "The role of dynamic modelling in drug abuse epidemiology", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 33-44.
12. J. G. Bachman, L. D. Johnston and P. M. O'Malley, "Explaining the recent decline in cocaine use among young adults: further evidence that perceived risks and disapproval lead to reduced drug use", *Journal of Health and Social Behaviour*, vol. 31, No. 2 (1990), pp. 173-184.
13. J. G. Bachman, L. D. Johnston and P. M. O'Malley, "Explaining recent increases in student's marijuana use: impacts of perceived risks and disapproval, 1976 through 1996", *American Journal of Public Health*, vol. 88, No. 6 (1998), pp. 887-892.
14. Z. Sloboda, "Changing patterns of drug abuse in the United States: connecting findings from macro- and micro-epidemiologic studies", *Substance Use and Misuse*, vol. 37, Nos. 8-10 (2002) pp. 1229-1251.
15. M. L. Van Etten and J. C Anthony, "Comparative epidemiology of initial drug opportunities and transitions to first use: marijuana, cocaine, hallucinogens and heroin", *Drug and Alcohol Dependence*, vol. 54, No. 2 (1999), pp. 117-125.
16. Z. Sloboda and N. J. Kozel, "Understanding drug trends in the United States of America: the role of the Community Epidemiology Work Group as part of a comprehensive drug information system", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).

17. A. Pach III and E. M. Gorman, "An ethno-epidemiological approach for the multi-site study of emerging drug abuse trends: the spread of methamphetamine in the United States of America", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 87-102.

18. "Drug trafficking in the United States", website of the United States Drug Enforcement Administration www.usdoj.gov/dea/pubs/state_factsheets.html.

19. W. A. Johnson and L. J. Ouellet, "The injection of crack cocaine among Chicago drug users", *American Journal of Public Health*, vol. 86, No. 2 (1996), p. 266.

20. G. M. Hunter, M. C. Donoghoe and G. V. Stimson, "Crack use and injection on the increase among injecting drug users in London", *Addiction*, vol. 90, No. 10 (1995), pp. 1397-1400.

21. L. J. Ouellet, W. W. Wiebel and A. D. Jimanez, "Team research methods for studying intranasal heroin use and its HIV risks", *National Institute on Drug Abuse Research Monograph*, vol. 57 (1995), pp. 182-211.

22. C. Beyrer, "Human immunodeficiency virus (HIV) infection rates and heroin trafficking: fearful symmetries", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 103-116.

23. L. Degenhardt, L. Topp and C. Day, "Issues surrounding the detection of a reduction in drug supply: the heroin shortage in Australia in 2001", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).

24. E. A. Koshkina, "Trends in the prevalence of psychoactive substance use in the Russian Federation", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).

25. National Institute on Drug Abuse, *Epidemiologic Trends in Drug Abuse, vol. I: Proceedings of the Community Epidemiology Work Group, Highlights and Executive Summary*, National Institutes of Health publication No. 02-5109, 2002.

26. V. Berridge, "Epidemiology and policy: the post-war context", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 143-151.

27. D. F. Musto and Z. Sloboda, "The influence of epidemiology on drug control policy", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).

28. D. B. Kandel and others, "Psychiatric comorbidity among adolescents with substance use disorders: findings from the MECA Study", *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 38, No. 6 (1999), pp. 693-699.

29. K. S. Kendler and C. A. Prescott, "Cannabis use, abuse and dependence in a population-based sample of female twins", *American Journal of Psychiatry*, vol. 155, No. 8 (1998), pp. 1016-1022.

30. K. Yamaguchi and D. B. Kandel, "Patterns of drug use from adolescence to young adulthood: III. Predictors of progression", *American Journal of Public Health*, vol. 74, No. 7 (1984), pp. 673-681.

31. N. Z. Weinberg and others, "Adolescent substance abuse: A review of the past 10 years", *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 37, No. 3 (1998), pp. 252-261.

32. J. D. Hawkins, R. F. Catalano and J. Y. Miller, "Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention", *Psychological Bulletin*, vol. 112, No. 1 (1992), pp. 64-105.

33. R. J. Pandina, "Risk and protective factor models in adolescent drug use: putting them to work for prevention", *National Conference on Drug Abuse Prevention Research*: *Presentations, Papers, and Recommendations*, National Institutes of Health publication No. 98-4293, 1998, pp. 17-26.

34. B. H. Bry, P. McKeon and R. J. Pandina, "Extent of drug use as a function of number of risk factors", *Journal of Abnormal Psychology*, vol. 91, No. 4 (1992), pp. 273-279.

35. J. S. Brook and others, "Young adult's drug use: a 17-year longitudinal inquiry of antecedents", *Psychological Reports*, vol. 80, No. 3, part 2 (1997), pp. 1235-1251.

36. M. D. Glantz and Z. Sloboda, "Analysis and reconceptualization of resiliency", in *Resilience and Development*: *Positive Life Adaptations*, M. D. Glantz and J. Johnson, eds. (New York, Kluwer Academic/Plenum Press Publishers, 1999), pp. 109-126.

37. P. L. Benson, J. Galbraith and P. Espeland, *What kids need to succeed: proven practical ways to raise good kids* (Minneapolis, Free Spirit Publishing, Inc., 1998).

38. D. B. Kandel, "Stages of adolescent involvement in drug use", *Science*, vol. 190, 1975, pp. 912-914.

39. Z. Sloboda, "Assessing substance abuse in the community", plenary presentation for the Second National Conference on Prevention, National Institute on Drug Abuse, Washington, D.C., 10 August 2001.

40. Z. Sloboda, "Building it together: qualitative and quantitative data collection methods", plenary presentation for the Treatment Needs Assessment Conference, Ohio Division for Drug Abuse and Alcohol Services, Columbus, Ohio, 22 March 2001.

41. L. Topp and R. McKetin, "Supporting evidence-based policy-making: a case study of the Illicit Drug Reporting System of Australia", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).

42. K. L. Dehne and others, "The need for a global understanding of epidemiological data to inform human immunodeficieny virus (HIV) prevention among injecting drug users", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 117-130.

43. M. Hickman and others, "Estimating the prevalence of problematic drug use: a review of methods and their application", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 15-32.

44. C. Fitch and others, "The role of rapid assessment methods in drug use epidemiology", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 61-72.

45. U. Kemmesies, "How to reach the unknown: the snowball sampling technique", in *Understanding and responding to drug use*: *the role of qualitative research*, J. Fountain, ed., European Monitoring Centre for Drugs and Drug Addiction, Scientific Monograph Series No. 4, 2000, pp. 265-271.

46. M. E. Medina-Mora and others, "Mexico: epidemiological diagnosis systems for drug abuse", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).

47. R. Hartnoll, "Drug epidemiology in the European institutions: historical background and key indicators", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).

48. M. Stauffacher, "Drug treatment data as an epidemiological indicator: methodological considerations and improved analysis", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 73-85.
49. C. D. H. Parry, A. Pluddemann and J. Strijdom, "Developing the Southern African Development Community Epidemiology Network on Drug Use: methods and issues", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).
50. K. Douglas, "The Caribbean Epidemiological Network: the complexities of developing a regional perspective", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).
51. R. Bless, "Experiences of the Pompidou Group of the multi-city network, 1983-2002", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).
52. M. Warner-Smith, "The challenge of developing drug information systems in Africa", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).
53. P. Griffiths and R. McKetin, "Developing a global perspective on drug consumption patterns: the challenge for drug epidemiology", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003).
54. M. Ramsey and A. Percy, "A national household survey of drug misuse in Britain: a decade of development", *Addiction*, vol. 92, No. 8 (1997), pp. 931-937.
55. J. Gfroefer, D. Wright and A. Kopstein, "Prevalence of youth substance use: the impact of methodological differences between two national surveys", *Drug and Alcohol Dependence*, vol. 47, No. 1 (1997), pp. 19-30.
56. J. Gfroefer, J. Lessler and T. Parsley, "Studies of non-response and measurement error in the National Household Survey on Drug Abuse", *National Institute on Drug Abuse Research Monograph*, No. 167, 1997, pp. 273-295.
57. C. L. Fry and W. Hall, "An ethical framework for drug epidemiology: identifying the issues", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002), pp. 131-142.
58. S. Darke, "Self-report among injecting drug users: a review", *Drug and Alcohol Dependence*, vol. 51, No. 3 (1998), pp. 253-263.
59. E. Finch and J. Strang, "Reliability and validity of self-report: on the importance of considering context", *Drug and Alcohol Dependence*, vol. 51, No. 3 (1998), p. 269.
60. A. M. Shillington and others, "Self-report stability for substance use over 10 years: data from the St. Louis epidemiologic catchment study", *Drug and Alcohol Dependence*, vol. 40, No. 2 (1995), pp. 103-109.

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

Estimating the prevalence of problematic drug use: a review of methods and their application

M. HICKMAN

Centre for Research on Drugs and Health Behaviour, Imperial College, University of London, London

C. TAYLOR

National Addiction Centre, London

A. CHATTERJEE

Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland

L. DEGENHARDT

National Drug and Alcohol Research Centre, University of New South Wales, New South Wales, Australia

M. FRISCHER

Department of Medicines Management, Keele University, Keele, United Kingdom of Great Britain and Northern Ireland

G. HAY

Centre for Drug Misuse Research, University of Glasgow, Glasgow, United Kingdom

K. TILLING

Department of Social Medicine, University of Bristol, Bristol, United Kingdom

L. WIESSING

European Monitoring Centre for Drugs and Drug Addiction, Lisbon

P. GRIFFITHS and R. McKETIN

United Nations Office on Drugs and Crime, Vienna

ABSTRACT

Policy makers increasingly need estimates of the prevalence of problematic drug use, such as injecting and the use of "crack". In the present article, the authors review indirect methods of estimating the prevalence of problematic drug use. Those methods utilize existing data on a sample of problematic drug users as "raw" material and then "indirectly" estimate the proportion of the total number of problematic drug

users sampled in the raw material; that is, the methods estimate the sampling intensity of the raw data. That analogy is used to explain a number of indirect estimation techniques, focusing on capture-recapture and multiplier methods, the methods most often used in settings in developing countries. Assumptions underpinning indirect estimation techniques are presented, together with examples of their application. In addition, there is a discussion of the need to develop routine data sources that can be used in indirect prevalence estimation procedures.

Keywords: injecting drug use; problem drug use; capture-recapture; multipliers; indirect estimation; prevalence; epidemiology.

Introduction

Injecting drug use affects over 135 countries. It is estimated that there are as many as 3 million injecting drug users worldwide who have been infected with the human immunodeficiency virus (HIV). Between 5 and 10 per cent of all HIV infections were acquired through injecting drug use; in some countries in Asia and Europe, injecting drug use accounts for over one half of HIV infections [1]. Policy makers need prevalence estimates to inform them and help guide them in formulating drug policy [2]; yet there is a dearth of reliable and valid estimates of injecting drug use around the world. Prevalence estimates, such as estimates of the number of heroin users or injecting drug users in the population, are required for several aspects of policy-making: planning and allocating resources for the control, treatment and prevention of problematic drug use and its consequences; monitoring key targets of drug policy, such as assessing coverage of treatment and harm reduction activities; and assisting in the interpretation of other data and research on drug-related consequences. Because of the limitations of the survey methodology in making such prevalence estimates [3], alternative, less costly techniques have been developed by drug abuse epidemiologists. That has led to an increasing number of manuals and review papers on prevalence estimation methods [4-8].

Limitations of survey approaches

Properly conducted population or general household surveys can be considered the "gold standard" for estimating how many people are in a target population and can be effective in monitoring common drug use behaviour, such as tobacco, alcohol or cannabis use, among the population. However, population surveys are less effective in estimating more problematic forms of drug use, such as heroin use or injecting drug use. There are several reasons for that. First, those forms of drug use are uncommon; even in surveys involving tens of thousands of people, only a few hundred people will experience problematic drug use behaviour, and even fewer will willingly report on such behaviour because of the stigma associated with it. Secondly, survey respondents tend to underreport the more serious forms of drug use. Thirdly, general household surveys tend to "miss" a lot of problematic drug users because of the lifestyle of the drug user. In particular, problematic drug users

tend to cluster in certain geographical areas and are less likely to reside in conventional housing; and usually, a certain proportion of such users are incarcerated in residential treatment or other similar facilities and are therefore not taken into account in such surveys. Such biases usually result in an underestimation of the population size of problematic drug users; that means that patterns of problematic drug use, such as injecting drug use, are simply not measured reliably enough or in sufficient numbers to measure prevalence or detect significant changes over time [9].

Indirect estimation methods

Several methods have been developed to make indirect estimates of the prevalence of problematic drug use and other socially stigmatized behaviours:

Capture-recapture methods—closed populations
Capture-recapture methods—open populations
Multiplier methods
Event-based multipliers
Synthetic estimation/multiple indicator methods
Truncated Poisson
Back-calculation

The methods are based on having access to existing data sources that correctly identify cases of problematic drug use and on being able to then establish what proportion of the population of problematic drug users is represented in those data sources.

To more accurately measure the size of the population of problematic drug users, a number of indirect estimation methods have been developed. In the present article, the authors review the more common indirect estimation methods, their assumptions and their application. The article focuses on heroin use and injecting drug use, as they are the subject of many indirect estimation techniques and are perhaps responsible for the greatest public health burden.

Capture-recapture

Bishop and others were among the first to identify the potential for capture- recapture methods in estimating the prevalence of addiction [10]. Capture-recapture methods had been developed by animal ecologists as one method among others of estimating animal abundance [11, 12]. Early in the twentieth century they were co-opted for use in estimating the undercount of census populations; more recently, they have been used extensively in epidemiological studies, in particular to estimate the underreporting of surveillance systems [13-16]. (For a fuller discussion of capture-recapture methods, see Fienberg [14] and Hook and Regal [15].)

In essence, capture-recapture involves the collection of two or more sources of data on problematic drug users. Information on the number of matches between the data sources (the number of people that occur in more than one data source)

is used to estimate the proportion of the total number of problematic drug users in the sample. These are then combined to generate the prevalence estimate. An example for two samples is shown in figure I.

Figure I shows the number of people in two data sources, n_1 and n_2, and m matches of people registered in both data sources. Together, those parameters are used to estimate N, the total population of problematic drug users. The method assumes that n_1/N is equivalent to m/n_2. For example, Mastro and others carried out a two-sample study in Bangkok in 1991 [17] (see figure II). The first sample consisted of 4,064 heroin users in methadone treatment, and the second consisted of 1,540 arrested persons whose urine samples were found to be opioid-positive. There were 171 people listed in both samples, giving an estimate of 36,600 opiate users, or 0.5 per cent of the total population of Bangkok in 1991. Capture-recapture can also be used with multiple samples and without lists of names as the data sources. Those methods will be discussed following consideration of the assumptions underpinning capture-recapture methodology.

Figure I. Using the capture-recapture method with two data sources

		Data source 2 (S2)		
		Yes	*No*	
Data source 1 (S1)	*Yes*	*a* (*m*)	*b*	n_1
	No	*c*	*? (x)*	
		n_2		*N*

Assuming $n_1/N = m/n_2$,

then the population estimate $N = (n_1 \times n_2)/m$

Number observed $n = a + b + c$

Number unobserved/hidden $x = N - n$, or $(c \times b)/a$

Confidence interval 95 per cent = $1.96 \sqrt{(n_1 \times n_2 \times b \times c)/m^3}$

Source: T. D. Mastro and others, "Estimating the number of HIV-infected injection drug users in Bangkok: a capture-recapture method", *American Journal of Public Health*, vol. 84, No. 7 (1994), pp. 1094-1099.

a or *m* = matches or marks; number of people in both S1 and S2

b = number of people in S1 but not in S2

c = number of people in S2 but not in S1

x = hidden population; number of people not in S1 or S2

n_1 = number of people in S1

n_2 = number of people in S2

N = total population

Figure II. Estimating the number of injecting drug users in Bangkok, 1991

Data Source		*Arrestees with opioid-positive urine samples (S2)* *Yes*	*No*	
Methadone maintenance (S1)	*Yes*	171	3 893	4 064
	No	1 369	*?* (*x*)	
		1 540		*N*

Population estimate $N = (n_1 \times n_2)/m = (4{,}064 \times 1{,}540)/171 = 36{,}599$

Number observed $n = a + b + c = 171 + 3{,}893 + 1{,}369 = 5{,}433$

Number hidden $x = N - n$ (or $(c \times b)/a$) = 36,599 – 5,433 (or (1,369 x 3,893)/171) = 31,166

Confidence interval 95 per cent = $1.96 \sqrt{(n_1 \times n_2 \times b \times c)/m^3}$

$= 1.96 \sqrt{(1{,}540 \times 4{,}064 \times 3{,}893 \times 1{,}369)/171^3} = 4{,}516$

Rounded estimate of the number of injecting drug users in Bangkok in 1991: 36,600 (32,000-40,800)

Source: T. D. Mastro and others, "Estimating the number of HIV-infected injection drug users in Bangkok: a capture-recapture method", *American Journal of Public Health*, vol. 84, No. 7 (1994), pp. 1094-1099.

a or *m* = matches or marks; number of people in both S1 and S2

b = number of people in S1 but not in S2

c = number of people in S2 but not in S1

x = hidden population; number of people not in S1 or S2

n_1 = number of people in S1

n_2 = number of people in S2

N = total population

The main assumptions of the capture-recapture method are: *(a)* independence of the data sources; *(b)* homogeneity of the data sources; *(c)* correct classification of cases; *(d)* having a closed population; and *(e)* the data sources being representative of the population that is being studied. In practice, it is inevitable that almost all of the assumptions are violated to some extent. That does not detract from the value of the prevalence estimate and the exercise, but it is important to "treat and tread cautiously" in critically reviewing capture-recapture studies of problematic drug use.

The assumption of independence refers to the likelihood that, if a person is listed in one data source, his or her listing in a second data source being used in

the capture-recapture estimation is random and independent of the first data source. That assumption cannot be tested with only two data sources; however, if three or more sources of data are available, then violating that assumption can be overcome to some extent. That advantage comes at a slight cost as more complex statistical methods are required. Log-linear models are used to fit "dependencies", or interactions between the data sources, and to generate an adjusted prevalence estimate. The methods can be taught relatively quickly to someone with rudimentary statistical knowledge using statistical software, such as SPSS or Stata or Generalized Linear Interactive Modelling (GLIM) [6]. Alternatively, estimates can be calculated using formulae provided by Bishop and others [10]. Instead of assuming independence of data sources, multi-source capture-recapture assumes that there is no interaction or interdependency between all of the data samples. Evidence of relationships between each pair of sources may suggest that that assumption has been violated.

Homogeneity of the data sources requires that all problematic drug users are equally likely to turn up on a data source. When certain subgroups of drug users are more likely to show up in a particular data source (for example, representation on treatment admission data would be affected by the accessibility of services, the severity of the drug problem and so on), then heterogeneity of the data sources becomes a problem. Hook and Regal have argued that heterogeneity is inevitable when using health data [15] and therefore comparisons of key variables across data sets should always be conducted prior to final analysis. The effects of heterogeneity can be limited by stratifying the subjects on those characteristics that may confound the analyses and running separate models (for males and females or for different age groups). The only problem with that solution is that the data required to run separate models may be missing. At the least, stratification should be carried out to check for evidence of heterogeneity. More complex models have been developed that allow the fitting of covariates within a single model, but they require greater statistical expertise and have not yet been conducted in studies of problematic drug use [18].

Correct classification refers to the extent to which all the subjects in the data sources are correctly identified as problematic drug users and that all matches between data sources are correctly identified. Bias due to misclassification is reduced if there are sufficient data to identify matches adequately and the data sources are accurate and reliable—both of which may be challenging when collecting data on problematic drug users.

Maintaining a closed population would require that there was no migration, no deaths and no new cases of problematic drug users during the study period. That is clearly impossible, but the bias can be limited if the study time interval is short in comparison with the life cycle of the subject (for example, one year or less).

The assumption of representativeness refers to whether the data sources used in the study are representative of the target population. That is less of an issue with animal studies or many epidemiological capture-recapture studies: for example, fish caught from a lake clearly represent the target population ("fish in the

lake"). There are implications, however, for problematic drug use as each potential data source only partially covers the target population. At any particular point in time, the definition of a problematic drug user as "a person who experiences social, psychological, physical or legal problems and/or dependence as a consequence of his or her own use of drugs" implies that studies should include treatment and criminal justice data sources. Similarly, it is questionable whether studies that use only data sources of problematic drug users in treatment are able to estimate the total population of "problematic drug users".

Capture-recapture studies of problematic drug use need to be scrutinized more carefully than studies of other health problems because of the difficulties in obtaining either large or representative samples of problematic drug users [19-21].

Finally, in capture-recapture there remains what Cormack calls a "leap of faith" [22]. It is assumed that the model that fits the observed data applies also to the "unobserved" population. But there is no way of testing that assumption. It is important, therefore, to use other knowledge and expertise to judge whether the estimates make sense—and ideally seek corroborating evidence.

Example of a capture-recapture study with multiple data sources

Data from a capture-recapture study with multiple data sources are shown in tables 1 and 2 [23]. Table 1 shows the data after matching four data sources: *(a)* HIV tests mentioning injecting drug use; *(b)* attendees of specialist drug treatment centres; *(c)* attendees of needle exchange programmes; and *(d)* police arrests for possession of heroin or benzodiazepines. All the data sources provided information on the date of birth, sex and initials of the drug user (the first character of his or her first name and the first character of his or her surname), which could be used for matching. A total of 3,760 records were collected from the four data sources, representing 2,866 individual reports after matching, 4 of which were represented in all four data sources.

Table 1. Capture-recapture study in Glasgow: 3,760 reports of injecting drug users from four data sources, 1990

Data source						
		HIV test	*Yes*	*Yes*	*No*	*No*
Police	*Needle exchange programme*	*Drug treatment centre*	*Yes*	*No*	*Yes*	*No*
Yes	*Yes*		4	2	13	56
Yes	*No*		8	17	50	358
No	*Yes*		41	52	147	864
No	*No*		116	267	871	X

Source: M. Frischer and others, "Estimating the population prevalence of injection drug use and infection with human immunodeficiency virus among injection drug abusers in Glasgow, Scotland", *American Journal of Epidemiology,* vol. 138, No. 3 (1993), pp. 170-181.

Table 2. Capture-recapture study in Glasgow: prevalence estimates overall and by sex and age group, 1990

Group	*Known*	*Estimate of hidden*	*Total*	*Prevalence (percentage)*
All	2 866	5 628	8 494	1.4
Males	1 977	3 567	5 544	1.8
Females	889	2 349	3 238	1.0
Age group				
15-19	264	640	904	1.0
20-24	1 137	1 613	2 750	2.6
25-29	878	1 724	2 602	2.7
30-34	342	796	1 138	1.4
35 or older	245	1 273	1 518	0.6

Source: M. Frischer and others, "Estimating the population prevalence of injection drug use and infection with human immunodeficiency virus among injection drug abusers in Glasgow, Scotland", *American Journal of Epidemiology*, vol. 138, No. 3 (1993), pp. 170-181.

Log-linear models were used to estimate *x* the hidden number of injecting drug users who were not included in any of the data sources in the study. Such analysis is aimed at selecting the model that is the simplest (the one with the fewest interactions) and the best fitting to estimate the prevalence. Model selection is tested in a number of ways. The absolute goodness of fit (G^2, or deviance between the observed and the expected values) of the model approximates to a chi-square distribution, with a lower deviance implying that the model fits the data better (that is, that the observed and the expected values were closer). Rival models also can be compared using a log-likelihood ratio test (LRT) [24] for models with different degrees of freedom (for example, independent versus a model with one interaction or a model with one interaction versus a model with two interactions). Recently, two other methods (the Akaike information criterion (AIC) and the Bayesian information criterion (BIC)) have been proposed to support G^2, which can be used to compare models that have the same degrees of freedom (for example, a model with interaction between data sources S1 and S2 compared with a model with interaction between data sources S2 and S3) [25, 26].

The best-fitting model in the Glasgow example (see table 2) found positive interactions between three data sources (three two-way interactions and one three-way interaction between HIV tests, needle exchange programmes and drug treatment centres). That means that injecting drug users on one of the lists were more likely to be on the others. The police source, arrests for drug possession, was independent of the other data sources. It was estimated that there were 8,500 injecting drug users (1.35 per cent of persons aged 15-54). There were sufficient data to provide prevalence estimates by sex and age group. Those estimates, which were used extensively by local policy makers, led to an increase in needle exchange services in the city [27]. Ninety-five per cent confidence intervals were calculated around the overall estimate as 7,500-9,700 (1.2-1.5 per cent).

Confidence intervals can be estimated directly from the models based on standard equations of variance, as given by Bishop and others [10]. Alternatively, a goodness-of-fit approach, where values for the lower and upper confidence intervals are inputted into the contingency table until a difference of 3.84 (95 per cent) is found in G^2 [28, 29]; or through bootstrap methods [30]. Confidence intervals are useful statistical measures of uncertainty about the sample, based on the size of the data sources, the number of matches and the complexity of the model; however, they say nothing about the reliability of the model or about how true the estimate is.

Capture-recapture: without lists or routine data sources

The examples of capture-recapture presented above (and many studies in the literature) are based on collecting data sources with names or some form of identifier (such as initials, date of birth and sex). That is not always possible, but there are other ways of doing capture-recapture studies. An example of one of those approaches was a study conducted in Bangladesh to estimate the number of street-based sex workers in Dhaka using ethnographic fieldwork. In the study, which was carried out as part of the SHAKTI Project established by CARE Bangladesh, several categories of key informants were interviewed (sex workers, pimps, rickshaw and taxi drivers, police and local *mastans* ("toughs")) to map the dimensions of the sex-work scene (by location and time). With the help of sex workers, red cards were distributed in all known locations of the city, from late in the evening to midnight. The red card could be used for a free health check-up in one of the clinics for reproductive health and primary health care in the city. Since the cards were numbered, a sample of the number of cards distributed (that is, the capture sample (n_1)), was easily obtained. Then green cards were distributed. The green cards could also be used for free health care (n_2). The sex workers were asked only one question: whether they had received a red card earlier (m). The estimated number of street-based sex workers (N) was around 5,000; that number was derived from the data using the simple calculation presented in figure I. If the study had been extended over a series of nights with different coloured cards corresponding to different sampling days, more sophisticated "open" capture-recapture models could have been used.

A more sophisticated version of that methodology was used to estimate the number of street prostitutes in Glasgow [31]. It involved identifying how many street prostitutes were working over a period of time and noting how many were working on each night and if they had been observed on previous nights. In total, 206 women (of whom 147, or 71 per cent, were injecting drug users) were interviewed. The capture histories of the women suggested that the population remained constant at around 200 per night but that the population changed by approximately 8 per cent per week, giving an annual total of about 1,150 prostitutes.

Multiplier methods

Multiplier methods, also referred to as ratio-estimation, come in a variety of guises [5]. In essence, they have two elements in common:

(a) The benchmark. The benchmark (*B*) is a data source that captures the number of problematic drug users who have experienced a particular event, such as the number of problematic drug users who have been in treatment, who have been arrested or who have died of an overdose;

(b) The multiplier. The multiplier (*M*) is an estimate of the proportion of problematic drug users who have experienced the event recorded by the benchmark, such as the proportion of such drug users who have been in treatment, who have been arrested or who have died of an overdose. That information is usually obtained independently of the benchmark data. The inverse of that proportion is the multiplier, which is an indirect estimate of the proportion of the total population of problematic drug users represented in the benchmark data. For example, if the proportion is 10 per cent, then the multiplier is 1/0.1, or 10, and the sampling intensity is 1 in 10.

The prevalence is calculated by multiplying the benchmark by the multiplier (*B* x *M*). For example, if the benchmark is 100 and the multiplier is 10, then the prevalence is 1,000.

Theoretically, the two-sample capture-recapture mentioned above (see figures I and II) could be reduced to a multiplier method—with the first data source as the benchmark and the proportion of problematic drug users from the first data source also found in the second data source as the multiplier. However, capture-recapture involves the collection and merging of data sources that are explicitly linked together. That is even true when the capture-recapture is done without identifiers, as linkage is formed by asking the second study sample whether they were included in the first sample. In contrast, it is not a necessity in multiplier studies for the benchmark and multiplier to be collected together; it is only necessary that they refer to each other. The benchmark could be simply a number given to the researchers by a service, and the multiplier could be obtained from another study or the literature.

Nomination also has been used to obtain a multiplier [5, 32, 33]. In such a case, a sample of people (for example, injecting drug users) is asked questions about their friends or acquaintances (that is, their "nominees") who also are injecting drug users. For example, Parker, Bakx and Newcombe conducted a study with 60 injecting drug users who were asked to nominate their five closest acquaintances and say how many of them had been in treatment during the previous year. The 60 injecting drug users reported 300 other injecting drug users. After removing duplicates, that number was reduced to 170, of whom 55 were identified as being in treatment. That gave a proportion of 32.4 per cent and a multiplier of 3.1 [32]. Hartnoll and others also used the nomination technique in one of their studies in London [33].

Figure III shows a multiplier study produced by Archibald and others for Toronto, Canada [34]. Laboratory reports of HIV tests indicating injecting drug use as the reason for testing were used as the benchmark; and a survey in which injecting drug users were asked whether they had been tested for HIV in the previous year was used as the multiplier. In multiplier studies of mortality, estimates of the overdose mortality rate among injecting drug users are used as the multiplier and the number of deaths caused by opiate overdose is used as the benchmark [5, 33]. Multiplier methods have also been applied in settings in developing countries [35].

Figure III. Multiplier study based on human immunodeficiency virus (HIV) tests in Toronto

Benchmark (*B*)	Number of HIV tests by injecting drug users in 1996:[a]	4,050
Multiplier (*M*)	Proportion of injecting drug users reporting having had an HIV test in the previous year:[b] 23 per cent 1/0.23 = 4.35	
Prevalence estimate (*B* x *M*)	4,050 x 4.35 = 17,600	

[a]Based on laboratory reports.
[b]Based on community-recruited survey of injecting drug users.

Assumptions

The key assumptions of the multiplier approach are that the estimate of the multiplier (or the estimate of the proportion of those in the target population who experienced the event recorded by the benchmark) is representative and unbiased. Ideally, the estimate is obtained from a representative sample of problematic drug users and collected in the time period and in a place corresponding to the benchmark. That rarely happens. Archibald and others [34] used a multiplier from a survey of injecting drug users that had been carried out in one city and assumed that it would be the same in Toronto. Truly random representative samples of problematic drug users do not exist. The best that can be done is to recruit subjects in a way that limits any potential bias [36]. For example, it would be foolhardy to attempt to generate an unbiased estimate of the proportion of registered injecting drug users for a multiplier estimate by recruiting injecting drug users directly outside a needle exchange facility and asking how many are registered with a needle exchange programme.

In practice, it is assumed that the benchmark event is common enough and significant enough to be remembered or detected in a sample of problematic drug

users and that it is truthfully reported. It is assumed, for example, that going to treatment or being arrested for drug possession will be accurately reported by a sample of problematic drug users (and that there is no conflict of interest that may influence willingness to respond). It is also assumed that the benchmark data are accurate and complete. Unfortunately, routine data sources can be notoriously unreliable because of underreporting or incomplete data collection. In the Canadian study described above, for example, it was noted that the laboratories might undercount the number of tests carried out and that clinicians ordering tests did not always specify whether the person to be tested was an injecting drug user. Therefore, the benchmark might need to be adjusted to take into account underreporting.

Violation of one or all of the above-mentioned assumptions is possible, introducing bias into the estimation. Studies need to be critically evaluated in terms of how the multiplier was obtained and how reliable it is for problematic drug users in a specific time and place, as well as how the benchmark was obtained and how reliable it is. Confidence intervals around the estimate may give a spurious sense of precision, since they do not take into account potential bias. It would be better to make an "evidence-based judgement"—that is, to compare the findings from a single multiplier study with other estimates generated in other multiplier studies or using other methodologies to give a range of prevalence estimates. Multimethod studies spread the risk and expand the evidence base. For example, in Tolgliatti, Russian Federation, approximately 1,000 km south of Moscow, a community-recruited survey of injecting drug users was used to collect saliva to estimate HIV prevalence [37], behavioural data and a number of multipliers. The multipliers included the proportion of injecting drug users in treatment (narcology), listed in a central addict register, registered with a needle exchange programme, arrested for possession of drugs, tested for HIV and treated at a hospital for overdose. In addition, data from treatment facilities, HIV tests and police arrests are being collected for a capture-recapture study.

Advanced estimation methods

The use of some indirect estimation methods in estimating the number of problematic drug users (see the list in the beginning of the present article) is still being developed. Current examples of their use are limited mostly to settings in developed countries.

An advanced "multiplier" method, utilizing the number of events as a multiplier, has been piloted in the United States of America, in Chicago [38]. The benchmark data are numbers of events (for example, the number of times that a heroin user has turned up at a shelter or the number of heroin users arrested by the police), which are collected as systematically as possible. The multiplier is a rate (the annual rate at which a heroin user turns up at a shelter or is arrested). What is novel in that approach is that advanced statistics are used to generate an "unbiased" event rate by "reweighting" three or more potentially biased samples

(for example, asking problematic drug users in police cells, shelters or treatment facilities whether they have been arrested). There are plans to pilot the procedure in Mexico [39]; if successful, that will have important implications for the application of the procedure in other settings.

While capture-recapture methods require two or more sources of data on drug use, a method known as truncated Poisson can sometimes provide prevalence estimates from a single source of data [40]. That method can be applied when data are available in the form of counts of individuals who appear in a single data source once, twice and so on. That comprises the "raw" material. Those who are never seen fall into the zero-frequency class and are missing from the observed series of frequencies. Naturally, the total population size equals the number of persons "ever seen" plus the number of persons "never seen". If the number of unseen drug users can be estimated then, as with the capture-recapture method, the total prevalence of drug use can be found. That can be done by fitting a Poisson distribution to the complete series of count data that estimate the probability of being seen once, twice and so on, giving the sampling intensity and a prediction of the number not seen. The Poisson process assumes that the counts are random and independent of each other, which some view as restricting its application in prevalence estimation [41].

Finally, Law and others have used back-calculation methods developed in acquired immunodeficiency syndrome (AIDS) epidemiology to estimate the prevalence of heroin injecting from long-term trends in overdose deaths [42]. The "raw" materials are the long-term trends in overdose mortality deaths; estimates of the sampling intensity are generated through the modelling process, which makes assumptions over the overdose mortality rate and cessation rate to estimate trends in the incidence of heroin injecting over time. That is, an estimate of the number of injecting heroin users over time that would yield the observed trends in overdose deaths. Projecting the estimates of the incidence forward, allowing for drug cessation and mortality, will give an estimate of the cumulative prevalence of heroin-injecting. The method is still in the initial throes of development; it is unlikely that the use of the method in developing countries will be feasible in the short term, but it could prove exciting and useful if empirical data on mortality and cessation rates of injecting drug users are strengthened.

City versus country

In general, indirect estimation methods are appropriate for towns and cities where there are sufficient numbers of problematic drug users and existing data sources to allow for a viable study. As an exception to that rule, Hay, McKeganey and Hutchinson updated the estimates for Glasgow and for the rest of Scotland [43] using four sources of data on drug misuse that were available—drug treatment services, general practitioners, the police (arrests for possession of opiates or benzodiazepines) and the court system (mainly in connection with crimes such as theft). The four-sample capture-recapture method was used in most areas of

Scotland. However, in some areas there were insufficient data from general practitioners; therefore those data were combined with drug treatment data in a single "treatment" source. There were only two small areas where there were insufficient data to undertake a capture-recapture analysis; in those areas, a modification of the multiplier method was used to obtain prevalence estimates. By combining the various local estimates from 77 models in 15 health boards and 32 local councils in Scotland, an estimate for Scotland as a whole was generated: 55,800 (43,700-78,400) persons aged 15-54. Thus, the estimate was built up from separate capture-recapture studies within discrete geographical areas.

The use of capture-recapture in cities and areas in Scotland was a unique piece of work that provided credible estimates for most of Scotland (except for certain rural or unpopulated areas, where there were not enough data to run any models). That may not be practicable or possible elsewhere. (An equivalent study in England, for example, would require over 150 separate capture-recapture studies and data collection exercises.) Synthetic estimation or the multiplier indicator method aims to derive a national estimate from prevalence estimations in selected sites (called "anchor points") and indicators of problematic drug use in all sites [44]. In terms of the definition of an indirect method provided here—the "anchor points" are the "raw" material or information about a proportion of the problematic drug use population. The indicators (that is, drug seizures, arrests, drug treatment, deaths resulting from overdose and laboratory reports) are used to estimate the sampling intensity (that is, the proportion of the total population of problematic drug users within the "anchor points"). Regression equations are generated between the "anchor points" and indicators. These are then applied to geographical areas with indicators but without "anchor points" to estimate the prevalence. Summing across all the geographical areas provides a national estimate. This method is possible only after several prevalence studies in areas or cities of a country have been carried out.

Potential data sources for capture-recapture and benchmarks for multiplier studies

There is a range of data sources that could be used in prevalence estimation, either as a data source for capture-recapture or as a benchmark for a multiplier study:

Data source	*Example*
Specialized drug treatment facilities	Drug users on methadone, attending treatment agencies or in residential care
Low-threshold drug agencies	Drug users attending drop-in sites or contacted by outreach workers
Needle exchange programmes	Drug users registered in needle exchange programmes
Hospital records	Drug users treated in hospitals because of an overdose

Laboratory reports	Drug users tested for HIV, hepatitis B or hepatitis C
Police or prisons	Drug users arrested or imprisoned for drug offences; drug users arrested or imprisoned for other crimes
Probation	Drug users on probation
Social services—assessments	Drug users assessed by local social services
Shelters (hostels) for drug users	Drug users living in shelters (hostels)
Addict registers	Drug users reported to a central register
Surveys of problematic drug users	Community surveys of drug users
Deaths resulting from overdose	Number of deaths due to opiate overdose

For capture-recapture, data sources can be underestimates of the total number of cases, but each person's name and date of birth need to be accurate in order to avoid misclassification errors. For multiplier studies, however, information on an individual can be missing or wrong but the total number of persons in contact with the service needs to be accurately recorded or capable of being estimated. In addition, the data sources for capture-recapture should be carefully chosen to reduce to a minimum both dependence and heterogeneity. Finally, if the available data sources are poor, it is recommended that steps be taken to improve them for future estimation work. Collecting the data is the most time-consuming part of prevalence estimation work. The time involved in doing such work could be dramatically reduced if contributing to prevalence estimates was one of the objectives of routine data on problematic drug use.

Conclusions

Indirect estimation techniques provide a relatively cost-effective and accurate method for estimating the extent of problematic drug use (injecting drug use, the use of heroin or "crack" etc.) when compared with conventional population surveys. As such, they provide a particularly attractive option for measuring the extent of problematic drug use in settings in developing countries. Two methods in particular, capture-recapture and multiplier methods, are likely to be appropriate in developing countries. It is important that policy makers and researchers are aware of the assumptions underpinning the methods to aid interpretation and critical evaluation of the findings obtained using an indirect method. Finally, the evidence base on the prevalence of problematic drug use is sorely lacking in many countries. The number of prevalence estimation exercises needs to be increased to address that shortcoming. During that process, attention needs to be given to improving the routine data sources from which estimates are derived and to using multiple methods where possible. Taking those steps will make it easier to derive updated estimates and to improve the robustness of estimates.

References

1. "Drug use and HIV/AIDS", fact sheet of the twenty-sixth special session of the General Assembly on HIV/AIDS, New York, 25-27 June 2001.
2. P. Reuter, "Prevalence estimation and policy formulation", *Journal of Drug Issues*, No. 23, 1993, pp. 167-184.
3. Z. Sloboda and N. J. Kozel, "Understanding drug trends in the United States of America: the role of the Community Epidemiology Work Group as part of a comprehensive drug information system", *Bulletin on Narcotics* (United Nations publication), vol. LIV, Nos. 1 and 2 (2002).
4. Y. Hser and others, *Techniques for the Estimation of Illicit Drug User Prevalence: an Overview of Relevant Issues* (Washington, D.C., National Institute of Justice, 1992).
5. European Monitoring Centre for Drugs and Drug Addiction, *Estimating the Prevalence of Problem Drug Use in Europe*, Scientific Monograph Series, No. 1 (Lisbon, 1997).
6. European Monitoring Centre for Drugs and Drug Addiction, *Methodological Guidelines to Estimate the Prevalence of Problem Drug Use on the Local Level* (Lisbon, 1999).
7. I. C. Maxwell, "Methods for estimating the number of hard core drug users", *Substance Use and Misuse*, No. 35, 2000, pp. 399-420.
8. M. L. Brecht and T. D. Wickens, "Application of multiple capture methods for estimating drug use prevalence", *Journal of Drug Issues*, No. 23, 1993, pp. 229-250.
9. J. Gfroerer and M. Brodsky, "The incidence of illicit drug use in the United States, 1962-1989", *British Journal of Addiction*, vol. 87, No. 9 (1992), pp. 1345-1351.
10. Y. M. M. Bishop, S. E. Fienberg and P. W. Holland, *Discrete Multivariate Analysis: Theory and Practice* (Cambridge, Massachusetts MIT Press, 1975).
11. M. Begon, *Investigating Animal Abundance: Capture-Recapture for Biologists* (London, Edward Arnold, 1979).
12. G. A. F. Seber, *The Estimation of Animal Abundance and Related Parameters*, 2nd ed. (London, Charles W. Griffin, 1982).
13. C. Chandra Sekar and W. Edwards Deming, "On a method of estimating birth and death rates and the extent of registration", *American Statistical Association Journal*, vol. 44, March 1949, pp. 101-115.
14. S. E. Fienberg, "Bibliography on capture-recapture modelling with application to census undercount adjustment", *Survey Methodology*, No. 18, 1992, pp. 143-154.
15. E. B. Hook and R. R. Regal, "Capture-recapture methods in epidemiology: methods and limitations", *Epidemiologic Reviews*, No. 17, 1995, pp. 243-264.
16. International Working Group for Disease Monitoring and Forecasting, "Capture-recapture and multiple record systems estimation I: history and theoretical development", *American Journal of Epidemiology*, vol. 142, 1995, pp. 1047-1057.
17. T. D. Mastro and others, "Estimating the number of HIV-infected injection drug users in Bangkok: a capture-recapture method", *American Journal of Public Health*, vol. 84, No. 7 (1994), pp. 1094-1099.
18. K. Tilling and J.A.C. Sterne, "Capture recapture models including covariate effects", *American Journal of Epidemiology*, vol. 149, No. 4 (1999), pp. 392-400.
19. S. Suzman, M. G. Sirken and C. D. Cowan, "Sampling rare and elusive populations", *Science*, No. 240, 1988, pp. 991-996.

20. E. B. Hook and R. R. Regal, "The value of capture-recapture methods even for apparent exhaustive surveys: the need for adjustments for source and ascertainment intersection in attempted complete prevalence studies", *American Journal of Epidemiology*, vol. 135, No. 9 (1992), pp. 1060-1067.
21. L. J. Schouten and others, "The capture-recapture method for estimation of cancer registry completeness: a useful tool?", *International Journal of Epidemiology*, vol. 23, No. 6 (1994), pp. 1111-1116.
22. R. M. Cormack, "Problems with using capture-recapture in epidemiology: an example of a measles epidemic", *Journal of Clinical Epidemiology*, No. 52, 1999, pp. 909-914.
23. M. Frischer and others, "Estimating the population prevalence of injection drug use and infection with human immunodeficiency virus among injection drug users in Glasgow, Scotland", *American Journal of Epidemiology*, vol. 138, No. 3 (1993), pp. 170-181.
24. D. Clayton and M. Hills, *Statistical Models in Epidemiology (*Oxford, Oxford University Press, 1993).
25. Y. Sakamoto and others, *Information Criterion Statistics* (Tokyo, KTK Scientific Publishers, 1986).
26. D. Draper, "Assessment and propagation of model uncertainty", *Journal of the Royal Statistical Society*, Series B, vol. 57, No. 1 (1995), pp. 45-97.
27. L. Gruer, J. Cameron and L. Elliott, "Building a city-wide service for exchanging needles and syringes", *British Medical Journal*, No. 306, 1993, pp. 1394-1397.
28. R. M. Cormack, "Interval estimation for mark-recapture studies of closed populations", *Biometrics*, No. 48, 1992, pp. 567-576.
29. R. R. Regal and E. B. Hook, "Goodness-of-fit based confidence intervals for estimates of the size of a closed population", *Statistics in Medicine*, No. 3, 1984, pp. 287-291.
30. S. T. Buckland and P. H. Garthwaite, "Quantifying precision of mark recapture estimates using the bootstrap and related methods", *Biometrics*, vol. 47, 1991, pp. 255-268.
31. N. McKeganey and others, "Female streetworking prostitution and HIV infection in Glasgow", *British Medical Journal*, No. 305, 1992, pp. 801-804.
32. H. Parker, K. Bakx and R. Newcombe, *Living with Heroin* (Milton Keynes, United Kingdom, Open University Press, 1988).
33. R. Hartnoll and others, "Estimating the prevalence of opioid dependence", *Lancet*, 26 January 1985, pp. 203-205.
34. C. P. Archibald and others, "Estimating the size of hard-to-reach populations: a novel method using HIV testing data compared to other methods", *AIDS*, vol. 15, No. 3 (suppl.) (3 April 2001), pp. 541-548.
35. United Nations, Office for Drug Control and Crime Prevention, *Drug Abuse in Pakistan* (Vienna, 2001).
36. P. Griffiths and others, "Reaching hidden populations of drug users by privileged access interviewers: methodological and practical issues", *Addiction*, No. 88, 1993, pp. 1617-1626.
37. T. Rhodes and others, "Explosive spread and high prevalence of HIV infection among injecting drug users in Togliatti City, Russian Federation: implications for HIV prevention", *AIDS*, No. 16, 2002, pp. F25-F32.
38. R. Simeone and others, *A Plan for Estimating the Number of "Hardcore" Drug Users in the United States* (Washington, D.C., Drug Policy Research Group, Office of National Drug Control Policy, 1997).

39. E. Medina-Mora, personal communication.

40. G. Hay and F. Smit, "Estimating the number of drug injectors from needle exchange data", *Addiction Research and Theory*, No. 96, 2001, pp. 1465-1476.

41. E. B. Hook and R. R. Regal, "Validity of bernoulli census, log-linear, and truncated binomial models for correcting of underestimates in prevalence studies", *American Journal of Epidemiology*, No. 116, 1982, pp. 168-176.

42. M. Law and others, "Back-projection estimates of the number of dependent heroin users in Australia", *Addiction*, No. 96, 2001, pp. 433-443.

43. G. Hay, N. McKeganey and S. Hutchinson, *Estimating the National and Local Prevalence of Problem Drug Misuse in Scotland: Executive Report* (Edinburgh, Scottish Office, 2001).

44. M. Frischer and others, "A comparison of different methods for estimating the prevalence of problematic drug misuse in Great Britain", *Addiction*, No. 96, 2001, pp. 1465-1476.

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

The role of dynamic modelling in drug abuse epidemiology

C. ROSSI

Department of Mathematics, University of Rome "Tor Vergata", Rome

ABSTRACT

Policy makers need information to describe and understand a situation involving problematic drug use, follow trends over time, design appropriate interventions and evaluate the results of the action taken. Monitoring drug use and drug problems involves complex information, based mostly on observational data and epidemiological indicators. Dynamic models can be used to generate estimates where data are sparse or to verify hypotheses or predict trends by means of "what if" scenario analyses. The simple act of building a model forces a researcher to make explicit statements about the process being studied, which usually leads to discussion and improved insight. The models that can be used effectively in the drug field are essentially models of epidemics that describe the spread of a disease in a population in order to provide evidence for public health-oriented interventions and policies. One such model proposed recently to reflect the spread of drug use in a population is described in the present article and used to demonstrate the potential of that approach.

Keywords: drug use; epidemics; compartmental modelling; scenario analysis; drug policy.

Introduction

When defined in general terms the needs of a policy maker seem remarkably clear: it is necessary to describe and understand a situation involving problematic drug use and follow trends over time, design appropriate interventions and evaluate the results of the action taken. Monitoring drug use and drug problems involves complex information, based mostly on epidemiological indicators derived from incomplete data.

The available data are incomplete because drug abuse is generally stigmatized and hidden, and trafficking in and possession of drugs are criminal offences in most countries. Consequently, there have been varying degrees of under reporting when standard epidemiological survey techniques such as household surveys of drug abuse or direct enumeration of cases (case finding) have been used. It is

therefore necessary to develop methods that allow the extent of the phenomenon to be estimated and the dynamics to be described from observational secondary data on drug abuse that are available in various forms. Secondary data can be defined as existing statistical and documentary information that is routinely collected and readily available, such as data concerning treatment presentations, drug seizures, infectious disease indicators or drug-related deaths. Estimation techniques refer mostly to models and methodologies used to estimate the extent and the dynamics of drug abuse in a community or at the regional or national level, or both, based on various observed phenomena and information received on a target population [1, 2].

This is achieved in general by describing the components of the process and the relationships between them in a formal or quantitative way, in particular by means of mathematical models. By manipulating or experimenting with the model, conclusions can be drawn that cannot be found by direct observation of the "real" process. Mathematical models can even aid in the design and choice of appropriate responses by providing a means of integrating data from different sources, describing a process to increase understanding and simulating policy experiments that are not possible in real life [3]. The models that can be used in the drug field are essentially dynamic models of epidemics aimed at describing the spread of a disease in a community.

There are similarities between the spread of drug use, in particular the spread of use of addictive drugs such as heroin, and that of infectious diseases [4-6]. The use of drugs is communicated, obviously not as an organic agent, but as a kind of "innovative" social practice or custom, and not to everyone but only to those who, for whatever reason, are not immune (prone individuals). Once the contagious nature of drug use is accepted, the epidemiological concepts of "incidence" (the rate of new cases occurring within a certain period of time) and "prevalence" (the number of cases at a particular time) are operationally valuable in studying illegal drug use. Unfortunately, as already mentioned, the hidden population of users cannot be properly studied by standard statistical methods. Mathematical models, in particular compartmental models, are therefore useful for studying problem drug use [2, 7-10]. Such models enable prevalence and incidence to be estimated, scenario analyses to be carried out and trends to be predicted, on the basis of indirect indicators such as therapy presentations, incarcerations and so forth.

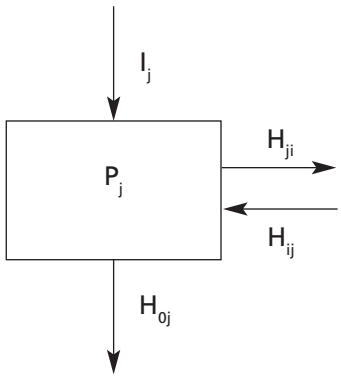
Compartmental models of epidemics

Compartmental models are a well-established and powerful mathematical tool for modelling the spread of "diseases" in a population [11]. They provide a framework in which numbers of people in different "compartments" (each one homogeneous with respect to some specified characteristic) and the relationships between such compartments, which model the dynamics of the population, can be described in mathematical terms. Two main types of model have commonly been used to describe the spread of diseases: deterministic models, expressed in terms of systems of differential equations, and stochastic models, expressed in terms of

stochastic equations or processes. Both assume that the population can be split into compartments that can be considered homogeneous with respect to a particular characteristic. Once the population has been split into compartments, it is easy to use suitable equations to describe mathematically how the size of the compartments will change over time, according to the basic hypotheses of the model describing the dynamics of the population under study. The result obtained from the model is usually the number of people in a specific compartment at a specific time (prevalence) or the number of people moving to or from, or moving to and from, a given compartment over a specified period of time (incidence).

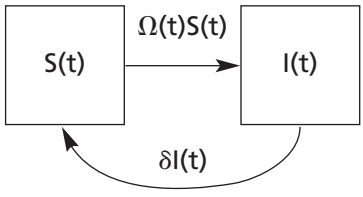
A simple general scheme of compartmental models of epidemics is shown in figure I. In the following section one such model is described and used to demonstrate the potential of the methodology in drug abuse epidemiology.

Figure I. Diagram of general compartmental models of epidemics



In general, a compartment is represented by a rectangle containing a capital letter, which represents the "level" of the compartment, that is, a state variable counting the number of individuals or the prevalence in the compartment, standardized if necessary

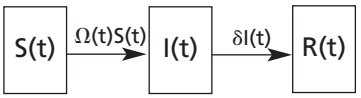
The arrows represent movement from or into the compartment. The inflows can originate either from another compartment (in which case the pair of ordered subscripts of the transition rate shown near the arrow indicate the originating and destination compartments), or from outside the system (in which case often only one subscript appears, or the originating compartment is indicated by 0). The outflow transition rates are indicated in the same way.



Two basic types of model can be used: S-I-S and S-I-R. S-I-S models can be used to represent situations in which susceptible individuals pass through a phase of infection and then recover, becoming again susceptible. S(t) is the state variable representing the susceptible population at time t, whereas I(t) is the state variable representing the infectious population.

Further transitions may take into account the possibility of death or any other relevant possibility.

Figure I (*continued)*



S-I-R models can be used to represent situations in which susceptible individuals pass through a phase of infectivity and then recover, becoming immune. S(t) is the state variable representing the susceptible population at time t, whereas I(t) is the state variable representing the infectious population and R(t) the recovered population.

Further transitions may take into account the possibility of death or any other relevant possibility.

A model for an epidemic of drug use

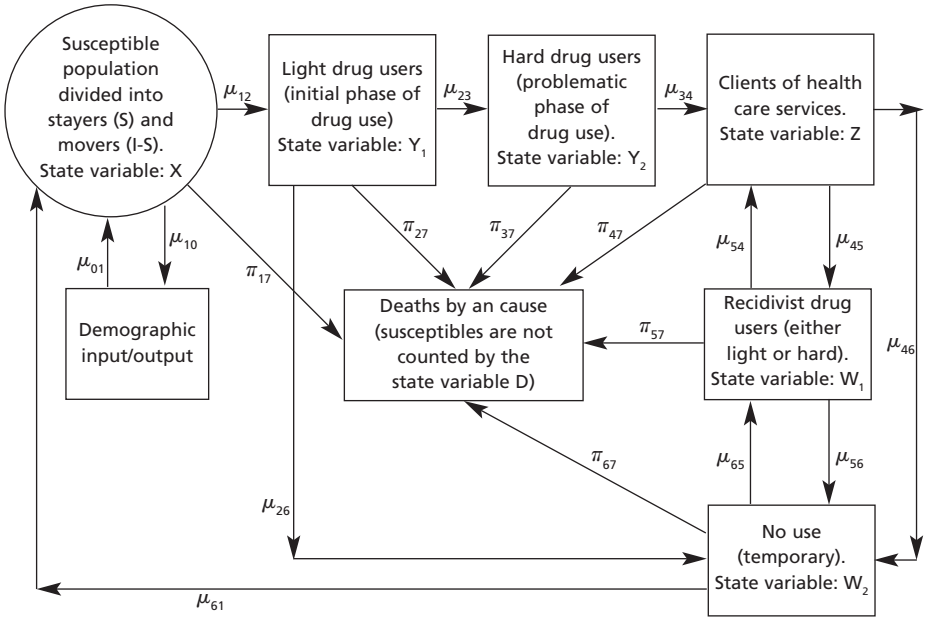
Figure II describes the main features of a model that was recently proposed [9, 10]. The model is of the "mover-stayer" type and allows for heterogeneous risk behaviour among the susceptibles. Such models consider the susceptible population as subdivided into two groups: the group of "stayers", that is, the group of individuals who, owing to their "prudent" behaviour, are considered not to be at risk of "infection" (these models are suitable for making scenario analyses in order to assess the impact of various proportions of vaccinated persons on the probability of extinction of a given epidemic) and the group of "movers", who are individuals at risk of infection. Owing to the interactions between infectious individuals (in the present case these are the problem drug users who are also pushers)[1] and the susceptibles, or owing to the pressure of the black market on the susceptibles, some of these may pass to the drug user compartments and begin a "drug user career". Like the model proposed by Behrens and others [7], the present model comprises two different stages of hidden drug use. The first (light use) stage is the initial, or non-problematic, stage of drug use, then light drug users can either stop using drugs or pass to hard drug use (or death). The other arrows represent the other possible transitions in a drug user career.

Using the flow chart (figure II), it is straightforward to write the equations of the model either in the form of deterministic (continuous or discrete) equations or in the form of stochastic (continuous or discrete) equations. Only the very general features of the model and policy-relevant conclusions are considered in the present article; the discrete stochastic equations are reported elsewhere [10]. The model is a combination of an S-I-S and an S-I-R model, with the second

[1]The surveys conducted among military conscripts reported in the annual report on the state of the drug problem in Italy for the year 1999 revealed that the two most-mentioned reasons for drug use were curiosity (over 40 per cent) and peer group pressure (over 30 per cent).

dominating. Thus, according to the general mathematical theory of epidemic models reported in Iannelli [11], it follows that an epidemic phase followed by an endemic phase is always obtained.

Figure II. A simple "mover-stayer" model for an epidemic of drug use



Using the simulation procedure [9, 10], a scenario analysis can be carried out to evaluate the influence of the various parameters on the course of the epidemic. Scenario analyses can also be used to evaluate numerically the impact of the various kinds of intervention (primary and secondary prevention, harm reduction and so forth).The results of the various scenario analyses [9, 10] show that the bigger the core group of "at risk" susceptibles, the faster the spread of the epidemic and the higher the prevalence and incidence curves, while the influence of the parameter that measures the pressure of the black market appears to be less important.

Some general qualitative results

Some qualitative analytical results have been obtained and scenario analyses carried out on the basis of various hypotheses on the parameter values and have been discussed elsewhere [9, 10]. The simulation procedure used to obtain the scenario analyses is written in S-plus. All the parameters can be modified at the beginning of each run and the total simulation time, which is measured in weeks, can be

chosen. The standard output comprises graphs of the prevalence curves in each compartment and of the incidence curves of major interest. The output also comprises the curves representing the main macroepidemic indicators such as the epidemic/endemic indicator, which is negative only during the epidemic phase, and the impact indicators related to primary and secondary prevention interventions and harm reduction, measuring the expected difference of the onset incidence (transitions from X to Y1) when the intervention is implemented compared with the basic situation (no intervention).

The numerical results reported in graphic form in figures III-IX enable the main features of the model to be presented. These results were obtained on the basis of the parameter values estimated for Italy. The first "epidemic wave" (figure III) relates to the incidence of light use, which generates the prevalence of light use, and (figure IV) by a transformation that produces a deformed wave. There follows a second incidence wave (figure V), of hard drug use, generating the second prevalence wave (figure VI), of hard drug use and so on for other compartments, such as therapy, recidivist use and so forth, which are not presented in figures III-IX. Incidence and prevalence per one million inhabitants are reported.

Figure III. Incidence curve from susceptibles to light users

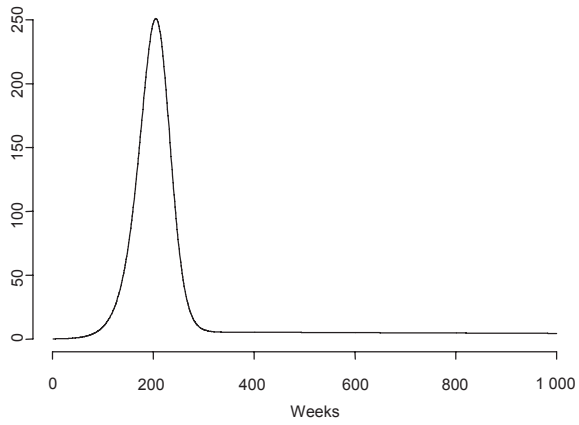


Figure IV. Prevalence curve of light users

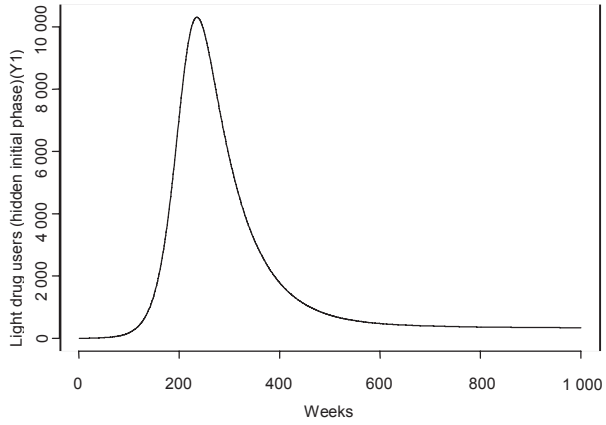


Figure V. Incidence curve from light users to hard users

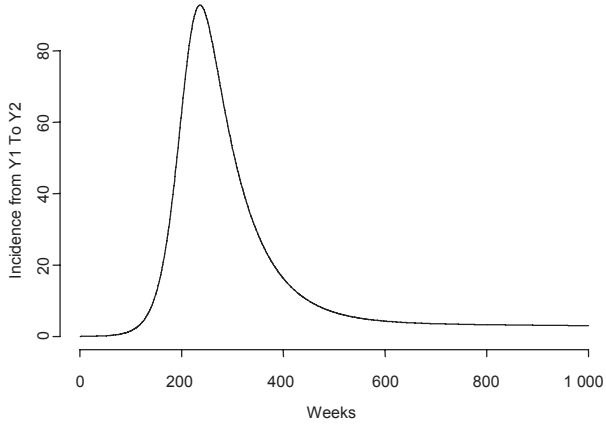


Figure VI. Prevalence curve of hard users

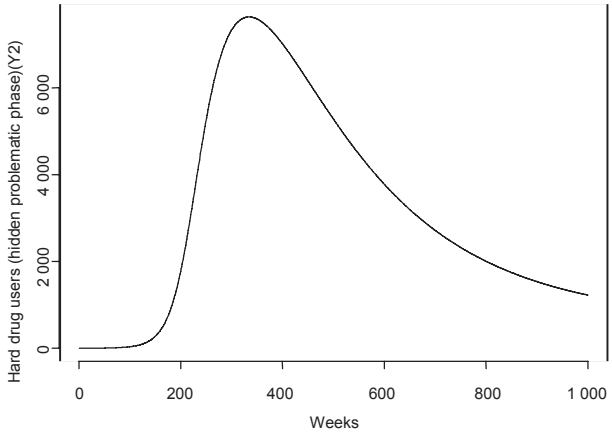


Figure VII. Epidemic endemic indicator

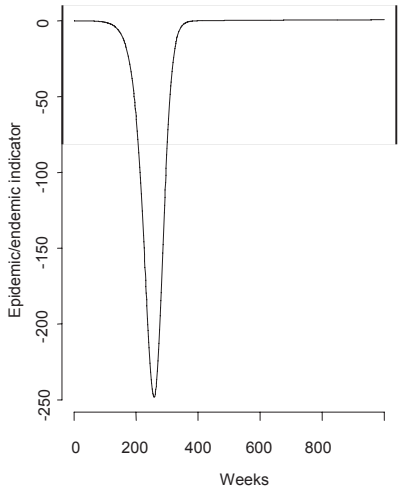


Figure VIII. Expected impact of a primary prevention intervention

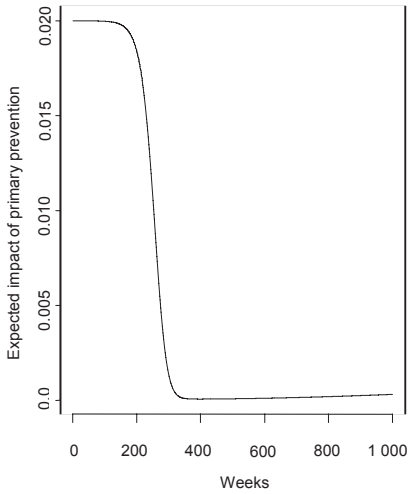
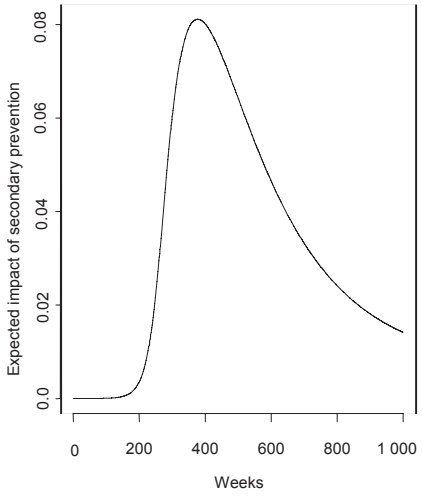


Figure IX. Expected impact of a secondary prevention intervention



The last three curves, which represent the behaviours of relevant macroindicators, enable some observations to be made that may be useful to policy makers. In figure VII, the behaviour of the epidemic/endemic indicator shows that the epidemic spreads fast and is then followed by the endemic phase, starting about seven years after the outbreak; in figure VIII, primary prevention interventions have a higher impact on the epidemic at the beginning, then the impact decreases rapidly; in figure IX, as might be expected, secondary prevention interventions have a higher impact when the prevalence of drug use is higher, but the efficiency decreases during the endemic phase.

These qualitative observations may be extended to any epidemic model similar to that presented, for any disease, and may therefore also apply to any infectious disease epidemic among drug users and, in particular, among injecting drug users, which spreads both by sharing needles and by sexual intercourse, such as the human immunodeficiency virus (HIV), hepatitis B and hepatitis C.

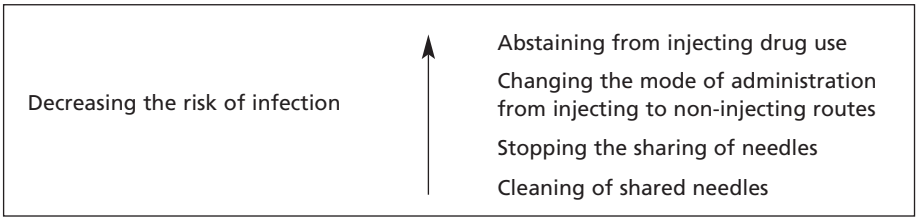
Conclusion

In the present article, a compartmental model for reproducing an epidemic of problematic drug use has been presented. Qualitative and quantitative analyses (by means of simulation) have been presented to demonstrate the potential usefulness of the model for decision makers. The quantitative analyses ("what if" scenarios) have been developed on the basis of knowledge gained from the heroin epidemic of the last 20 years in Italy, which enabled several parameter values to be derived from various epidemiological and statistical studies that were already available. The results obtained, from both the qualitative and the quantitative analyses, apply to any kind of epidemic of drug use, even involving new drugs, at least from a qualitative point of view. In particular, the qualitative evaluation of the effectiveness of different types of intervention over the course of the epidemic is valid for any epidemic of problematic drug use; for example, the indications arising from the behaviour of the epidemic/endemic indicator, which changes abruptly to coincide with changes in the efficacy of policy interventions (notably prevention), can be considered. This suggests that policy makers should monitor such parameters, which are not directly observable, using indirect estimates, such as onset incidence estimates obtained by various estimation methods or survey data [12, 13]. In fact, the epidemic part of the indicator is essentially based on the incidence of new use. Consequently, decision makers should put in place real-time monitoring systems to estimate onset incidence.

The qualitative results obtained also enable the possible impact of harm-reduction interventions to be indirectly evaluated. Harm reduction is a public health approach that gives priority to reducing the adverse consequences of drug use for the individual, the community and society, rather than to eliminating drug use or ensuring abstinence. Although the aim is still to reduce drug use in general, emphasis is placed on the elimination of the potential harmful effects of drug-taking behaviour.

With regard to HIV or hepatitis, for example, a harm-reduction strategy will first try to reduce the transmission of infections by means of the cleaning of shared injecting equipment or the cessation of that sharing, rather than by means of promoting abstention from drug use (figure X). Achieving those immediate and realistic goals is usually viewed as a first step towards risk-free use; abstinence may be considered a final aim. It follows that a harm-reduction intervention aimed at reducing the transmission of infection can be viewed as a secondary intervention with respect to the epidemic of drug use, but as a primary prevention intervention with respect to infection.

Figure X. Hierarchy of intervention goals for reducing the transmission of HIV



This kind of intervention therefore has a high impact with respect to the onset incidence of injecting drug use when the prevalence of injecting drug users is high, but it can really prevent infectious diseases only if applied at the very beginning of the epidemic of injecting drug use. It can therefore be concluded that harm-reduction measures aimed at preventing the spread of infectious diseases among injecting drug users should be taken as early as possible. The empirical evidence from the various European Union countries where such measures have been taken confirms these general qualitative results obtained a priori on the basis of the dynamic model [14].

Finally, the model demonstrates that the spread of the infectious disease epidemic among injecting drug users is related for the most part to the hidden part of the drug user's career (compartments Y1 and Y2 of figure II). This suggests that the objective of interventions should be to reduce the duration of that period, which is called the latency period. According to analyses carried out in several European Union countries, the latency period appears to be remarkably similar in different cities, with a median of between four and six years and an average of between five and seven years [12, 13]. However, the true lapse appears to be much longer than this for young drug users, which implies a need for specific interventions aimed at young drug users in order to reduce their latency period and prevent the spread of infectious diseases.

References

1. United Nations, Office for Drug Control and Crime Prevention, *Drug Abuse Rapid Situation Assessment and Responses* (Vienna, 1999).
2. C. Rossi, "Using available data with dynamic models to estimate the extent of problem drug use", *Modelling Drug Use: Methods to Quantify and Understand Hidden Processes*, Scientific Monograph Series, No. 6, C. Rossi, ed. (European Monitoring Centre for Drugs and Drug Addiction, 2001), pp. 67-90.
3. D. A. Behrens and others, "Optimal control of drug epidemics: prevent and treat—but not at the same time?", *Management Science*, vol. 46, No. 3 (2000), pp. 333-347.
4. R. de Alarcón, "The spread of heroin abuse in a community", *Bulletin on Narcotics* (United Nations publication), vol. XXI, No. 3 (1969), pp. 17-22.
5. D. R. Mackintosh and G. T. Stewart, "A mathematical model of a heroin epidemic: implications for control policies", *Journal of Epidemiology and Community Health*, vol. 33, 1979, pp. 299-304.

6. L. G. Hunt and C. D. Chambers, *The Heroin Epidemics* (New York, Spectrum Publications, 1976).
7. D. A. Behrens and others, "A dynamic model of drug initiation: implications for treatment and drug control", *Mathematical Biosciences*, vol. 159, No. 1 (1999), pp. 1-20.
8. L. Billard and P. W. A. Dayananda, "Drug addiction-pushers generated from addicts", *Biometrical Journal*, vol. 35, 1993, pp. 227-244.
9. C. Rossi, "A mover-stayer type model for epidemics of problematic drug use", *Bulletin on Narcotics* (United Nations publication), vol. LIII, Nos. 1 and 2 (2001), pp. 39-64.
10. C. Rossi, "Operational models for epidemics of problematic drug use: the mover-stayer approach to heterogeneity", *Socio-Economic Planning Sciences*, forthcoming.
11. M. Iannelli, "The mathematical description of epidemics: some basic models and problems", *Mathematical Aspects of Human Diseases*, G. Da Prato, ed. (Pisa, Giardini, 1992), pp. 11-32.
12. European Monitoring Centre for Drugs and Drug Addiction, "Pilot project to estimate time trends and incidence of problem drug use in the European Union", final report, Lisbon, 1999.
13. European Monitoring Centre for Drugs and Drug Addiction, "Study on the incidence of problem drug use and latency time to treatment in the European Union", final report, Lisbon, 2000.
14. M. Owel, "Harm reduction within six European national drug policies: a necessary option or a deficiency?", Master of Arts dissertation, University of Ghent, 2000.

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

Drug use, drug abuse and heterogeneity

M. D. GLANTZ and J. D. COLLIVER*

Division of Epidemiology, Services and Prevention Research, National Institute on Drug Abuse, National Institutes of Health, Bethesda, Maryland, United States of America

ABSTRACT

Concepts and criteria for drug abuse have significant impacts on research and the development of interventions. The empirically supported differentiation of drug use and drug abuse has led to important gains in understanding the aetiology (etiology) of drug abuse. The need for further distinctions that reflect the heterogeneity of drug abuse is explored. The need for more sophisticated models of drug abuse is discussed.

Keywords: drug use (drug abuse); drug dependence; epidemiology; etiology; heterogeneity; developmental psychopathology.

Drug use and drug abuse

At one point, every drug abuser was a non-user, an abstainer who initiated drug use, later became an escalating user and eventually became an abuser. Every drug abuser follows this progression. While there is considerable variation in drug abuse (the use of different drugs, the quantity and frequency of use etc.), there are also critical commonalities. Drug abuse intrinsically involves the compulsive use of illicit psychoactive substances despite significant potential or already manifest harm.

This commonality suggests to some that drug abuse and drug abusers are fundamentally homogeneous in terms of the primary inherent determining factors. While acknowledging the variable patterns of drug abuse behaviour, this perspective assumes that drug abuse by different individuals is essentially alike in terms of the critical characteristics. These might be any of a number of possible factors, for example, a common underlying neurobiological process, comparable environmental conditions or equivalent behavioural conditioning. Taking this a step further, some believe that there is one basic pathway or trajectory leading to drug abuse and one set of predisposing or risk factors for vulnerability to drug use as well as abuse.

*The authors would like to thank Zili Sloboda for her helpful comments on an earlier version of the present article.

However, drug abuse is not just a lot of drug use. Obviously, drug use and drug abuse are not independent discrete phenomena. Nevertheless, despite the unquestionable relation of use to abuse, research shows that vulnerability to use seems to be largely distinct from vulnerability to abuse. Further, the quantitative and qualitative differences between drug use and drug abuse are so critical that for heuristic and research purposes, they are in many ways fundamentally distinct phenomena despite their quantitative and behavioural association. This might be thought of as the contrast between a problem behaviour and a psychopathological condition.*

During the late 1960s and 1970s, there was a dramatic growth of illegal drug use in the United States of America and in many other countries. In response, there was a major increase in research attempting to describe and explain both drug abuse in general and the new "drug culture" phenomenon, which involved the use of drugs by population subgroups not previously associated with illegal drugs. Whereas the majority of early investigations had typically studied addicts in treatment and other heavily drug-dependent individuals, the subjects of the new explosion of research were often students in their second year of college (aged 18-19). A common research design of the time tested single factor models of drug abuse by attempting to find a correlation between assessments of a trait or characteristic and a minimal measure of drug involvement, both assessed at a single point in time. Most of the subjects in those studies had relatively limited experience with illegal drug use, which resulted in relatively little variability on the measure of drug experience. Perhaps as a means of achieving variability, the drug abuse measures usually categorized the subjects either as abstainers and experimenters who had used an illegal drug only a few times or as abusers, a category including most other degrees and forms of drug involvement. The published investigations often reported that the researchers' theory of drug abuse had been supported by research because a correlation had been found between the trait being studied and the drug abuse that it "caused". Not surprisingly, the studies contributed little to the understanding of drug abuse. Fortunately, some researchers used more sophisticated longitudinal research designs and took a more sophisticated approach to measuring drug involvement.

In the early 1990s, under the co-sponsorship of the National Institute on Drug Abuse (National Institutes of Health) and the American Psychological

*Use of an illegal drug automatically qualifies as "abuse" in some nomenclatures because it involves criminal activity. This is in contrast to categorizations of alcohol use. Alcohol is legally available although there are restrictions such as age limitations or prohibitions on use while driving. Alcohol "abuse" usually refers to pathological and/or problematic use of alcohol. For the purposes of this article, the term "drug use" is used analogously to the term "alcohol use". "Drug use" refers to experimentation or low-frequency, typically irregular use of illicit drugs that has not (or at least not yet) resulted in significant negative consequences. "Drug abuse" refers to regular and/or compulsive use of illicit drugs. While such a distinction may seem artificial to some, it serves to distinguish people on the basis of whether their illicit drug use has or has not become a significant feature of their lifestyle and whether it is likely to have a psychopathological character. This is not intended to discount the significance of engaging in an illegal behaviour as opposed to one that is legal nor is it intended to minimize the potential danger of even a single use of illicit drugs, particularly under certain circumstances. The intention is to focus on the characteristics and differences underlying the behaviours of use and abuse and to facilitate the discussion of ideas.

Association, drug abuse researchers conducting major longitudinal studies of adolescents were asked to conduct special analyses to explore vulnerability to drug abuse, focusing specifically on those factors and patterns associated with the transition from drug use to abuse. More sophisticated differentiations of use and abuse were encouraged. Despite the variations in the studies, their designs and their subjects, the research projects consensually found that the predisposing or risk factors for drug use were different than those for drug abuse. In addition, the sets of risk factors identified by the diverse studies followed a single basic pattern [1]. Research in the past decade has further corroborated and extended those findings.

The consensus that emerges from the research is that drug use (including initiation and lower levels of involvement) appears to be more a result of social, peer and environmental factors, while drug abuse (including higher levels of drug involvement, drug dependence and drug use disorders) appears to be more a function of biological psychological and psychiatric factors. The risk factors for drug use do not necessarily predict the transition to drug abuse, addiction or disorder. This categorical differentiation of drug use from drug abuse is not just semantic; it highlights different paths and different phenomena. Risk factors for drug use and abuse include the following:

Risk factors for initiation/low involvement

Some "problem behaviours", "bad" conduct

Friends with problem behaviour

Friends involved with drugs, particularly friends with "problem behaviours"; peer influences encouraging and facilitating drug involvement

Drug availability

Unconventionality; rebelliousness

Low involvement with traditional value-oriented institutions (i.e. family, religious institutions, school)

Poor academic achievement

Poor quality relations with (and attachment to) parents and/or having parents with problems; poor parenting

Risk factors for abuse/addiction

Multiple risk factors

Multiple and/or severe problem behaviours

Early age of onset or initiation of drug use

High frequency use of drugs

Parental substance abuse and/or antisocial behaviour, sibling drug abuse

Family history of psychopathology

Severe family disruption and/or dysfunction, including significantly problematic divorce

Neurobiologic dysfunction related factors

Some psychopathologies: antisocial personality disorder, conduct disorder and criminal behaviour; affective disorders, including depression, bipolar disorder, anxiety disorders (in particular, post-traumatic stress disorder)

Severe childhood conduct/behaviour problems involving aggressivity, acting out and a high childhood activity level

Multiple psychopathologies, that is, more than one childhood psychiatric disorder, in particular a combination of an internalizing and externalizing disorder

Emotional/behavioural arousal self-regulation difficulties (possibly including sensation seeking), impulsivity and attention deficit disorder (called attention deficit/hyperactivity disorder in the United States) if coupled with aggressivity/conduct disorder; deficits in executive cognitive function and affect regulation may also be predisposing conditions

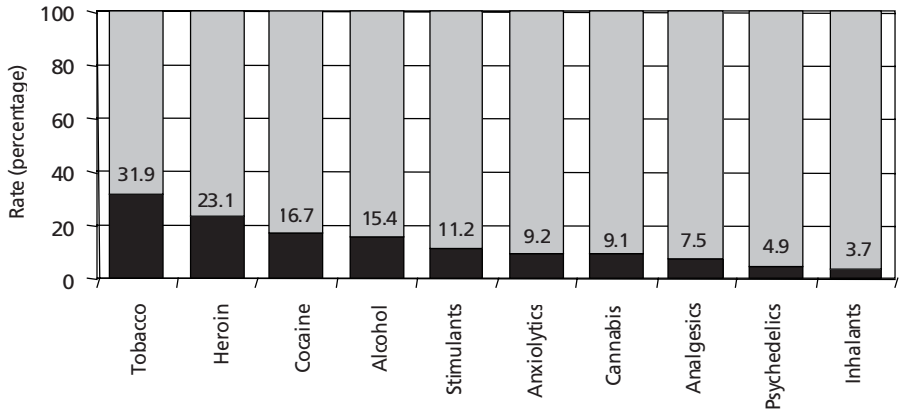
Traumatic experiences, including childhood physical and sexual abuse, particularly if the experience results in post-traumatic stress disorder

Other factors: generally poor function and difficulties in coping, social isolation, interpersonal difficulties

This means that only some individuals who use drugs will be at high risk for making the transition to drug abuse or addiction. In fact, data from general population surveys reveal that only a minority of illicit drug users meet the criteria for dependence. The National Comorbidity Survey conducted in the period 1990-1992 indicated that in the United States 51.0 per cent of the population 15-54 years of age reported illicit drug use at some time in their lives but only 7.5 per cent of the population had a history of drug dependence based on meeting diagnostic criteria at some time in their lives [2].

Not only has research consistently shown that the large majority of those who use drugs are not drug-dependent abusers, but it is also clear that in the large majority of cases, such drug use does not escalate to drug abuse. Given the difference in rates of dependence among users of different substances, it is not surprising that the latency of onset of dependence varies by substance. In a recent re-analysis of National Comorbidity Survey data [3], survival analysis was used to track the development of dependence on cannabis, cocaine and alcohol. Focusing on the transition from use to dependence, cocaine dependence (reported for 16.7 per cent of users) exhibited rapid development, 5-6 per cent of users having met the criteria within one year after initial use and approximately 11 per cent (the majority of eventual cases) within three years. Within 10 years, approximately 15 per cent of cocaine users had developed dependence. Cannabis dependence (reported for 9.1 per cent of users) developed less rapidly, less than 2 per cent of users having met criteria within one year and approximately 5 per cent in three years; rates were close to asymptotic within 10 years.

Rates of lifetime dependence among lifetime users 15-44 years of age



Source: J. C. Anthony, L. A. Warner and R. C. Kessler, "Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: basic findings from the National Comorbidity Survey", *Experimental and Clinical Psychopharmocology,* vol. 2, 1994, pp. 244-268.

For alcohol (reported for 15.4 per cent of all users), dependence developed at rates similar to those for cannabis in the first four years (reaching 5-6 per cent). Unlike cannabis dependence, alcohol dependence continued to increase at steady rates up to around seven years, when it reached around 11 per cent, and to continue to increase, though at somewhat diminishing rates, 15-20 years after initial use and beyond [3].

A relatively large number of people engage in some drug use, including drug use that involves experience with substances that have a high abuse liability (i.e. substances shown to be the most addictive). In the United States, for example, 39 per cent of the population 12 years of age and older (some 87 million persons) have tried illicit drugs at some time in their lives, 67 per cent have used cigarettes and 81 per cent have used alcohol (2000 estimates from the National Household Survey on Drug Abuse) [4]. Among the illicit drugs, lifetime exposure to cannabis is reported by 34 per cent, 11 per cent have used cocaine and 1.2 per cent have used heroin. Initiation of cigarette, alcohol and illicit drug use frequently occurs at an early age. The 2001 Monitoring the Future Study revealed that, in the United States, 53 per cent of students in grade 10 (aged 15-16) had used cigarettes in their lifetime, 70 per cent had consumed alcohol, 48 per cent had "been drunk" on alcohol, 46 per cent had used an illicit drug, 40 per cent had used cannabis, 6 per cent had been exposed to cocaine and 1.7 per cent had used heroin [5].

The prevalence rates vary considerably across countries. Paralleling the estimates for students 15-16 years old in the United States, the 1999 European School Survey Project on Alcohol and Other Drugs [6] reported that, among the 30 participating countries, the percentage of students reporting any cigarette smoking in their lifetime varied from 50 per cent in Cyprus to 86 per cent in Greenland

(Denmark); lifetime consumption of alcohol ranged from 68 per cent in the former Yugoslav Republic of Macedonia to 98 per cent in Greece and the Czech Republic; lifetime rates of having been drunk ranged from 32 per cent in Cyprus to 89 per cent in Denmark; use of any illicit drug at least once ranged from 3 per cent in Cyprus to 36 per cent in the United Kingdom of Great Britain and Northern Ireland; and lifetime cannabis use varied from 2 per cent in Cyprus to 34 per cent in the United Kingdom. The wide ranges may reflect the influences of several factors. One obvious factor is availability; rates of perceived availability of cannabis, for example, ranged from 5 per cent in Romania to 57 per cent in Denmark. The most recent cannabis availability rate in the United States was much higher—77 per cent, corresponding to the higher rates of lifetime exposure (40 per cent).

Drug use, abuse and prevention

Obviously, it is common for adolescents and young adults to have experience with drug and alcohol use; one important question is why more drug use does not escalate to drug abuse and dependence. There appear to be two probably complementary reasons why most drug users do not become drug abusers. The influence of protective factors may be the first reason. It is very likely that no single factor "causes" drug abuse. Evidence indicates that drug abuse develops from the interaction of multiple influences, including biological, psychological and social/environmental factors. In addition, models that consider only risk factors ignore an important set of influences in the system of factors determining risk for drug abuse. Protective factors also play a critical role and they appear to work in opposition to risk factors. If the right protective factors are present at the right time and if they are sufficient in influence, they may nullify the predispositional influence of even potent risk factors. Research on protective factors provides support for interventions that attempt to increase these factors as an approach to prevention.

Secondly, examination of the risk factors associated with drug abuse and addiction shows that for the most part, these factors are powerful but, fortunately, relatively uncommon, at least at higher levels. This suggests that there are significant individual differences between those who make the transition from drug use to abuse and those who do not. Those at higher risk for drug abuse are likely to be influenced by both drug use and drug abuse predisposing factors. That provides encouragement for developing targeted interventions for children and adolescents showing evidence of drug abuse risk factors to supplement universal prevention programmes. It also suggests an explanation for those studies reporting findings that universal prevention interventions benefit even individuals at high risk.

Many prevention programmes target drug use initiation as a prevention strategy. The logic that someone who never uses drugs can never become a drug abuser seems irrefutable. This may be sufficient for those who are primarily at risk only for drug use. Even if they do have some drug involvement, there is a good chance it will be mitigated by the preventive intervention. Alternatively, children

and adolescents who are subject to abuse risk factors may benefit—but benefit much less—from the same prevention programmes because a substantial part of the influence towards drug involvement is not ameliorated. Even before any drug experience, those who are more vulnerable to future drug abuse are likely to already be on different paths because they are also influenced by the drug abuse risk factors.

Further, prevention programmes that target drug use initiation usually do not include an intervention for those who are already past the point of initiation and already have some drug involvement and/or some experience with illegal drugs or with other (usually illicit) psychoactive substances such as alcohol and/or nicotine (whose purchase and/or use is prohibited for most adolescents). As discussed above, this includes a large percentage of adolescents and young adults and is likely to include a high percentage of those at higher risk for drug abuse. Incorporating escalation intervention components into prevention programmes may significantly increase their ultimate effectiveness.

It is also important to remember that drug use has consequences that in turn become predispositional influences on subsequent behaviour. It is likely, for example, that drug use impairs an adolescent's social and psychological development, making them more vulnerable to maladaptive behaviour and other problems, including more drug use [7]. Curtailing or even stopping drug abuse will not restore a person to his or her pre-abuse state. Drug abuse is not just the use of drugs that got out of control and that, if brought back under control, will leave no lasting consequences. A person who stops drug use will not simply become the same person with the same development and coping resources that they would have had if they had not used drugs. Heavier drug users whose use has impaired their maturation will not, for example, suddenly catch up on the development of problem-solving, coping and other skills [8]. For those who are vulnerable to abuse or who are already beginning to escalate their use, this is even more the case. Successful drug abuse prevention programmes for vulnerable adolescents may also need to help those individuals mitigate the impaired development and functional maturation resulting from their drug involvement.

The dynamic interaction between risk and protective factors occurs in the context of the individual's development. In other words, etiologic factors are not merely antecedents exerting influence in a simple linear cause-effect fashion. Individual and environmental influences interact at any given developmental stage of the child or adolescent to lead them to certain behaviours. Those behaviours have consequences and the consequences then become influences on subsequent interactions and behaviours. For example, a young child with a difficult temperament being raised by parents who are not able to help the child to develop self-regulation may then have a more difficult time making friends and doing well in school. In turn, this may predispose the child to associate with peers who engage in delinquent behaviour, creating an opportunity for ready availability of cigarettes and alcohol and encouragement for their use [9, 10]. A more extensive discussion of a developmental conceptualization of drug abuse aetiology can be found in Blackson [11] and Glantz [12].

Many models of drug abuse aetiology assume that there is a single primary pathway or trajectory leading to drug abuse. Many versions of the single trajectory perspective tacitly assume that the aetiology of drug abuse is a linear continuum, a single trajectory that could potentially end in drug abuse for most people who start down the path, that is, individuals who are drug users. Individual differences in vulnerability are assumed to be minimal; that is, given sufficient drug use experience, most people will become drug abusers. Single trajectory concepts typically minimize the role of pre-drug-use differences as determinants of later drug use/abuse.

Conceptualizations of drug abuse aetiology that are based primarily on a single risk factor or even a set of risk factors, even conceptualizations that include the role of protective factors, usually assume that the primary etiological factors are relatively unchanging over the course of the aetiology of drug abuse and that most of the determination of the individual's course of drug behaviour is exerted at the beginning of the chain of behaviour. They are "ballistic" models that suggest that the aetiology of drug abuse is much like the firing of a cannonball. Once fired, the trajectory of the cannonball is largely predetermined and changes in course are possible only with extreme external influences. The best hope for stopping the cannonball from striking the point at the end of its trajectory is to prevent it from being fired in the first place. While this is the case for cannonballs, it is not true for a complex course of human behaviour.

Aetiology is not a straight line course but more like a spiralling line with a feedback loop that keeps reorienting before it moves again, possibly in a new direction [13]. Every consequence becomes a "cause" influencing the next step, and all of those influences are complex interactions taking place in the context of development. In more severe instances, the course of normal development may be impaired or redirected. Furthermore, changes in trajectory for an individual can occur at any point because of the influence of strong positive or negative factors in the environment. Risk factor models conceptualizing drug abuse aetiology as a ballistic event ignore both the developmental nature of the aetiology and the developmental and environmental contexts in which it usually occurs.

The best theoretical formulation to account for the development of drug abuse appears to be an interactive risk and protective factor model that takes into account individuals' developmental stage and processes. That type of aetiology model is most useful when viewed with the perspective of patterns of factors and the results of their interaction [14]. It places more emphasis on the outcome of the interactions of risk and resilience and individual and environmental factors rather than on the specific original factors themselves. Adding in a developmental framework that views the course of the patterns over time supports a pattern or path model. Some patterns will be more common and, as the patterns involve interactions of factors over time, the metaphor of "paths" is particularly helpful.

Kazdin [15, p. 180] has provided a brief definition of developmental psychopathology as "the study of clinical dysfunction in the context of maturational and developmental processes". (For more comprehensive expositions of developmental psychopathology, see Cicchetti [16], Cicchetti and Rogosch [17] and Luthar

and others [18].) This perspective is particularly helpful in understanding drug abuse aetiology, especially because most of the formative steps in the aetiology of drug abuse typically occur during the periods of child and adolescent development [19].

Such a developmental psychopathology model of interacting factors implies that more than one pattern may lead to drug abuse. Evidence indicates that there is not a single path or combination of factors that leads to drug abuse; instead, there are numerous possible etiological patterns. Some paths, however, are more common than others. Evidence indicates that one particularly common pathway to drug abuse may be through a path involving conduct problems while another may be through a childhood psychopathology pathway. It is probably the case that a small number of paths or patterns will account for the drug abuse aetiology of the majority of individuals [20, 21].

Differentiating drug use from abuse is only a part of a larger issue of the heterogeneity of drug abuse. Just as it is important to recognize the qualitative differences of use compared with abuse, other distinctions may also be important. There is no question that there are variations in a number of dimensions of drug use/abuse. There is diversity in the drugs and combinations of drugs used, the routes of administration, the quantity and frequency patterns of use, the "career" stage of abuse, the age when use and abuse began, whether there is a comorbid psychiatric condition, what the consequences are, whether there is a physical dependence etc. Clearly, drug use/abuse is highly heterogeneous.

Concepts, models and research on drug abuse, however, do not always reflect this heterogeneity. For example, research measurements often view drug abuse as a dichotomous variable with few or no variations and as a final outcome, a static endpoint. Individual differences in drug abuse and related factors are often minimized in data analysis and are often considered to be "noise", measurement error to be controlled for. While some variability probably reflects differences in less significant factors, some of the variability in drug abuse is likely to be related to crucial distinctions (see Zucker, Fitzgerald and Moses [22] for a discussion of this and associated issues of heterogeneity related to alcoholism). Even if there are critical factors common to all cases of drug abuse and related disorders (for example, the involvement of particular neurological processes), the variations may denote significantly different subtypes or patterns. To use an analogy, the general category of "cancers" includes a large number of relatively distinct subtypes. While there are some significant core commonalities among the different types of cancer, there are also critical differences that distinguish the subtypes. Failure to consider the different clinical entities represented by the nosology of cancers would result in a limited understanding and ability to treat the different conditions. The same principle applies to drug abuse and the variable characteristics associated with it.

The question is, which if any of the characteristics relate to critical differences in drug involvement and which are relatively non-critical variations? There is no consensus among clinicians or researchers. The research data point to the probability that some variations may relate to critical heterogeneities. Often, however,

findings indicate some variations are more associated with degrees of severity without providing information about whether there is a qualitative difference involved. Identifying a common characteristic is helpful but it falls far short of identifying the set of fundamentally distinguishing characteristics of a phenotype. A big part of the problem is that it is not possible to determine whether a particular variation relates to a critical defining characteristic of drug abuse when there is no agreed upon set of critical defining characteristics of drug abuse. This issue is often referred to as the phenotype question.

The phenotype question

The concept of phenotype is prominent in genetics research, where the goal is often to relate a particular "genotype", a particular genetic constellation, to a particular "phenotype", the observable traits or characteristics of an organism. A phenotype is not, however, a direct manifestation of the genotype. The phenotype results from the genotype as it manifests and develops in a particular environment. Differences in environmental influences and experiences are the reason that, for example, even though identical (monozygotic) twins have identical deoxyribonucleic acid sequences (DNA), they have similar but not identical fingerprints and people who know them well can usually tell them apart even as infants. The relationship of the genotype to the phenotype varies with different traits. Eye colour, for example, is highly related to specific gene characteristics, while other phenotypes, particularly those that are more behavioural in nature, may be less directly related. For some traits and behaviours, there may not be a specific genotype. In other cases, a particular genotype may be less directly related to a phenotype but is more strongly associated with an intermediary, an "endophenotype", which is an underlying biological characteristic or mechanism.

Obviously, determining a particular genotype-phenotype relationship requires a clearly distinct and preferably consensually accepted definition of both the genotype and the phenotype. In some cases, it can be difficult to formulate a consensus of a reliably identifiable, appropriately inclusive and concretely defined phenotype. The greater the heterogeneity of the phenotypic characteristic, the more difficult the problem. Not surprisingly, this has been an issue for genetic research related to drug abuse.

These are not esoteric theoretical questions. Defining a phenotype for drug abuse is an operational necessity for investigations of the relationship of genetic characteristics to drug abuse, and there is no simply defined drug abuse phenotype. Many researchers have used an individual's meeting the diagnostic criteria of the International Classification of Disease (ICD) or the Diagnostic and Statistical Manual (DSM) for a drug abuse disorder as the phenotype while others have relied on other criteria. Even when these types of standardized assessments and categorizations are used, questions remain. For example:

(a) Should the criteria for the phenotype be drug-specific?

(b) Is there a predispositional endophenotypic intermediary that must be identified in order to determine the genetic contribution to drug abuse and perhaps to understand vulnerability to drug abuse?

(c) Is the phenotype of drug abuse too narrow and should it be broadened to include other forms of substance abuse, other antisocial behaviours or other psychopathological characteristics?

(d) As individuals go through different stages of drug involvement, are different characterizations necessary to correctly identify drug abuse? To what extent is it necessary to modify the characterizations necessary to correctly identify drug abuse in order to take into account the age, developmental level and life circumstances of the individual?

(e) Is the behavioural drug involvement pattern (defined by such factors as age of onset, rate of escalation, relapse history, experience of and tolerance for negative consequences and quantity and frequency of use) a more useful phenotype?

(f) Does using a phenotype defined by an individual's meeting the criteria for a DSM drug abuse disorder exclude significant alternative drug abuse manifestations?

(g) Is a more useful phenotype of drug abuse defined by an individual's reaching some ultimate common condition or process such as a particular neurochemical or neuroanatomical equifinal state?

(h) Is a currently abstinent drug abuser still appropriately classified as a drug abuser; is this classification appropriate even if they never relapse?

(i) Should drug abuse be thought of as having multiple phenotypes?

(j) Should the basic model of a drug abuse phenotype be based on the characteristics of a disease? A psychopathology? An antisocial behaviour?

There has been general recognition of the importance and lack of resolution of the phenotype issue in drug abuse genetics research. There has been less attention and concern in many other areas of drug abuse research. Some researchers have considered the issue to be little more than a semantic disagreement, but the way that drug abuse is defined has significant pragmatic and heuristic consequences: for example, using "drug use" as the phenotype for drug abuse obscured critical differences. By differentiating drug use and abuse, researchers were able to determine that different factors, in fact different paths, led to one versus the other. The importance of formulating an empirically based phenotype for application in other areas of drug abuse research besides genetic research is at least as great.

Asking the question "What is the phenotype of drug abuse?" is asking "What are the critical defining characteristics of drug abuse?". Each different phenotype conceptualization focuses on different aspects and models of drug abuse and sets different criteria for the determination of drug abuse. An important issue is whether it is appropriate to use different defining criteria depending on the

question being asked. In some cases this appears to be necessary. Some research questions seem to call for different phenotype definitions because they focus on different aspects or stages of drug abuse: for example, comparisons of drug abuse among adolescents versus adults. Other circumstances seem to require still other phenotype definitions: for example, identifications of marijuana abusers versus cocaine abusers.

The general principle of using different phenotypes in relation to different questions seems reasonable to some, in that it responds to the variations in population and behaviour. This leaves the question of whether a phenotype must have some core characteristics in order to be a drug abuse phenotype. Consider the phenotype of skin cancer. Different diagnostic criteria are applied for different stages and variations of the illness and different populations may have different risks and manifestations. Certain environmental factors and behaviours may change the likelihood and course of the disorder and therefore different prevention and treatment strategies may be most effective depending on a variety of factors. Nevertheless, there is a core concept of defining criteria that denote skin cancer and those criteria are involved in all concepts, forms and interventions for skin cancer. Whether there is more heuristic benefit in formulating one phenotype for drug abuse varying by stage and variation or in developing multiple phenotypes is unknown. Ultimately, the resolution must be driven by empirical findings.

The encouragement for the recognition and exploration of the heterogeneity of drug abuse is not intended to minimize the commonalities inherent in drug abuse. It is obviously the case that each instance of drug abuse is not unique. The issue is that identifying critical characteristics common to all instances of drug abuse has not been accomplished. Even for a single drug-abusing individual, the defining characteristics of their involvement with drugs changes over time. If the cardinal characteristics of a phenomenon cannot be specified, it is difficult to understand and develop the most effective interventions for that phenomenon. It is essential to determine which characteristics of drug abuse are universal, which are common, which are particular to specific subgroups and which are idiosyncratic. However, even seemingly simple classifications can present considerable challenges.

To use a mundane example, it would be difficult to research the nature of desserts without first determining their cardinal characteristics. While everyone knows what "dessert" is, developing a classification that would facilitate research and understanding requires some organized conceptualization. To illustrate, it is not possible to formulate a single recipe for all types of desserts, for example, none of the ingredients in a fruit salad are found in a chocolate cake. A particular response to the food would help to discriminate desserts from other foods. Most (and for the purposes of the discussion, all) desserts are sweet. This could be agreed to be a universal characteristic of desserts even though the experience of sweetness is not achieved by the same ingredient (although the sweetening agent is usually sugar in some form or a related substance). Interestingly, this does not resolve the question of whether it is the taste of sweetness or the particular component (sugar) that is the defining characteristic of desserts. This is similar to the

controversy over whether the characteristics of drugs interacting with biological systems or some more behavioural or subjective phenomenon "cause" or at least determine drug abuse. Further, not all sweet foods and not all foods containing sugar (sweet and sour chicken, chutney, barbecued pork, etc.) are desserts. Nevertheless, identifying an almost universal characteristic (sweetness) and an almost universal component creating that characteristic (sugar or a sugar substitute) would make research on desserts practical and more productive. It would not, however, identify a dessert phenotype.

Breaking the general category of desserts into distinctive major subtypes—creating a nosology of desserts—would be very helpful, especially if distinguishing critical characteristics of each subtype could be identified. Public consensus alone would probably not be adequate. Research would have to be done to identify some underlying principles or critical characteristics, in order to differentiate the subtypes into relatively exclusive categories. Classification would not be easy. For example, establishing the categories "pastry", "ice cream", "pudding", "candy" and "fruit desserts" would lead to such questions as whether fruit pies are a separate category, a form of pastry, or a fruit dessert. However, if a categorization were developed that described the "real" subtypes, a prototype recipe for each subtype might be possible and that would greatly facilitate research on desserts. Without this classification step, the heuristic potential of dessert research would be limited. This is approximately the current status of drug abuse research.

The discussion of formulating a phenotype and related nosology for desserts is facetious and intended only as a non-controversial example. Unfortunately, in the area of behavioural research, not having a phenotype and classification of subtypes for many behaviours has been a serious obstacle. For example, the field of resilience research has not successfully developed either a phenotype or a classification for resilience; even the identification of a universal characteristic has proven to be elusive. For this reason, at least in part, the extensive research done on resilience has not been nearly as productive as once expected. (See Glantz and Sloboda [23] and Kaplan [24] for in-depth discussions of resilience.)

While a considerable understanding of compulsive psychoactive drug use, despite negative consequences (the universal characteristic of drug abuse), has been achieved and a great deal of knowledge about drugs of abuse (the universal component) has been ascertained, new quantum steps forward in progress probably require identifying the subtypes or patterns of drug abuse. This is the next frontier for drug abuse research. In all likelihood, it is a pivotal accomplishment that will be the foundation of future critical progress.

References

1. M. D. Glantz and R. Pickens, eds., *Vulnerability to Drug Abuse* (Washington, D.C., American Psychological Association, 1992).
2. J. C. Anthony, L. A. Warner and R. C. Kessler, "Comparative epidemiology of dependence on tobacco, alcohol, controlled substances and inhalants: basic findings from

the National Comorbidity Survey", *Experimental and Clinical Psychopharmacology,* vol. 2, 1994, pp. 244-268.

3. F. A. Wagner and J. C. Anthony, "From first drug use to drug dependence: developmental periods of risk for dependence upon marijuana, cocaine and alcohol", *Neuropsychopharmacology,* vol. 26, No. 4 (2002), pp. 479-488.
4. United States of America, Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, *Summary of Findings from the 2000 National Household Survey on Drug Abuse* (Rockville, Maryland, Substance Abuse and Mental Health Services Administration, 2001).
5. L. D. Johnston, P. M. O'Malley and J. G. Bachman, *Monitoring the Future National Survey Results on Drug Use, 1975-2001; Volume I: Secondary School Students* (Bethesda, Maryland, National Institute on Drug Abuse, 2002).
6. B. Hibell and others, *The 1999 ESPAD Report: Alcohol and Other Drug Use among Students in 30 European Countries* (Stockholm, Swedish Council for Information on Alcohol and Other Drugs and Pompidou Group of the Council of Europe, 2000).
7. M. D. Glantz and A. I. Leshner, "Drug abuse and developmental psychopathology", *Development and Psychopathology,* vol. 12, No. 4 (2000), pp. 795-814.
8. L. Chassin, S. C. Pitts and C. DeLucia, "The relation of adolescent substance use to young adult autonomy, positive activity involvement, and perceived competence", *Development and Psychopathology,* vol. 11, No. 4 (1999), pp. 915-932.
9. J. A. Andrews and others, "The influence of peers on young adult substance use", Health Psychology, vol. 21, No. 4 (2002), pp. 349-357.
10. T. A. Wills and S. D. Cleary, "Peer and adolescent substance use among 6th-9th graders: latent growth analyses of influence versus selection mechanisms", *Health Psychology,* vol. 18, No. 5 (1999), pp. 453-463.
11. T. C. Blackson, "Temperament: a salient correlate of risk factors for alcohol and drug abuse", *Drug and Alcohol Dependence,* vol. 36, No. 3 (1994), pp. 205-214.
12. M. D. Glantz, "A developmental psychopathology model of drug abuse vulnerability", *Vulnerability to Drug Abuse,* M. D. Glantz and R. Pickens, eds. (Washington, D.C., American Psychological Association, 1992), pp. 389-418.
13. M. D. Glantz and others, "The origins of drug abuse: mapping the paths", *Drug Abuse: Origins and Interventions,* M. D. Glantz and C. R. Hartel, eds. (Washington, D.C., American Psychological Association, 1999), pp. 3-45.
14. D. Cicchetti and S. L. Toth, "Transactional ecological systems in developmental psychopathology", *Developmental Psychopathology: Perspectives on Adjustment, Risk and Disorder,* S. S. Luthar and others, eds. (Cambridge, Cambridge University Press, 1997), pp. 317-349.
15. A. E. Kazdin, "Developmental psychopathology: current research, issues and directions", *American Psychologist,* vol. 44, No. 2 (1989), pp. 180-187.
16. D. Cicchetti, "The emergence of developmental psychopathology", *Child Development,* No. 55, 1984, pp. 1-7.
17. D. Cicchetti and F. A. Rogosch, "Conceptual and methodological issues in developmental psychopathology research", *Handbook of Research Methods in Clinical Psychology,* 2nd ed., P. C. Kendall, J. N. Butcher, and G. N. Holmbeck, eds. (New York, Wiley, 1999), pp. 433-465.

18. S. S. Luthar and others, eds., *Developmental Psychopathology: Perspectives on Adjustment, Risk and Disorder* (Cambridge, Cambridge University Press, 1997).

19. D. Cicchetti, "A developmental psychopathology perspective on drug abuse", *Drug Abuse: Origins and Interventions,* M. D. Glantz and C. R. Hartel, eds. (Washington, D.C., American Psychological Association, 1999), pp. 97-117.

20. M. D. Glantz, "The impact of treatment for childhood psychopathology on subsequent drug abuse: an introduction to an important question", *Journal of Consulting and Clinical Psychology,* in press.

21. R. C. Kessler and others, "Mental-substance comorbidities in the ICPE surveys", *Psychiatria Fennica,* No. 32, 2001, pp. 62-80.

22. R. A. Zucker, H. E. Fitzgerald and H. D. Moses, "Emergence of alcohol problems and the several alcoholisms: a developmental perspective on etiologic theory and life course trajectory", *Developmental Psychopathology: Vol. 2. Risk, Disorder and Adaptation,* D. Cicchetti and D. J. Cohen, eds. (New York, Wiley, 1995), pp. 677-711.

23. M. D. Glantz and Z. Sloboda, "Analysis and reconceptualization of resilience", *Resilience and Development: Positive Life Adaptations,* M. D. Glantz and J. L. Johnson, eds. (New York, Kluwer Academic/Plenum Press, 1999), pp. 109-126.

24. H. Kaplan, "Toward an understanding of resilience: a critical review of definitions and models", *Resilience and Development: Positive Life Adaptations,* M. D. Glantz and J. L. Johnson, eds. (New York, Kluwer Academic/Plenum Press, 1999), pp. 17-83.

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

The role of rapid assessment methods in drug use epidemiology

C. FITCH, T. RHODES, V. HOPE, G. V. STIMSON and A. RENTON

Centre for Research on Drugs and Health Behaviour,
Department of Social Science and Medicine,
Faculty of Medicine, Imperial College, University of London, London

ABSTRACT

"Rapid assessment" methods have the potential to generate important public health information that can be used with monitoring, surveillance and other available data systems to develop intervention programmes. That potential is now the subject of discussion within the field of substance use. First emerging in the early 1990s, the last three years have seen the approach endorsed as an expedient method for profiling drug-related problems, mobilizing human immunodeficiency virus (HIV) prevention efforts, initiating policy change and service reorientation and, more recently, as a potential component of "second-generation" surveillance. In the present article, the authors consider the role of rapid assessment in generating knowledge for public health action and, more specifically, the relationship between rapid approaches and the cornerstone of public health science, epidemiology. Drawing on case studies and examples, the authors propose that rapid assessment is best understood not as a new method, but as a practical convergence of existing research and intervention traditions (including field epidemiology and anthropology). Six roles for rapid assessment in either informing or complementing drug use epidemiology are outlined: (a) in information poor situations; (b) as a means of informing ongoing monitoring; (c) optimizing community involvement; (d) informing quantitative research; (e) questioning quantitative research; and (f) as a tool for responding to emerging health problems.

Keywords: rapid assessment; epidemiology; drug use; surveillance; developing countries; intervention; anthropology; community development.

Introduction

"Rapid assessment" methods have the potential to generate important public health information that can be used with monitoring, surveillance and other available data systems to develop intervention programmes. Drawing on both qualitative and quantitative research techniques, such assessments are typically undertaken in situations where data are needed quickly, where local resource

constraints rule out more conventional research approaches and where agencies require information to develop, monitor and evaluate intervention programmes [1].

The potential of rapid assessment is now the subject of discussion within the substance use field. First emerging in the early 1990s, the last three years have seen the approach utilized by national, international and United Nations agencies as an expedient method for profiling drug-related problems [2-4], mobilizing human immunodeficiency virus (HIV) prevention efforts [5-8], initiating policy change and service reorientation [9-11], and, more recently, as a potential component of "second-generation" surveillance [12-13]. This interest has manifested itself in the method's global diffusion (with documented use in 70 countries) [14] and in attempts to systematize and refine the methodology through the production of various field guidelines [5, 6, 8, 9].

However, despite that activity, there is much that is not known about rapid assessment. Elsewhere previously unavailable data are presented on the known number, composition, and outcomes of rapid assessments on substance use [14]. In the present article, the role of rapid assessment in generating knowledge for public health action is reviewed and, more specifically, consideration is given to the relationship between rapid approaches and the cornerstone of public health science, epidemiology.

To achieve this, the article addresses two key questions: *(a)* what is rapid assessment?; and *(b)* what is its potential role within drug use epidemiology?

Rapid assessment: method

Rapid assessment is not a new method. The approach has evolved as a practical convergence of different research and intervention traditions. Among the most important of these are "field epidemiology", anthropological research, community development and formative research.

Field epidemiology: rapidity and response

Field epidemiology represents a core methodological foundation for rapid assessment in the field of substance use. Typically employed in the investigation of outbreaks of infectious disease [15, 16] and in the formulation of post-disaster intervention programmes [17, 18], Bammer and others propose that the process can be encapsulated in 10 steps [19]:

(a) Determine that there is a problem;

(b) Confirm the problem is correctly described;

(c) Define what constitutes having a problem and what data need to be collected to respond effectively and estimate the numbers affected;

(d) Orientate the data in terms of time, place and person;

(e) Determine the risk factors and who is at risk of developing the problem;

(f) Develop a hypothesis explaining why the problem has emerged and test this by appropriate statistical methods;

(g) Compare the hypothesis with the established facts;

(h) Plan a more systematic study that includes data collection to guide action;

(i) Prepare a written report with options for action;

(j) Plan and execute control and prevention measures.

Such steps reflect the basic processes and principles underpinning a rapid assessment (although, as mentioned below, they are typically supplemented with additional techniques).

Firstly, owing to external time and resource pressures, the speed of assessment and response is necessarily accelerated [17]. Within the field of drug use the continued occurrence of HIV epidemics among injecting drug users has been a key driving force in the emergence of rapid assessment. As Friedman and Des Jarlais pointed out over a decade ago, once HIV prevalence among intravenous drug users (IDUs) reaches 10 per cent, without intervention it can then surpass 40-50 per cent in between one and four years [11]. Acting quickly to minimize risk in order to prevent harm is a fundamental feature of the logic of the public health movement and rapid assessment approaches have been promoted as one of the methods of achieving those aims. Consequently, both field epidemiology and rapid assessment studies aim to adopt pragmatic and streamlined methodological designs. In practical terms, this means that an ongoing trade-off is made between analytical depth and rapidity. In both approaches, priority is typically given to establishing a basic description of the problem in order to allow initial decisions about response development to be made, rather than implementing sophisticated (and more lengthy) methodological designs. Significantly, this does not mean that conventional research techniques, protocols and standards are discarded. Instead, those basic techniques are adapted and reformulated in the light of the often difficult and "resource-poor" settings in which assessments are conducted.

Secondly, the central aim of both field epidemiology and rapid assessment approaches is to initiate intervention responses, rather than for knowledge generation itself. As Glass and Noji note [17]:

> "The success of the ... investigation ... can be measured directly by how rapidly data collected and analysed can identify prevention strategies, and how effectively these strategies can then be implemented by decision makers to direct relief and decrease ongoing mortality."

Arguably, rapid assessment's utility is therefore evaluated both on its ability to inform effective response development as well as on methodological rigour.

Recent evaluation studies of rapid assessment in the field of substance use underscore the method's ability to quickly produce systematic data that can be used in response development. A review of 12 different methodological guidelines

on "drug" rapid assessments indicate that most assessments should take no longer than 12-16 weeks, while data retrospectively collected on 83 rapid assessments suggest an average study length of 18 weeks [14]. Typically employing multi-method approaches, the majority of studies are also reported as resulting in the development of a range of intervention responses [14].

Anthropology: meaning and context

Field epidemiology represents a core methodological foundation for rapid assessment in the field of substance use. However, as with any method, the approach also has limitations. As recent discussion acknowledges [20, 21], one limitation lies in the extent to which measures of the social and cultural contexts in which substance use, risk behaviour and HIV infection take place can be captured in epidemiological assessments. The importance of "situating" epidemiological research in the social and cultural context of health behaviours and problems has been met through numerous initiatives [22, 23]. However, within rapid assessment, two significant contributions have been made by anthropologists interested in developing "short-cut" approaches to accumulating social knowledge.

Firstly, the work of the Latin American Centre at the University of California at Los Angeles in the early 1980s began making strides towards adapting traditional ethnographic research techniques [24-26]. Working at a time when increasing interest among policy makers in capturing the social context of health behaviours was tempered by frustration with the time-scale required to produce that knowledge, researchers adapted techniques such as observation, in-depth interviews, mapping and group discussion so that they could be readily applied. Importantly, those methods did not aim to achieve the analytical depth of intensive anthropological studies, but traded this off against the rapid production of basic situational descriptions.

Secondly, the same period witnessed the development of various field manuals and guidelines on how to undertake rapid assessments [27-33]. Rapid assessments had previously been relatively uncodified and frequently reflected the disciplinary interests and skills of the investigator. The development of these simple "road maps" [34] to social knowledge provided a reference point for assessment design and allowed assessments to be undertaken in areas where a lack of trained social scientists had meant that opportunities to profile and respond to health problems had either not been fully exploited or were based upon conventional survey methodologies. The simplified research protocols contained within the guidelines for rapid assessment procedures allowed such work to take place and arguably contributed to an element of technology transfer.

Community development and formative research: participation and action

Rapid assessment has also evolved as a mode of "community development" and "community diagnosis". Popularized in the late 1970s, those approaches

emphasize that the efficacy of assessment of welfare problems, and of the interventions developed in response, not only depend on knowledge of local situations and community practices, but also require the active participation of local communities in the research and intervention process. "Rapid rural appraisal" provides an example [35]. Here, rapid assessment is envisaged as a mode of community organization and change, and not merely "top-down" research promoted by "outside" experts.

Rapid assessment has also featured as part of the development of local interventions through programmes of "formative research" and "needs assessment". The aim of such research is to identify the interests, attributes and needs of different populations within a community (primarily those people likely to use the intervention, but also other groups with a stake or concern in the response) and to develop interventions that both meet those needs and are culturally acceptable and feasible. Importantly, such programmes are set up either prior to the design or implementation of an intervention or they are used to refine and fine-tune the intervention while it is being conducted.

Emergence in the substance use field

As described above, rapid assessment approaches represent a convergence of experience and techniques derived from different research and intervention traditions. Taken together, those approaches emphasize the importance of:

- Rapidity
- Pragmatism, simplicity and adaptation
- Using multiple methods and data sources
- Hypothesis testing, but remaining open to new discoveries (induction)
- Understanding individual behaviour in terms of local meaning and context
- Optimizing community participation
- Undertaking research that leads to local action and intervention

However, it is important to note that, while rapid assessment draws on various disciplines, the resulting composite does not represent a superior "meta-methodology" that overcomes and supersedes the limitations of individual disciplines. Instead, rapid assessment acts less as a substitute for in-depth research and more as a tool for generating basic knowledge and initiating prevention activity.

The role of rapid assessment

Currently, rapid assessment has a low profile among public health and scientific audiences. This has been a consequence of the approach prioritizing the generation of knowledge for local action, rather than for publication in peer-reviewed

journals. The present authors argue that rapid assessment should be given a more central role in public health data collection and in this section outline six potential contributions it could make.

In "information-poor" situations

In the absence of data from previous research or surveillance systems, rapid assessment can be used to generate descriptions of the extent and nature of local substance use. One example of this is provided by a series of rapid assessments conducted in India in 1998 [36]. Despite few existing data sources and governmental denial of the problem of IDU, researchers in five Indian cities undertook rapid assessments with the dual aims of bringing attention to the problem and also of highlighting the resources necessary to respond to it. Significantly, that assessment did not represent a "one-off" and instead provided a base from which further research was undertaken, with rapid assessments being undertaken in 1999 in the same five cities (plus five additional locations) with a view to filling gaps in a now growing knowledge base [36].

Informing ongoing monitoring

Rapid assessment can make an important contribution to drug monitoring systems. Monitoring systems draw on the routine collection of data to provide an ongoing description of trends in health behaviours and their adverse health consequences. However, such monitoring can fail to deliver information of practical relevance. Routine systems for the monitoring of substance use generally identify "new" drug problems long after they first emerge. Usually located in drug treatment, medical care or criminal justice agencies, they only cover people who make contact (who are a selected minority of drug users) and there is a considerable time lapse between new drug trends and related agency contact. The use of fixed data fields, while essential for the routinization of data collection and analysis, can also mean that subtle changes in drug use, as well as new forms, can be missed.

Periodic rapid assessment studies may overcome some of these shortfallings. Through multi-method assessments conducted with users and substance use professionals, descriptions can be generated of current drug use situations, contributing to the revision of data collection procedures, providing an early warning of drug problems on the horizon and locating trends within the broader socioeconomic context. Although similar systems are in operation in countries such as Australia [37], they currently do not go beyond sentinel surveys and existing data analysis.

Involving the local community

Community endorsement is key to effective research and intervention programmes. Work undertaken in Madras, India, in 1998 provides an example of how rapid assessments can initiate and optimize local participation. The assessment

emphasized the importance of not only soliciting local opinion, but also formalizing this in a Community Advisory Body, which would meet regularly to help plan, undertake and act on the results of the research. The body consisted of 20 members, including, a Catholic priest, a community worker, a pharmacist, an elected political representative, a leader of the Ambedkar Movement, a representative of the fishermen's association and representatives of the community council. The community board was assisted by a small group of technical experts in the field of drug use and acquired immunodeficiency syndrome (AIDS) and the decision-making process was seen entirely as negotiations around priority issues for assessment and intervention [38]. Although it was recognized that such collaboration required careful facilitation, the assessment was used not only as a vehicle for data collection, but also as a focal point for incorporating multisectoral input and action.

Informing quantitative research

Rapid assessments aim to complement and inform longer-term research design. An example of this is the use of such assessments within the WHO Phase II Drug Injecting Study. The WHO Drug Injecting Study is a comparative study of drug injecting behaviour and health and social consequences in 14 developing countries and countries with economies in transition. The aim of the study is to produce information on drug injecting behaviours that can inform regional and national policy and practice and data from phase I have resulted in important epidemiological, prevention and policy advances [39].

Importantly, the study has employed a design that links rapid assessment and epidemiological survey approaches. This has involved rapid assessments in collecting both behavioural data through qualitative methods and describing the context in which research was conducted and the use of survey methodology to provide descriptions of drug use behaviours in populations of injectors both in and out of contact with treatment agencies. Data from the rapid assessment have been used to enhance the design of locally relevant questions for inclusion in the cross-sectional survey, inform sampling design (including possible recruitment locations), win the support of influential professional and community leaders and provide data based on which epidemiological analysis could explain key findings.

Questioning quantitative research

Rapid assessment can also be used to question and explore the findings of quantitative research. As noted earlier, assessments can be undertaken that seek to "unpack" some of the social and cultural factors underpinning statistical trends in substance use. One example of this has been the Rapid Assessment Response and Evaluation (RARE) programme conducted in 11 metropolitan cities of the United States of America between 1999 and 2001 [40]. The catalyst for the study was data from routine surveillance indicating that HIV infection rates in each city were

disproportionate among ethnic minority groups. However, there were no clear indications why this was the case. In response, each site employed a combination of geo-social mapping techniques (to physically describe the layout of drug "hotspots"), interviews with users and cultural experts, extended observations (often over 24-hour-long periods) and street interviews to explain the cultural and social reasons for the statistical relationship [41] and on the basis of these to identify appropriate interventions. This adopted an inductive approach to hypothesis testing, where key research questions were identified prior to assessment but were both refined as data were collected and those questions were supplemented by additional questions as new discoveries were made.

Responding to emerging health problems

Finally, although rapid assessment can play important roles in relation to conventional epidemiological research, this approach also represents an important tool in its own right for profiling and responding to rapidly emerging public health problems.

An example of rapid assessment being employed to quickly mobilize resources to assess and respond to emerging drug use problems is provided by work undertaken by Care Bangladesh. In December 1998, an IDU harm-reduction strategy was launched in Dhaka. The programme encountered significant risk behaviours among IDUs prompting concern about the potential for wide-scale spread of HIV infection among drug injectors in Dhaka and elsewhere in Bangladesh [42]. Consequently, in January 1999, rapid assessments were undertaken in Dhaka, six other cities in Bangladesh and five smaller areas with the aim of documenting drug using behaviours and establishing how limited prevention resources could be best allocated. Completed over a period of six weeks, the assessments employed interviews with 82 substance use professionals, 113 "formal" user interviews, several hundred "informal" user interviews and targeted observation studies. They revealed that the potential for HIV diffusion among IDUs was pronounced in the northern region of Bangladesh and that there was particular need in the city of Rajshahi (in which resources were subsequently concentrated). This example underscores not only the rapidity with which rapid assessments can be mobilized, but also the ability of assessments to produce comparative data across multiple sites for use in intervention development decisions.

A further example of intervention development also underscores the importance of accepting the process of rapid assessment as the beginning of the response itself. Rather than envisioning a chain of activities of which intervention development is the final link, rapid assessments emphasize the importance of introducing interventions at any stage of the assessment if there is sufficient evidence for action. Conducted in Sydney, Australia, in 2000 the project aimed to strengthen the capacity of service providers to better respond to the health needs of IDUs from non-English-speaking backgrounds, with particular emphasis on prevention of hepatitis C virus [43]. During the project it was found in focus groups that Spanish-speaking injectors had limited knowledge about the routes of

transmission of HIV and hepatitis C virus. Consequently, in the week following the focus group, the research team organized a number of health education sessions with that population of injectors.

Conclusion

In the present article, it has been suggested that rapid assessment provides a means of undertaking multi-method research in time- and resource-poor situations. Representing a practical convergence of different research and intervention traditions, rapid assessments do not provide a substitute for long-term or in-depth studies, but instead offer a useful tool for establishing basic descriptions of substance use problems and initiating intervention development. Furthermore, six roles have been described in which the method is directly complementary to conventional epidemiological research.

It is important to recognize, however, that rapid assessments also have their limitations. Simple research questions, relatively small sample sizes, the potential biases inherent in becoming both observer researcher and interventionist actor and an ongoing trade-off between rapidity and analytical depth mean that assessments have to walk a careful line between being expedient, pragmatic and informative and being simply "quick, practical—and wrong" [44].

However, perhaps the greatest constraint currently placed upon rapid assessment studies is their low profile among scientific and public health audiences. The consequence of the approach prioritizing the generation of knowledge for local action, rather than for publication in peer reviewed journals, compounded by the relatively few platforms in which such accounts could be published, means that rapid assessment faces a trinity of obstacles: countering the assumption that the approach's low public profile is somehow representative of its public health contribution; avoiding the danger of an inward-looking discipline, where practitioners unknowingly replicate the previous endeavours of others and fail to learn from past experience; and developing a coherent account of its emergence and development in the substance use field. To overcome those obstacles, rapid assessment practitioners face what is perhaps their greatest challenge: introducing a culture of learning, reflection and discussion into a methodology primarily premised on rapidity and pragmatism.

References

1. L. Manderson, *Population and Reproductive Health Programmes*: *Applying Rapid Anthropological Assessment Procedures* (New York, United Nations Population Fund, 1996).
2. H. Riper, "Quick Scan: recreational drug use, clubbing and prevention", unpublished document, 2001.
3. F. Trautmann, "Rapid assessment on nature and extent of substance use among asylum seekers and refugees", personal communication, 2000.

4. F. Trautmann, "Rapid assessment on nature and extent of substance use among asylum seekers and refugees", unpublished document, 2000.
5. J. Howard and others, *The Rapid Assessment and Response Guide on Psychoactive Substance Use and Especially Vulnerable Young People* (EVYP-RAR), draft for field-testing (Geneva, World Health Organization, 1998).
6. G. V. Stimson, C. Fitch and T. Rhodes, *The Rapid Assessment and Response Guide on Injecting Drug Use*, draft for field-testing (Geneva, World Health Organization, 1998).
7. T. Rhodes and others, "Rapid assessment, injecting drug use, and public health", *Lancet*, vol. 354, 1999, pp. 65-68.
8. T. Rhodes, C. Fitch and G. V. Stimson, *The Rapid Assessment and Response Guide on Psychoactive Substance Use and Sexual Risk Behaviour* (Geneva, World Health Organization, 2002).
9. United Nations International Drug Control Programme, *Guidelines for the Development and Implementation of Drug Abuse Rapid Situation Assessments* (Vienna, 1998).
10. R. N. Needle and others, "Rapid assessment, response and evaluation (RARE): a public health strategy to reduce the impact of communities", National HIV Prevention Conference, 29 August-1 September 1999, Abstract No. 525.
11. D. C. Des Jarlais and S. R. Friedman, "HIV among drug injectors: the epidemic and the response", AIDS Care, vol. 3, 1991, pp. 239-250.
12. Tobi Saidel, "Behavioural surveillance surveys and rapid assessments: complementarities in addressing data needs for HIV epidemics among IDUs in Asia", Twelfth International Harm Reduction Conference, Delhi 2001.
13. D. C. Des Jarlais, K. Dehne and J. Casabona, "HIV surveillance among injecting drug users", *AIDS*, vol. 15 (suppl. 3), 2001, pp. S13-S22.
14. C. Fitch and G. V. Stimson, *An International Review of Rapid Assessments Conducted on Drug Use: A Report from the WHO Drug Injection Study Phase II* (Geneva, World Health Organization, 2003).
15. J. Giesecke, *Modern Infectious Disease Epidemiology* (London, Arnold, 2001).
16. M. A. Bellis and others, "Unexplained illness and deaths among injecting drug users in England: a case control study using regional drug misuse databases", *Journal of Epidemiology and Community Health*, vol. 55, 2002, pp. 843-844.
17. R. I. Glass and E. K. Noji, "Epidemiologic surveillance following disasters", Public Health Surveillance, W. E. Halperin, E. L. Baker and R. R. Monson, eds. (New York, Van Nostrand Reinhold, 1992).
18. R. Pearson and M. D. Kessler, "Use of rapid assessment procedures for evaluation by UNICEF", *Rapid Assessment Procedures: Qualitative Methodologies for Planning and Evaluation of Health Related Programmes*, Nevin S. Scrimshaw and Gary R. Gleason, eds. (Boston, Massachusetts, International Nutrition Foundation for Developing Countries, 1992), pp. 387-401.
19. G. Bammer, M. Beers and M. Patel, "The role of field epidemiology in the investigation of drug problems", *Drug and Alcohol Review*, vol. 18, 1999, pp. 329-31.
20. M. Susser and E. Susser, "Choosing a future for epidemiology: from black box to Chinese boxes to eco epidemiology", *American Journal of Public Health*, vol. 86, 1996, pp. 674-677.

21. M. Agar. "Recasting the 'ethno' in 'epidemiology'", *Medical Anthropology*, vol. 16, 1995, pp. 1-13.

22. P. Bourgois, "The moral economies of homeless heroin addicts: confronting ethnography, HIV risk, and everyday violence in San Francisco shooting encampments", *Substance Use and Misuse*, vol. 33, 1998, pp. 2323-2325.

23. P. Bourgois and J. Bruneau, "Needle exchange, HIV infection, and the politics of science: confronting Canada's cocaine injection epidemic with participant observation", *Medical Anthropology,* vol. 18, 2000, pp. 325-350.

24. S. C. M. Scrimshaw, "Adaptation of anthropological methodologies to rapid assessment of nutrition and primary health care", *Rapid Assessment Procedures: Qualitative Methodologies for Planning and Evaluation of Health Related Programmes*, Nevin S. Scrimshaw and Gary R. Gleason, eds. (Boston, Massachusetts, International Nutrition Foundation for Developing Countries, 1992), pp. 25-38.

25. R. Chambers, "Rapid but relaxed and participatory rural appraisal: towards application in health and nutrition", *Rapid Assessment Procedures: Qualitative Methodologies for Planning and Evaluation of Health Related Programmes*, Nevin S. Scrimshaw and Gary R. Gleason, eds. (Boston, Massachusetts, International Nutrition Foundation for Developing Countries, 1992), pp. 295-305.

26. M. E. Bentley and others, "Rapid ethnographic assessment: applications in a diarrhea management program", *Social Science and Medicine,* vol. 27, 1988, pp. 107-116.

27. S. C. M. Scrimshaw and others, "The AIDS rapid anthropological assessment procedures: a tool for health education planning and evaluation", *Health Education Quarterly*, vol. 18, No. 1 (Spring 1991), pp. 111-123.

28. S. C. M. Scrimshaw and others, *HIV/AIDS Assessment Procedures: Rapid Anthropological Approaches for Studying AIDS—Related Beliefs, Attitudes and Behaviours* (Geneva, World Health Organization, 1990).

29. I. A. Agyepong and others, *The Malaria Manual Guidelines for the Rapid Assessment of Social, Economic and Cultural Aspects of Malaria* (Geneva, World Health Organization, 1995.

30. A. M. Almedom, U. Blumenthal and L. Manderson, *Hygiene Evaluation Procedures, Approaches and Methods for Assessing Water and Sanitation Related Hygiene Practices* (Boston, Masachusetts, International Foundation for Developing Countries, 1997).

31. L. Manderson and A. Larson, *Contextual Assessment Procedures for STD and HIV/AIDS Prevention Programmes* (Brisbane, Queensland, Australian Centre for International and Tropical Health and Nutrition, 1997).

32. E. Herman and M. Bentley, *Rapid Assessment Procedures (RAP): to improve the Household Management of Diarrhea* (Boston, Massachusetts, International Nutrition Foundation for Developing Countries, 1993).

33. A. Long, S. C. M. Scrimshaw and E. Hurtado, *Epilepsy Rapid Assessment Procedures. (ERAP): Rapid Assessment Procedures for the Evaluation of Epilepsy Specific Beliefs, Attitudes, and Behaviours* (Landover, Maryland, Epilepsy Foundation of America, 1988).

34. L. Manderson, "Applying medical anthropology in the control of infectious disease", *Tropical Medicine and International Health*, vol. 3, 1998, pp. 1020-1027.

35. M. Nichter, "Project community diagnosis: participatory research as a first step toward community involvement in primary health care", *Social Science and Medicine*, vol. 19, 2002, pp. 237-252.

36. J. Dorabjee and others, "Rapid situation assessment of injecting drug use in Delhi", unpublished report, 1998.

37. R. McKetin, "Use of RAM to monitor illicit drug use in developed countries: can rapid assessment result in a rapid response?", *Drug and Alcohol Review*, vol. 18, 1999, pp. 333-335.

38. M. S. Kumar and others, "Rapid assessment and response to injecting drug use in Madras, south India", *International Journal of Drug Policy*, March 2000.

39. G. V. Stimson, D. C. Des Jarlais and A. L. Ball, eds., *Drug Injecting and HIV infection: Global Dimensions and Local Responses* (London, Taylor and Francis, 1998).

40. R. N. Needle and others, "Methodologically sound rapid assessment and response: providing timely data for policy development on drug use interventions and HIV prevention", *International Journal of Drug Policy*, vol. 11, 2000, pp. 19-23

41. E. Quimby, "District of Columbia rapid assessment, response and evaluation", unpublished report, 2001.

42. C. Jenkins, "A six city rapid assessment in Bangladesh (short version)", unpublished report, 2001.

43. L. Maher, P. Sargent and M. Cunningham, "Rapid assessment of hepatitis C prevention strategies among injecting drug users from non-English speaking backgrounds (NESB)", unpublished report, 2000.

44. H. Lambert, "Methods and meanings in anthropological, epidemiological and clinical encounters: the case of sexually transmitted disease and human immunodeficiency virus control and prevention in India", *Tropical Medicine and International Health*, vol. 3, 1988, pp. 1002-1010.

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

Drug treatment data as an epidemiological indicator: methodological considerations and improved analyses

M. STAUFFACHER

Institute for Cultural Anthropology, University of Basel, Basel and Natural and Social Science Interface, Swiss Federal Institute of Technology, Zurich, Switzerland

ABSTRACT

Data collected in drug treatment agencies have played a prominent role in informing policy makers. Such data often represent the only information collected regularly and consistently over a certain period of time. The first treatment reporting systems were introduced in countries in the late 1960s and early 1970s. In Europe, the 1990s was the most important period in terms of the standardization and widespread implementation of such reporting systems.

In the present article, the advantages and limitations of treatment reporting systems are discussed; the shortfalls of certain methodological approaches are also described. Despite its limitations, the treatment reporting system is a simple but powerful instrument for tracking the changing patterns of problematic drug use and, as such, is a valuable epidemiological tool.

As data need to be adequately exploited, a concrete four-step model is presented that analyses data with the aim of building incrementally upon the expanding knowledge base that exists with regard to drug use behaviour. Good knowledge of the local drug situation is essential to the sound interpretation of results.

The conclusion reached is that treatment data need to be supplemented by additional information from other indicators and from qualitative studies in order for that data to contribute to the still developing theory of trends in drug use behaviour.

Keywords: drug information systems; drug treatment; drug trends; epidemiology; statistical method.

Introduction

Reviewing epidemiological information systems on drug use behaviour reveals a great number of different indicators stemming from a variety of sources, including population surveys, the police and justice system, mental health institutions and specific research projects. In particular, data collected in drug treatment agencies have played a prominent role in informing policy makers about actual situations

and recent developments. In many regions of the world, such data often represent the only information collected regularly and consistently over a certain period of time.

Historical overview

Starting in the late 1960s and early 1970s, several countries introduced the first systems for collecting data from treatment institutions. In the United Kingdom of Great Britain and Northern Ireland, the Home Office launched the Addicts Index, requiring doctors to list patients whom they considered to be addicted to illegal drugs ([1], p. 186). In the United States of America, the first such system was set up in 1972, collecting data on the admission and discharge of clients from federally funded treatment programmes ([2], p. 3). Also in 1972, the Mental Health Reporting Programme was introduced at the national level in Indonesia, making it possible to question all patients from 35 mental care institutions in that country ([3], p. 267). In 1976, the territory of Hong Kong started its Central Registry of Drug Abuse ([3], p. 273), utilizing a wide range of sources that included treatment data and information from law enforcement and welfare agencies. A similar but more comprehensive system was established in Malaysia in 1977 [4]; again, law enforcement agencies (such as the police, prisons and customs) were the focus, but treatment institutions (including hospitals, private doctors, offices and rehabilitation centres) were also included. At that time, similar case registry systems existed in many communist countries, although the information contained in them was rarely used for epidemiological research.

In Europe, the 1980s saw the introduction of several specialized treatment reporting systems in different countries:

(a) In Germany, the treatment-unit-based information system known as the einrichtungsbezogenes Informationssystem (EBIS) was set up in 1980 [5];

(b) In the Netherlands, the Central Methadone Registry was introduced in 1980 and the National Information System on Alcohol and Drugs (LADIS) was set up in 1986 [6];

(c) In Spain, the Sistema Estatal de Información sobre las Toxicomanías (SEIT) was established in 1987 [7];

(d) In the United Kingdom, the Drug Misuse Database was developed in Manchester in 1986 and used at the national level in 1990 [1].

While such developments were taking place, pilot studies were being undertaken with the aim of establishing a European standard for treatment reporting [8, 9].

Standardization

In Europe, the 1990s represented the most important period in the standardization and widespread implementation of treatment reporting systems. The

Treatment Demand Protocol, for example, was developed within the framework of the activities of the group of epidemiology experts in drug problems of the Pompidou Group of the Council of Europe.* The aim of the protocol for treatment reporting systems, the first that is Europe-wide, was to provide professionals and researchers with a standardized methodology for collecting and reporting core data on the profile of drug users in contact with treatment services, so that the data would be comparable between different treatment services, cities and countries [11]. The main steps towards producing the final protocol included a pilot study carried out in 1991 in Dublin and London, a developmental project with the participation of 11 cities, and several meetings of the expert group where standard definitions, a core data set for a model questionnaire and data collection procedures were agreed upon. The protocol was finalized in 1994 [12].

The Treatment Demand Protocol received further attention in Europe with the establishment of the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), which placed a Europe-wide treatment demand indicator among its first priorities. The Pompidou Group protocol and the experience gained through its application served as a starting point for the initiative of EMCDDA in the late 1990s to harmonize treatment reporting systems in Europe. In 2000, the joint EMCDDA and Pompidou Group protocol was published [13, 14]. That protocol introduces all core definitions and provides extensive methodological guidelines. It offers the most up-to-date and comprehensive guidance for the establishment of a treatment reporting system. The present article, therefore, focuses on issues other than the practical ones addressed in the protocol.

Implementation at the municipal and national levels

Producing a protocol in itself is not enough: a guarantee that it is implemented and used later for international comparisons is also necessary. For that reason, the Pompidou Group initiated a six-year project in Europe, beginning in 1991. In that project, the treatment demand protocol was implemented in an increasing number of countries in Eastern and Western Europe, primarily at the municipal level. In addition, data checking routines and internal consistency checks were developed. The computer programs checked for answers that were out-of-range or inconsistent and those that represented impossible combinations. Data aggregation at the municipal level was simplified and unified by provision of another small program. Guidelines for data reporting were made available in spreadsheet format to automate the calculation of percentages and minimize possible mistakes. Over the years, data from major cities across Europe were compiled in six annual reports.

*The Pompidou Group, formed in 1971 and operating within the framework of the Council of Europe, provides a forum in which it is possible to discuss and exchange information and ideas on a range of problems concerning drug use and drug trafficking. There are more than 40 members of the Pompidou Group, including all of the States in Western Europe and many of the States in Central and Eastern Europe. The main tasks and obejctives of the epidemiology expers are to improve the quality and comparability of data on drug use and to collect, monitor, compare and interpret trends in drug use in Europe [10].

In the first year (that is, comprising data from 1992), nine cities provided data; in 1997, more than 20 cities from all over Europe were covered and data were compiled and analysed in a final project report [15]. Besides annual data reports presenting basic descriptive statistics, more detailed analyses were conducted and results were exchanged at annual meetings. Three volumes of proceedings show many interesting examples of such analyses. The papers focused on methodological aspects of treatment reporting (such as the reliability of self-reporting), the practical implementation of a treatment reporting system, drug use patterns (including the evolution of new use patterns in Eastern Europe) and trends comparing findings at the municipal and national levels.

Since 2000, the new joint protocol of the Pompidou Group and EMCDDA has been implemented at the national level in all member States of the European Union, in many more countries in Central and Eastern Europe and in major cities of the Russian Federation. Results comparing national levels are published as part of EMCDDA annual reports on the state of the drug situation in Europe [16].

Outside Europe, treatment attendance data play a considerable role in several community epidemiology work groups, modelled on the Community Epidemiology Work Group based in the United States and established in 1976 [17], including the Epidemiologic Surveillance System of Addictions in Mexico; the Canadian Community Epidemiology Network on Drug Use; the South African Community Epidemiology Network on Drug Use; the Asian Multicity Epidemiology Program on Drug Abuse; and the Queensland Community Based Drug Reporting Working Groups in Australia. In a comprehensive and easy-to-read document, all indicators used in those systems, including treatment data, are presented and explained in detail [18].

Methodological considerations

Both the original Pompidou Group and the new joint protocol of the Pompidou Group and EMCDDA share major methodological principles. All starting treatment episodes over one year are collected (such collecting involves full coverage and does not allow sampling). In that sense, it is an event reporting system [3], although it goes a step further, as only one episode per person is actually reported. Controlling such a variable is normally done within a treatment centre using personal identifiers. In various centres, a generic code that includes date of birth, age, gender and initials is used. Ideally, a reliable register of individuals is the result. Treatment is broadly conceptualized to allow the inclusion of all drug users with institutional contacts.

Advantages

In principle, treatment data provide results on two levels: a direct measurement of the demand on treatment services and an indirect measurement of trends in drug use. Treatment data provide an insight into what sorts of clients go to which services for service planners and treatment providers. In addition, they provide a

basis for evaluating attempts to attract particular subgroups into treatment. With sufficient coverage, such data provide the actual incidence of treated drug use and—if data on treatment admissions are complemented with a yearly census of the population undergoing treatment—treated prevalence. Data on first treatment demand can also indicate changing patterns of more severe, problematic drug use in the communities or populations ([12], p. 5). In that respect, they provide a complement to the results from the general population and school surveys. The latter provide an overall picture of, for example, lifetime drug use, as no information is typically available about different patterns of drug use (such as routes of administration or combinations of different drugs). Furthermore, because the use of heroin or cocaine is rare in most countries, very little information about problematic drug use is available. In contrast, comparing profiles of drug users and drug use patterns in different treatment centres, regions or cities, taking into consideration reporting years and age cohorts, can provide a detailed picture of drug users in treatment centres.

It is important to distinguish first treatment admissions from subsequent, or repeated, ones. Tracking changes among those having received treatment for the first time can indicate new developments. To that end, the period of time between the first use of the primary drug and the first treatment needs to be followed carefully. Findings suggest that some subgroups with a shorter period of time between the two stages indicate evolving trends in the drug user population.

In summary, treatment data provide a great deal of relevant information on the actual situation and prevalent trends within the group of problematic drug users. The treatment indicator is the best developed indicator. It is easy to implement, because it fits readily into the routine administration at the beginning of a treatment. It can be adapted to different situations, such as treatment modalities, or cities and countries. The size of the instrument used by the Pompidou Group (it fits on a normal size page) partly explains its success. Larger instruments involve more work and prevent many treatment centres from participating. There is often a conflict between the researchers, who look for more data, and treatment personnel, who are reluctant to collect more. One possibility for addressing such a dilemma is to work with different modules: the core module could be filled in for all clients, while additional questions could be used for a (defined) sample only.

The cost of data collection is low, as much of the information needs to be collected and recorded as part of the treatment procedure. The low cost means that continuous collection of data for treatment related information is possible. For most other data sources, including surveys, such data collection would be impossible. However, of perhaps greater importance than cost is the continuity over almost 10 years and the dedication of several individuals within the Pompidou Group in implementing and continuously improving the indicator. In general terms, close contact and regular communication between data collectors or researchers and (potential) users of results are a prerequisite to a successful system. Systematic and ongoing collection, analysis and result reporting of treatment related data within a monitoring system offer a stable and long-term means of data collection. Those elements increase validity and make trend analyses possible.

Limitations

Firstly, given that there can be a delay or time lag, sometimes of several years, between the first time a drug is used and the first demand for treatment ([12], p. 6), any very recent developments are excluded. It is important to note that, in some regions, such a time lag is less than two years; even if it is longer, it applies to the overall population. Given a mean of, for example, six years, approximately one third of the first treatments can be administered within up to three years after their first use. Again, comparing different subgroups of the treated population is important. Nevertheless, changes in the time lag are crucial. A change between different reporting years may be traced back to the introduction of new treatment approaches or new drug trends. If new treatment centres attract those drug users who have not previously been in contact with a treatment centre, the effect may be an increase in the mean time lag. That may cause an overestimation of certain trends, including trends in the mean age of the treated population or changes in the pattern of drug use. In such a situation, adequate analytical techniques are important in helping to distinguish between observed and real effects.

The proportion of the population of drug users included is dependent on the treatment facilities available in different countries, on the accessibility of the centres and on existing waiting lists. Changes in treatment policies, which influence availability and accessibility of treatment, can also have an important impact on treatment data. That means that with the establishment of, for example, methadone programmes attracting mainly long-term heroin users, changes in the group studied are bound to occur. If the coverage is restricted either to State-owned specialized or residential treatment centres, for example, an important part of the information is probably missed. Usually institutions run by non-governmental organizations are reluctant to collect and communicate data. General practitioners involved in the treatment of drug users constitute another group that is not always willing to cooperate. Inadequate coverage can be addressed in two ways: by making data collection compulsory by law or by convincing all treatment managers (individually or as a group) to provide data. If coverage cannot be complete (which will usually be the case), several responses are necessary. First, a complete and updated list of all treatment centres is indispensable. Most treatment agencies collect basic data and provide figures in annual reports. Those results can be used to assess potential differences within the population covered (for example, age, gender and drugs used). It may be possible to influence those collection agencies to collect more data. If they are not willing to collect information continuously, they might agree to collect for a sample only (such as a random sample or for a period of one month each year). All such information enables the assessment of existing biases. There again, changes are more important than continuous underreporting.

A further problem is that drug users often attend different treatment centres in the course of one year. That leads to double counting, overestimation of the total treatment demand and possible biases in several variables. Many possibilities exist to avoid double counting; the easiest at first glance would be to create a case

register. Not all countries have the possibility of using a national identification system (registry), owing to practical constraints (such as high costs), legal problems (including confidentiality issues) or lack of trust (for example, some drug users may be afraid of the misuse of data). In the guidelines [12], a code consisting of an individual's initials, date of birth and gender is proposed which is distinctive but which conceals that individual's identity. Such a code or something similar is used successfully in many places. It is important to remember that it is not necessary for the code to be fully successful in eliminating all double counting. The objective is to reduce the probability of double counting to a level that gives a reasonably accurate count of the number of treatment admissions in a given year.

On a more general level, it is important to take into account that the data come from treatment reporting systems. Possible conclusions are, therefore, limited to generalizations about the treated population. As in any other study based on treatment statistics, the hidden or out-of-treatment population is excluded. For that reason, treatment data need to be supplemented with qualitative data (for example, snowball sampling, observation or other ethnographic research methods) as well as other indicators (such as drug-related deaths, data about infectious diseases and police arrests).

Other problems, apart from those involving the indicator itself, include, for example, the process of data collection, analyses and presentation of results. Like any other routine statistics, the administration of the project exploits most of the available resources. Normally, the time available is not sufficient for further and deeper statistical analyses. That situation weakens the possible benefit of the treatment reporting system and requires additional effort, at the municipal as well as at the international level. Much more information is available than most experts actually use.

Finally, the general limitations of the chosen methodological approach must be acknowledged. As the study is restricted to starting treatment episodes, an outcome measurement is not possible. In addition, the questionnaire utilized is not a clinical instrument; thus, no diagnoses are generally recorded. For those reasons, further instruments would be necessary (see, for example, the Maudsley Addiction Profile [19] or the Addiction Severity Index [20]). As no register of all people in treatment exists and no assessment of the waiting list is made, the information on the actual treatment demand is inadequate. Such information, however, can be added using other modules. Furthermore, treatment data collected during administrative procedures are no substitute for thoroughly designed research instruments. Such data, therefore, will not be able to explain, for example, why people use illicit drugs. To summarize, the treatment reporting system is a simple but powerful instrument for tracking patterns of problematic drug use and as such is a valuable epidemiological tool.

General lack of analyses

It is important to recognize that the analysis of treatment data requires more than descriptive tables and graphs. Much more can be and needs to be done.

Significant resources are invested in the development and running of treatment reporting systems. Ideally, equally significant resources should be invested in the actual use of the collected data. Treatment data still represent a major source of information with regard to epidemiology. Such data, however, need to be exploited adequately. Having reliable data is essential to any statistical analysis. Another prerequisite is a researcher's comprehensive knowledge of the data relevant to the research being undertaken. The use of statistical software programs does not replace competent statistical analyses.

Why statistical analyses?

The rationale for using statistical methods to analyse treatment data is sometimes questioned. It can be argued that, on a general level, real data are random. Suppose, for example, that there is a difference in mean age between groups with different patterns of drug use. Statistical analyses can indicate how significant such a difference is. Furthermore, the difference in the mean age may be attributed to other causes, such as a selection effect owing to insufficient coverage of specific treatment modalities. To reduce such threats, statisticians have developed techniques that control for the difference in other variables. The latter refers to analyses concerned with the comparison of more than two variables, called multivariate analyses, controlling for group differences and spurious relationships.

More specifically, in the field of treatment data, it is important to emphasize that treatment monitoring systems do more than mere bookkeeping. As with any other routine statistics, administration of the project exploits a large part of the available resources. That weakens the possible benefits of such a system. Much more information is available than most system managers actually use. It is important to frame specific questions, with the intention of pursuing specific answers. For that reason, it is not sufficient to be content with describing frequency distributions. It is essential to develop concrete hypotheses and test them. Similarly, annual result reporting is not sufficient: building incrementally upon an expanding knowledge base should be the broad aim in analysing treatment data.

Four steps to analysing data from treatment reporting systems

Reliability of data

An important prerequisite to any statistical analysis is the reliability of the data. The first stage is to agree on uniform key definitions, that is, to use definitions that are well known and used in all treatment centres. That requires continuous training efforts. An important element here is regular feedback on the quality of data and results. Data collectors aware of the potential uses of treatment monitoring are likely to support the idea and contribute to rigorous data reporting.

Reported data need careful checking, a process known as data cleaning. For example, if a multiple answer was given where only one answer was possible, either a missing value needs to be assigned or a decision needs to be made about applying a clear rule when recoding takes place. If a question concerns an

individual's living situation and the answer given is both "alone" and "in an institution", the decision can be made to recode all such answers as "in an institution". It is then necessary to verify that only valid answer codes have been used (for example, checking that an answer such as "six" is not given to a question concerning employment status, where there are only three possibilities). The verification stage also includes conducting consistency checks and inspecting out-of-range data in continuous variables. The most reliable method is computer-assisted data collection, that is, using specific software for data input, data checking and reporting.

A statistics program is strongly recommended. Such a program involves defining the data by attaching meaningful labels to variables (that is, questions) and codes. It is important to remember that even correct data can be wrong. The assumption underpinning treatment monitoring systems, for example, is that self-reporting by drug users is valid—an assumption, in fact, that is supported by evidence. Nevertheless, specific studies that question and test the validity of data are helpful and can provide essential details.

Descriptive analysis

Statistical analyses should begin with basic analyses, illustrated in the following steps:

(a) Computing frequency distributions on all variables; summarizing data using measures for central tendency (mean, median) and variation (standard deviation, range);

(b) Cross-tabulating the most important variables (for example, route of administration by primary drug, level of education by injecting behaviour);

(c) Visualizing the results and using graphs to gain more insight into the data.

The steps of the basic analysis outlined above should lead to a thorough understanding and knowledge of the data set in question and can be considered a prerequisite to advanced statistical analysis. The ease of using statistical software, however, should not lead to conducting tests indiscriminately, simply because they are available and easy to run. Good results and sound interpretation are unlikely to come from such efforts.

Testing hypotheses (bivariate and multivariate analysis)

At the stage of testing hypotheses, the data set should have been verified and the data themselves should be completely familiar. Testing hypotheses at the bivariate level can begin, as can undertaking more advanced statistical analyses. As treatment data are collected consecutively over years, most analyses will deal with changes over years, or trends. Some analyses involve describing frequency distributions, as well as testing hypotheses. A few good examples include: Smyth and others [21], who show the change in the route administration from injecting to smoking heroin reflected in their treatment reporting system in Dublin; de la

Fuente and others [22] followed the same question, comparing different regions in Spain; Falcato and others [23] showed the impact of the closure of an open drug scene for Zurich, Switzerland; Agar and Schacht Reisinger [24], in a wide-ranging account of some important developments in heroin use in the Baltimore area, in Maryland, United States, showed how important regional and local breakdowns of collected data are. Certain trends, for example, may show up only in the inner city or in the suburbs. To that, further breakdowns as age cohort or cohorts of calendar years of first drug use could easily be added.

Interpretation

Good knowledge of the local drug situation is essential to the sound interpretation of results. Treatment data need to be supplemented with additional information from other sources. As for any study based on treatment statistics, the conclusions should be restricted to the treated population, or restricted to those involved in the study. It is always necessary to cross-check results with information from other sources, such as expert opinions, surveys, police data and information from field observation. It is also necessary to conduct qualitative studies that can fill in the detail of the picture of current drug use. The complementary use of quantitative and qualitative methods is crucial to the development of effective responses to drug use. Close links with ethnographers and other field researchers can both help interpret results and provide ideas for further analysis. Again, the work of Agar and Schacht Reisinger [24] is important. They combined and cross-checked different indicators with qualitative information from mass media and oral accounts from students.

A new project supported by the Council of Europe develops those ideas [25]. The project looks (retrospectively) for important trends found in treatment data from cities and tries to describe and explain them in detail. The project takes an epidemiological perspective, using treatment data as an indicator for drug-related problems in cities (for example, incidence and prevalence of high-risk use and morbidity). The aim is to develop detailed descriptions of emerging trends and explanations of those trends. It is hoped that the project will contribute to the developing theory on trends in drug use [26, 27].

Conclusions: combining with other indicators and further information

In the field of drug use epidemiology, treatment-based indicator data remain a source of some of the most valuable information available. The collection of such data is relatively easy and cost-effective and can be combined readily with the administrative work of personnel involved in the treatment process. Accepting the methodological framework and analysing in detail the growing body of data can give a wealth of important and useful information.

Among the observations made, it was noted that treatment data need to be supplemented with information from other indicators to allow for sound

conclusions to be drawn. There are several frameworks that describe in detail the combined use of different indicators [9, 18]. Most of them, however, remain on a purely descriptive level, presenting data from all indicators either in parallel or consecutively. In contrast, Hartnoll [28] cross-validates information from different indicators to an extent, although the result remains descriptive and restricted to the macro level of data (that is, the data are not broken down). In some chapters in the third multi-city study of the Pompidou Group [29] (in particular, those sections written by Uwe Kemmesies), correlations between different indicators and different groups of countries are analysed. A model of epidemiological stages is also explored. Further efforts, nevertheless, continue to be necessary, in order to explore trends in drug treatment data and to use different kinds of breakdowns (including age, primary drug, route of administration and calendar year of first use).

As for any study based on treatment statistics, the hidden population is excluded. Access to such users can be gained using, for example, snowball sampling. Experience shows that in many cities, there is a pressing need to develop a more detailed understanding of patterns of drug use. Such an understanding is unlikely to come from quantitative methods alone. What is required is the implementation of qualitative studies. The complementary use of quantitative and qualitative methods is crucial to the development of effective responses to drug use [30]. Once again, the work by Agar and Schacht Reisinger [31] deserves mention: results are cross-checked with their independent analysis of all indicators and many other sources of information, comparing them with different parts of the country and assimilating the elements into a preliminary theoretical model of trends in drug use behaviour. The latter is especially exceptional in the field, where descriptive analyses and fractional testing of arbitrary hypotheses still prevail. Only by formalizing knowledge and testing it with real data will our understanding about changing drug use behaviour successively increase.

References

1. M. Donmall, "UK monitoring of problem drug users: the Drug Misuse Database—a system based on regional centres", *European Addiction Research*, vol. 5, No. 4 (1999), pp. 185-190.
2. Substance Abuse and Mental Health Services Administration, Office of Applied Studies, *Treatment Episode Data Set (TEDS): 1994-1999; National Admissions to Substance Abuse Treatment Services*, DASIS Series S-14, DHHS Publication No. (SMA) 01-3550 (Rockville, Maryland, 2001).
3. R. G. Smart, "Reporting systems", *Guide to Drug Abuse Epidemiology* (Geneva, World Health Organization, 2000), pp. 249-325.
4. V. Navaratnam and K. Foong, "Development and application of a system for monitoring drug abuse: the Malaysian experience", *Bulletin on Narcotics* (United Nations publication), vol. 40, No. 1 (1989), pp. 53-65.

5. R. Simon, E. Hoch and A. Holz, "The German monitoring and reporting system for the treatment of substance-related problems: a national system on the basis of aggregated data", *European Addiction Research*, vol. 5, No. 4 (1999), pp. 167-172.
6. A. W. Ouwehand and V. C. M. van Alem, "The Dutch national alcohol and drugs database: progress in monitoring, monitoring in progress", *European Addiction Research*, vol. 5, No. 4 (1999), pp. 173-179.
7. A. Alvarez-Requejo and others, "Monitoring treatment demand for drug abuse in Spain: perspective over a decade", *European Addiction Research*, vol. 5, No. 4 (1999), pp. 179-184.
8. R. Hartnoll, "Current situation relating to drug abuse assessment in European countries", *Bulletin on Narcotics* (United Nations publication), vol. 38, No. 1 (1986), pp. 65-80.
9. R. Hartnoll and others, "A multi-city study of drug misuse in Europe", *Bulletin on Narcotics* (United Nations publication), vol. 40, No. 1 (1989), pp. 3-27.
10. R. Hartnoll, "The work of the Pompidou Group in the field of drug research", *European Addiction Research*, vol. 1, Nos. 1-2 (1995), pp. 71-73.
11. M. Stauffacher and A. Kokkevi, "The Pompidou Group Treatment Demand Protocol: the first pan-European standard in the field", *European Addiction Research*, vol. 5, No. 4 (1999), pp. 191-196.
12. R. Hartnoll, *Drug Treatment Reporting Systems and the First Treatment Demand Indicator. Definitive Protocol* (Strasbourg, Council of Europe Publishing, 1994).
13. R. Simon and others, "The EMCDDA/Pompidou Group Treatment Demand Indicator: a European core item set for treatment monitoring and reporting", *European Addiction Research*, vol. 5, No. 4 (1999), pp. 197-207.
14. European Monitoring Centre for Drugs and Drug Addiction and Pompidou Group, *Treatment Demand Indicator. Standard Protocol 2.0* (Lisbon, 2000).
15. M. Stauffacher, *Pompidou Group Project on Treatment Demands: Treated Drug Users in 23 European Cities; Data 1997, Trends 1996-1997* (Strasbourg, Council of Europe Publishing, 1999).
16. European Monitoring Centre for Drugs and Drug Addiction, *Annual Report on the State of the Drugs Problem in the European Union 2001* (Luxembourg, Office for Official Publications of the European Communities, 2001).
17. National Institute on Drug Abuse, *International Epidemiology Work Group on Drug Abuse, June 1999: Proceedings* (Rockville, Maryland, 1999).
18. *Assessing Drug Abuse within and across Communities: Community Epidemiology Surveillance Networks on Drug Abuse* (Rockville, Maryland, National Institute on Drug Abuse, 1998).
19. J. Marsden and others, "The Maudsley Addiction Profile (MAP): a brief instrument for assessing treatment outcome", *Addiction*, vol. 93, No. 12 (1998), pp. 1857-1867.
20. A. T. McLellan and others, "The fifth edition of the Addiction Severity Index", *Journal of Substance Abuse Treatment*, vol. 9, 1992, pp. 199-213.
21. B. P. Smyth, M. O'Brien and J. Barry, "Trends in treated opiate misuse in Dublin: the emergence of chasing the dragon", *Addiction*, vol. 95, No. 8 (2000), pp. 1217-1223.
22. L. de la Fuente and others, "Declining prevalence of injection as main route of administration among heroin users treated in Spain, 1991-1993", *European Journal of Public Health*, vol. 7, 1997, pp. 421-426.

23. L. Falcato and others, "Closure of an open drug scene: a case register-based analysis of the impact on the demand for methadone maintenance treatment", *Addiction*, vol. 96, No. 4 (2001), pp. 623-628.
24. M. Agar and H. Schacht Reisinger, "Numbers and patterns: heroin indicators and what they represent", *Human Organization*, vol. 58, No. 4 (1999), pp. 365-374.
25. M. Stauffacher, *Pompidou Group Project on Treatment Demand: Tracking Long-term Trends; Introduction and Guidelines (Basic Module)* (P-PG/Epid (2002) 01 E) (Strasbourg, Council of Europe/Pompidou Group, 2002).
26. M. Agar, "Towards a trend theory", *Understanding and Responding to Drug Use: the Role of Qualitative Research*, EMCDDA Scientific Monograph Series No. 4, J. Fountain, ed. (Luxembourg, Office for Official Publications of the European Communities, 2000), pp. 67-72.
27. D. J. Korf, "Substance careers: winners and losers on the market of 'new' drugs", *Illicit Drugs, Patterns of Use: Patterns of Response*, Proceedings of the 10th Annual ESSD Conference on Drug Use and Drug Policy in Europe, A. Springer and A. Uhl, eds. (Innsbruck, Vienna and Munich, Studienverlag, 2000), pp. 123-141.
28. R. Hartnoll, "Cross-validating at local level", *Estimating the Prevalence of Problem Drug Use in Europe*, EMCDDA Scientific Monograph Series No. 1 (Luxembourg, Office for Official Publications of the European Communities, 1997), pp. 247-261.
29. R. Bless, *3rd Multi-city Study: Trends in 42 European Cities in the 1990s* (Strasbourg, Council of Europe Publishing, 2000).
30. J. Fountain, ed., *Understanding and Responding to Drug Use: the Role of Qualitative Research*, EMCDDA Scientific Monograph Series No. 4 (Luxembourg, Office for Official Publications of the European Communities, 2000).
31. M. Agar and H. Schacht Reisinger, "Explaining drug use trends: suburban heroin use in Baltimore county'', *Illicit Drugs: Patterns of Use; Patterns of Response*, Proceedings of the 10th Annual ESSD Conference on Drug Use and Drug Policy in Europe, A. Springer and A. Uhl, eds. (Innsbruck, Vienna and Munich, Studienverlag, 2000), pp. 143-165.

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

An ethno-epidemiological approach for the multi-site study of emerging drug abuse trends: the spread of methamphetamine in the United States of America*

A. PACH III

Kathmandu, Nepal

E. M. GORMAN

College of Social Work, San Jose State University, San Jose, California, United States of America

ABSTRACT

In the present article, the authors describe a multi-site, ethno-epidemiological study of methamphetamine use in five cities in the United States of America. They present techniques and strategies used to develop and link short-term, focused qualitative studies. Both local variations and comparable issues germane to methamphetamine use and its public health consequences were examined in the studies. The article contains a description of the methods utilized to identify and access various subgroups of drug users, collect sensitive information on drug use and risky sexual and drug use behaviour and establish links between the study sites. In the study, data collection was tailored to fit the local contexts of methamphetamine use and involved the use of a shared protocol and reporting format. Collaborative participation in developmental meetings and regular communication by Internet facilitated the collection of comparable data. The diverse and socially embedded nature of methamphetamine use indicated: (a) *a wide array of user groups and patterns of methamphetamine use;* (b) *multiple health and social consequences of methamphetamine use; and* (c) *the need for ongoing qualitative and quantitative monitoring of trends in methamphetamine use. The approach followed in the study suggests methods for enhancing the collection of information in current surveillance networks.*

Keywords: methamphetamine; ethnography; emerging drug trends.

*The authors gratefully acknowledge the continuous support and direction of Rich Needle and Elizabeth Lambert of the National Institute on Drug Abuse of the United States of America throughout the study. Mike Agar, Michael Clatts, Ross Gibson, Stephen Koester and Claire Sterk were members of the multi-site study group who made invaluable contributions to the development and conduct of the study and to the preparation of the present article. The study was funded by grant number R01DA1087 and contractual arrangements with the National Institute on Drug Abuse of the United States of America.

Multi-site study of methamphetamine use

Amphetamine-related stimulants have become the most widely used illicit synthetic drugs in the United States of America and in much of the world [1]. In the United States, there was an explosive rise in methamphetamine use in the 1990s, especially in parts of California, Washington and the Midwest [2, 3]. The total number of treatment admissions for methamphetamine expanded by fourfold in the period 1992-1997 [4]. The level of methamphetamine use eclipsed that of other illicit substances in a number of settings [5], and the rate of human immunodeficiency virus (HIV) infection grew exponentially among methamphetamine users in certain subpopulations such as homosexual and bisexual men [6, 7]. Furthermore, methamphetamine use was associated with an increase in new and younger users [8]. Accompanying those trends was an expansion in small local laboratories used to manufacture methamphetamine and the involvement of Mexican criminal organizations in trafficking and distributing methamphetamine throughout the United States [9]. Ethnographic and media reports indicated that methamphetamine use was growing in popularity among user groups in a number of cities on the eastern coast of the United States, such as Atlanta, New York and Washington, D.C. [2]. Methamphetamine thus presented a serious drug abuse problem with significant public health consequences, and it emerged in the 1990s as a major concern of national drug policy makers [9].

There was much uncertainty about how to deal with the growing problem of methamphetamine use. That was related in part to the dearth of studies on the characteristics of methamphetamine users and the patterns of its use [10] The few intensive studies that existed [11-13] indicated an enormous diversity of user groups. Those groups included homosexual and bisexual men, college students, white-collar businessmen, young adults and youth in the rave or club culture and living on the street and longtime user groups such as outlaw motorcycle gangs. Within those groups there was much variation by class, ethnicity, rural/urban residence and social identity, many social groups presenting different patterns of drug use and health risks [14-16]. Methamphetamine use identified in the initial studies ranged from occasional use to binging and chronic use, which varied depending on the user subgroup [17, 18].

Thus, there was a need to document patterns of methamphetamine use among diverse subpopulations of users and to identify the associated health risks. The widely varying patterns of use and user subgroups necessitated responsive methods of data collection. Given the time constraints of survey research, it was decided that a short-term, highly focused ethnographic study implemented in multiple sites could rapidly access diverse, hard-to-reach populations and collect comparable information.

With support from the National Institute on Drug Abuse of the United States, the authors launched a short-term, exploratory ethnographic study with researchers in five metropolitan areas in the United States from 1997 to 1998; the study was expanded to include two other sites in 1999. In the present article, the authors examine the strategies and challenges involved in designing an

ethnographic project in multiple sites in an effort to explore newly emerging drug abuse trends in methamphetamine use. The article includes a description of the conceptual and methodological framework of the study, its sampling of "hidden" populations, its approach to collecting sensitive information and the strategies that supported comparative linkages in data collection and analysis between the sites. The study has contributed to the growing area of "ethno-epidemiology", in which risk behaviours are examined in relation to the user population and the social contexts in which risky behaviours occur, in order to improve the content and delivery of public health preventive services and interventions [19-21].

Methamphetamine

In the United States, amphetamine-type substances vary in their effects and composition. Popularly known as "speed", "crystal", "crack" or "ice", such substances are swallowed, snorted, smoked or injected. Gradations in the quality of methamphetamine relate to the precursor chemicals involved, the manufacturing techniques used and the channels used for trafficking and distributing the substance [9]. For instance, the red phosphorus method produces high-quality methamphetamine, which is distributed from the south-western part of the United States and California and is believed to be trafficked by Mexican drug trafficking organizations. Lower-quality methamphetamine (such as that manufactured using the 1-phenyl-2-propanone (P-2-P)) method is distributed by motorcycle gangs and small laboratory manufacturers in various areas. The wide range of sources and forms of methamphetamine spread throughout much of the United States in the 1990s [22].

The use of methamphetamine affects neural transmitters and receptors, producing euphoria and increased energy and mental activity. However, long-term use of the substance can produce psychosis, depression and drug dependency and may lead to death [23]. Studies have shown that many of the compounds and by-products of methamphetamine manufacture contain lead and heavy metals [24]. Those substances may be associated with neurological damage that has been documented by a number of studies of long-term or high-volume methamphetamine use [25, 26]. Chronic methamphetamine use may also lead to significant health and social problems such as job loss, interpersonal problems, infection with sexually transmitted diseases and HIV infection [6, 12, 27-29]. However, patterns of methamphetamine use vary greatly, depending on age, gender, sexual preference, occupation and social group; that applies to both highly addictive and more intermittent and recreational patterns of use [30, 31].

Methodological background

The multi-site methamphetamine study included geographical locations across the United States with populations at high risk for methamphetamine use. The initial study sites included Atlanta, Denver, New York, Sacramento and Seattle;

subsequently, San Jose, California and south-western Missouri were added. The sites were selected to provide a broad sampling of methamphetamine contexts with the intent of gaining insight into a variety of geographical, demographic and cultural drug ecologies while making the best use of the limited time and resources available. Each site targeted a different combination of high-risk populations. The study was to take six months to complete, from beginning to end. The aim of the study was to collect information on characteristics of users, patterns of methamphetamine use and health risks associated with methamphetamine use.

There were a number of advantages to using ethnographic methods. Short-term ethnographic studies provided the means to delineate various social categories of methamphetamine users and emerging issues concerning the continued and increasing use of methamphetamine. In the framework of an ethno-epidemiological approach [20], the ethnographers worked at the community level to identify categories of users and contexts of use by integrating data gathered through multiple qualitative methods and from secondary sources of data on methamphetamine use, including problem indicator data (for example, arrest and treatment data) and media reports. The use of qualitative methods allowed the ethnographers to grasp the meanings, motivations and consequences of methamphetamine use from the perspective of users, as well as from the perspective of substance abuse professionals. Those perspectives placed methamphetamine use in meaningful sociocultural and public health contexts.

The study design consisted of four components: *(a)* a semi-structured guide that would be used with a sample of at least 20 methamphetamine users; *(b)* unobtrusive observations in sites where users congregated; *(c)* interviews with professionals and lay individuals knowledgeable about methamphetamine use; and *(d)* the collection of problem indicator data and reports on methamphetamine use. The completion of data collection in six months was facilitated by the nature of the study design, which focused on particular issues, a limited sample of respondents and ready access to professional and street contacts knowledgeable about or involved in drug scenes in certain communities [32].

At least two "scenes" of methamphetamine users were described at each site. Scenes were defined by the social characteristics of users and their identifying sexual, interaction and expressive practices (for example, sexual preferences, music, clothing or occupational attributes). Those features influenced subgroup norms and behaviours that shaped their patterns of methamphetamine use (such as the frequency and mode of ingestion) and the contexts in which they used methamphetamine (such as in clubs or on the street) [19, 33].

In accessing various groups of drug users and collecting information, it was essential that the ethnographers develop rapport with the users in order to collect credible data in a limited amount of time. That was only possible because all of the ethnographers had pre-established knowledge of local drug scenes and contacts with drug users within their communities. In addition, the ethnographers regularly interacted with local drug surveillance, treatment and prevention organizations. Thus, their local knowledge of drug use and established contacts with professionals and drug users from various drug scenes were essential aspects of the

short-term projects. Issues of access to various user subpopulations were thus related to the depth and credibility of the information that the ethnographers were able to collect. Those issues are discussed in the sections below.

Project linkages

User subpopulations and patterns of drug use specific to the contexts of each setting were identified. Among other user groups, the study focused on rural and female users in Atlanta [34], a comparison of heroin and cocaine injectors with methamphetamine injectors and the identification of long-time users of methamphetamine in Denver [35], polydrug users and minority female users of methamphetamine in Sacramento [36], homosexual and bisexual male and young adult and female heterosexual users in Seattle [14, 31, 37], and street youth and homosexual and bisexual men in New York who visited clubs and used methamphetamine [18]. In California, ethnographers investigated the diffusion of methamphetamine use in tandem with drug use in clubs (for example, methylenedioxymethamphetamine (MDMA, commonly known as "Ecstasy"), *gamma*-hydroxybutyric acid (GHB)) among homosexual and bisexual men in the San Francisco Bay area and central California [38]. The wide range of drug users and contexts of use necessitated flexibility in the approach used to access each subpopulation and to collect the data. However, there was also a need to produce comparative information on relationships between patterns of use and the distribution of health and social consequences. Therefore, the project initiated a number of processes that linked forms of data collection and analysis. These efforts extended throughout the study and involved utilizing a shared protocol, having a number of planning, site visit and presentation meetings and using the Internet. Those activities generated shared frameworks for data collection and analysis.

1. Project design and protocol

The projects shared a study design that drew on rapid assessment [38] and community identification procedures [39]. Those approaches utilized a complementary blend of methods for identifying poorly understood health and social problems. The identification and recruitment of small samples of methamphetamine users and the emergent and "hidden" nature of their drug use necessitated the use of multiple data sources. Those sources included informal, open-ended and semi-structured interviews and participant observations with users, exploratory interviews with methamphetamine specialists and problem indicator data (police reports, emergency department mentions, drug treatment data and media reports).

The interview protocol was designed by the authors to examine drug use patterns and health risks. It was collaboratively expanded and adapted to the contexts of each project in order to identify connections between drug use and its health and social consequences. It included questions on drug use behaviour (patterns and frequencies of use), social aspects of drug use (characteristics of users,

purposes of use (dance, sex, work)), forms of drug distribution (procurement and the substances used) and HIV risk behaviour and other health and social consequences of methamphetamine use.

2. Meetings

A number of meetings were held to promote the development of a shared framework for data collection and analysis [40]. The authors initially met with staff at the National Institute on Drug Abuse to identify issues of concern regarding methamphetamine use and its public health consequences. The ethnographers also discussed those issues and their approach to examining methamphetamine use. Following that meeting, the ethnographers met among themselves to decide on the project design and the topics to be examined.

In the course of the project, project coordinators visited sites to observe their research settings and gain a sense of the information being collected at each site. During those visits, they also shared a series of issues that reflected the public health concerns of the National Institute on Drug Abuse with regard to methamphetamine use. The on-site discussions, together with follow-up communication by means of electronic mail (e-mail), produced a reporting format that guided the ethnographers' analyses and reports.

Following data collection, there was a series of meetings in which the researchers met among themselves and with staff of the National Institute on Drug Abuse. The ethnographers first shared among themselves key findings and emerging conceptual issues and implications. In the discussion, the shared and distinct drug-use patterns and health risks of the sites were identified. That meeting was followed by presentations in which the ethnographers discussed their preliminary findings with staff of the National Institute on Drug Abuse and then jointly participated in a discussion on the insights and implications of their research. Following those meetings, the ethnographers completed their analyses and reports.

All of the meetings and discussions led to an agreement on the issues to be examined, analysed and presented in final reports. Regular communication played an essential role in that process. It allowed the ethnographers to agree on their approach to the research and to share perspectives and concepts that emerged in their analyses. Regular communication using the Internet was a key factor in that process.

3. Use of the Internet

Regular communication using the Internet facilitated project development in a number of ways. It provided a means of immediately sharing information (such as references and initial findings), supported collaborative planning (regarding site visits, schedules and deadlines) and facilitated the project's conceptual development through ongoing discussions of emerging issues, analytical categories and substantive findings. Through monthly communication, the Internet served as a

forum for discussing concepts and strategies of data collection and analysis. The discussions included comments on the interview guide, the nature of user groups and social aspects related to drug use [41].

Such communication generated a number of conceptual perspectives. Two useful concepts were the multifunctional uses of methamphetamine and the "scenes", or configurations, of social contexts and groups of methamphetamine users. The concept of multifunctional uses of methamphetamine was especially useful given the wide range of forms, functions and effects of methamphetamine use among user groups [17]. For example, methamphetamine use was found to enhance recreational dancing, work performance, sexual encounters, social interactions and everyday functioning. The various purposes and perceptions of methamphetamine use were related to its initiation, continued use and variability in its patterns of use. Those diverse associations were found to be related to particular health and social consequences.

The notion of "scenes" of methamphetamine users acted as a unifying methodological and analytical concept. Drawing on theories of cultural production [42] and the expression of subcultural identities [43], the ethnographers identified "scenes" of drug users by the social characteristics of the users (age, gender, sexual preference, occupation) and their social, occupational and expressive practices. The concept of scenes of methamphetamine use and users provided a framework for identifying social categories of users and related contexts and practices of methamphetamine use that were meaningful from "insider" social and analytical perspectives. Identification of these contexts and practices provided insight into configurations of situations and norms that led to risky health and social behaviours.

Methods

1. Identifying and recruiting methamphetamine users

The initial challenge for each site was to identify meaningful social categories of users in order to identify patterns of methamphetamine use and the distribution of health risks among representatives of those categories [44]. Methamphetamine users do not make up a visible or readily identifiable population, and a random sample cannot be drawn from an unknown universe of users. In order to develop a baseline of relevant types of methamphetamine users, the projects utilized a series of approaches that had been developed in applied ethnographic research with undersampled and hard-to-reach subpopulations. In approaches that combined targeted sampling [45] and community identification processes [18, 39], project ethnographers consulted available research reports, the "internal" knowledge of their outreach staff and community contacts knowledgeable about methamphetamine use, including current users. They also observed the settings that users frequented. Information from those sources enabled the ethnographers to begin to identify distinctions in types of users in their communities.

Those efforts entailed convenience, snowball and modified targeted sampling activities. For instance, a convenience sample of prior research contacts with methamphetamine users led to initial interviews with such users. The initial interviews and other sources of information began to indicate distinct groups of methamphetamine users and provided contacts for other interviews with users. As information accumulated, the nature of categories and "scenes" of users were increasingly delineated and revised in an evolving process. The attributes of users and their patterns of behaviour were thus revealed in a process in which further interviews confirmed categories or suggested other categories of users [44, 46].

Various approaches were used to access and recruit individuals from the various "scenes" of methamphetamine users, such as members of club or street cultures. For instance, some users could be recruited directly through a chain referral process that stemmed from previously established contacts. Other categories of users had to be recruited on a one-to-one basis more gradually from contacts developed at congregation sites or through other social intermediaries, such as bartenders. To recruit individuals from other groups, it was necessary to rely on the reputation and status of a staff member who had strong ties to often wary and hidden user groups. Initial locating and recruiting activities also involved the use of various public media, such as advertisements in local newspapers and trade newsletters that certain groups (for example, homosexual and bisexual men) were known to read. Referrals from public health, social service and community-based organizations and other professional agencies such as substance abuse treatment programmes or HIV outreach and support programmes were also used to recruit individuals.

The various forms of recruitment were necessary for identifying and accessing important at-risk groups and networks of users. However, some user groups (such as truck drivers, motorcyclists and upper middle-class users) were especially difficult to recruit in the time allotted because of sociocultural barriers. Again, the development of rapport and trust among community contacts was critical to the efforts to recruit and interview methamphetamine users at the sites.

2. Data collection

Ethnographic data collection began in February and March 1998 and continued until the end of September 1998 at the initial sites and until the end of 1999 at the additional sites. Over 200 interviews were conducted at the five sites with active or recent users of methamphetamine who were over the age of 18. The interviews lasted between one and two hours and respondents were reimbursed for their time. The exigencies of interviews during participant observation sessions varied in their length and degree of formality, as did interviews with knowledgeable community professional and lay contacts in the community.

Ethnographic data collection occurred initially as part of the process of identifying and recruiting users. It evolved from interviews aimed at identifying categories of methamphetamine users and locales where such users congregated to interviews focused more on particular aspects of methamphetamine use and users.

In the same manner, exploratory participant observation in bars and clubs evolved to become observation in particular clubs or on the street and, over time, in more long-standing projects, in users' homes. Just as sampling categories were scrutinized and revised throughout the study, questions on methamphetamine use were supplemented and respondents' answers were probed as issues arose throughout the data collection process [47].

The multi-site protocol contained both closed and open-ended questions and respondents were encouraged to speak at length about any subject that they thought was relevant to their own use of methamphetamine, and to their perceptions of methamphetamine users in their community [48]. Socio-demographic data and information on HIV risk behaviours were collected through closed-ended questions, while statements containing information on current and past methamphetamine use and other drug use were elicited through semi-structured questions. Accounts of incidents of methamphetamine use or sexual episodes associated with methamphetamine use were gathered in an open-ended, conversational format. Respondents were encouraged to elaborate on issues of interest to them throughout the interview; that approach yielded information that was unexpected and that indicated their perspectives on methamphetamine. Many of the interviews were audiotaped and transcribed. Listening to audiotaped interviews provided non-verbal information on voice tone, pauses in discussions and other communicative information not available in hard copies of transcripts. In addition to the interviews, the ethnographers collected field notes that provided information about where an interview or observation occurred, the physical appearance and behavioural nuances of the respondents and other impressions and experiences of the ethnographers in user settings.

Given the sensitive nature of the information, research relationships and the setting of the interviews influenced the type of information collected [49]. For example, interviews that occurred informally during participant observation at clubs or bars with users who had just been contacted could cover only a limited range of topics. In collecting sensitive information, some ethnographers, at the beginning of the interview, initiated open-ended discussions with respondents about their life history and about methamphetamine users in their community; towards the end of the interview, more specific personal questions (such as HIV and drug risk behaviours) were covered. Where possible, some ethnographers would utilize open-ended and conversational formats throughout the initial interview and then interview respondents a second time on topics that had not been covered or that had emerged during the first interview.

The above-mentioned approaches to data collection elicited responses on topics that were less threatening, allowed rapport to develop during the interview process and led to the collection of unexpected information as users shared their views on methamphetamine use. The various formats and developmental approaches to collecting information also allowed researchers to frame questions and allowed the respondents to consider issues from varying perspectives, thus providing the ethnographers with a means to cross-check sensitive or complex information, such as HIV or drug risk behaviour.

Summary of findings

The ethnographic, multi-site approach of the study demonstrated that methamphetamine use had become entrenched and that methamphetamine was becoming an increasingly popular drug among a number of categories of users in diverse areas of the United States. Methamphetamine use was related to perceptions that it enhanced sexual, social and occupational performance; thus, it was often integrated into the social fabric and identities of users' lives. The use of methamphetamine was found to be less stigmatized and often cheaper than the use of other drugs. Methamphetamine was readily available, it was usually obtained through private distribution channels rather than on illicit drug markets on the street. Those features made methamphetamine a less visible drug in public contexts and made it more difficult for institutional sources of information and law enforcement to reach methamphetamine users.

There was considerable variation in user groups, user settings and patterns of use. The configurations of drug use characteristics and social contexts underpinned users' involvement in behaviours that posed various health and social risks. The preliminary findings of the study are presented below.

1. Variation in patterns of use and user groups

The user groups and associated patterns of use varied according to age, gender, sexual preference, ethnicity, occupation and residence. The sampling techniques used did not produce a representative sample of the user population, but did reflect much of the diversity in user groups and patterns of use. For example, although two thirds of the users studied were male, in some sites (such as Atlanta and Denver), rates of female users rivalled those of male users. Furthermore, the study showed that growing numbers of youth were becoming involved in methamphetamine use: one quarter of the study population was below the age of 24, most of them having used the drug for less than five years. Youth and young adult users displayed patterns characterized by both a rapid progression to daily use and more recreational, intermittent use, depending on the social context of use. That trend is particularly serious because many long-term users become involved in methamphetamine use in their youth, suggesting that initiation to methamphetamine use during youth may lead to lifetime use of the drug.

Methamphetamine was predominantly used by Whites (67 per cent), although there was an increase in methamphetamine use among Blacks and Hispanics. Those trends appear to reflect, in part, a shifting in the traffickers and distributors of methamphetamine, from largely White motorcycle gangs to local producers and small-scale distributors, and the involvement of Mexican criminal organizations and gangs in trafficking and distributing the drug.

There was also much variety in the user subgroups. For example, homosexual and heterosexual men at some sites (such as New York [18]) were generally young, educated and employed and rarely injected methamphetamine, while at other sites (such as Seattle [28]), they were older and unemployed and usually

injected methamphetamine and used it in sexual encounters; or, if the users were HIV-positive, they would use methamphetamine to help them in their daily functioning. Thus, some users were involved in chronic and highly addictive patterns of use, while others preferred intermittent and seemingly more controlled drug use.

The study showed that the aetiology and patterns of methamphetamine use are not equivalent in all categories of users or determined simply by the biochemical properties of the drug. Because of the impact of social contexts, biography and meaning in the use of the drug, especially among new and intermittent users [17], it is important that further research be carried out to examine specific categories of methamphetamine users in particular settings and related to particular practices of drug use. Such knowledge is critical to understanding variations in risk behaviour and to designing relevant and effective intervention programmes.

2. Health and social consequences of methamphetamine use

The multi-site study indicated that methamphetamine use presents serious public health concerns, including increased risk of: the transmission of HIV, hepatitis B and hepatitis C; reproductive problems for women; psychiatric comorbidity; and violence. It also has a number of consequences for the social and educational development of youth and the occupational and social functioning of chronic users. Furthermore, a number of subpopulations, such as heterosexual youth and young and HIV-positive homosexual and bisexual men, are especially vulnerable to initiating use of the drug. Health risks associated with the use of methamphetamine include the following:

(a) Intravenous use of methamphetamine is widespread and increasing in various social categories of users and areas in the United States. The multi-site study indicated that methamphetamine users engage in direct and indirect sharing of drug paraphernalia, and that the rate of syringe-sharing is often higher among methamphetamine injectors than among heroin injectors. Observations in the studies showed that methamphetamine injectors may inject the drug over long periods of time and in groups in which there is often inadvertent sharing of used syringes. Many users also have incomplete or inaccurate knowledge of modes of HIV transmission and lack strategies for risk reduction;

(b) For almost all categories of drug users, homosexual and bisexual males, heterosexual males and females and youths, methamphetamine is associated with unsafe sexual behaviour in the form of unprotected sex, sex with multiple partners and with strangers, commercial sex and experimental and prolonged sexual encounters. Those practices were distributed across the sites but also varied by social group. For example, methamphetamine use within sexual encounters was more common among homosexual and bisexual males than among females. Nonetheless, the frequency of sexual encounters and involvement with multiple partners was shown to be higher among methamphetamine injectors than among heroin injectors, and users of most other drugs. Moreover, alcohol use was com-

bined with drug use in sexual liaisons more frequently among methamphetamine users than among users of other drugs; that further heightened involvement in risky behaviour and the potential for transmitting and contracting blood-borne infections;

(c) The social groups that use methamphetamine are involved in contexts in which various subpopulations interact with one another during drug and sexual behaviours. Those subpopulations and mixing patterns vary by mode of administration—injection drug users using drugs with non-injection drug users; they also vary in drug-use interactions among younger and older cohorts of users, rural and urban users, homosexual and bisexual subpopulations and heterosexual subpopulations, and HIV-positive and HIV-negative individuals. Those social, drug and sexual mixing patterns, combined with frequent changes of sexual and drug partners, together with the synergistic effects of injection drug use and unsafe sex, were shown to be facilitated by the use of methamphetamine. The interaction of those behaviours and relationships placed those groups at higher risk of contracting blood-borne infections, regardless of the subpopulations and area involved;

(d) There appears to be an association between methamphetamine abuse and mental health problems. Methamphetamine users have reported using the drug to relieve depression, anxiety and personality disorders. They also use it to relieve symptoms of acquired immunodeficiency syndrome (AIDS). Methamphetamine use is considered to exacerbate predispositions to emotional and cognitive problems. That is a complex area that deserves rigorous, clinical and epidemiological assessment. However, sleep deprivation from continued methamphetamine use and the depressing effects of the process of withdrawal, or "coming down", from such use put a strain on the emotional and cognitive states of most users. Moreover, many chronic, long-term and high-volume methamphetamine users were found to be unemployed and to lack intimate relationships, indicating that, over time, the drug has deteriorating effects on the users' psychosocial adjustment and that the use of methamphetamine, like the use of other drugs, may lead to increased social isolation;

(e) Methamphetamine use has been associated with domestic violence between couples and within families and with organizational forms of violence related to its distribution. At some of the sites, there were reports in the media associating murders and violent subgroups (such as White supremacists) with methamphetamine use. However, studies focusing on the relationship between violence and methamphetamine have yet to be conducted. Nevertheless, as Mexican criminal organizations and gangs have expanded their involvement in methamphetamine trafficking and distribution, the level of violence associated with its distribution has increased;

(f) Female methamphetamine users are becoming increasingly common; in a number of areas, the proportion of female users is comparable to that of male users. Initial findings at two of the sites indicated that female users tend to suffer from reproductive problems, such as disturbed menstruation cycles and miscarriages, and experience problems in parenting. They were also found to be the

victims of domestic violence and to have proportionally more blood-borne infections. Those who become pregnant risk transmitting infections such as HIV to their children.

The above-mentioned health and social risks are closely related to particular forms and frequencies of methamphetamine use (such as chronic, high-volume and injection drug use) and the social contexts in which the drug is used and risk behaviours occur. Those factors are critical in planning and evaluating interventions for methamphetamine users.

3. Surveillance

The prevalence of methamphetamine is not adequately reported or monitored by local and national drug surveillance systems. Variations in the scale of manufacture of methamphetamine, its distribution through private channels and friendship networks and its association with social activities make it less publicly visible than other drugs. Since indicator surveillance systems operate in public spaces, innovative approaches are needed to track changes in the use of methamphetamine. In the multi-site study, it was possible to access various user subpopulations, to collect data through the use of qualitative methods and to use those sources to gain access to established community contacts. Mixed methods of qualitative contextual and quantitative monitoring in particular communities and with groups with whom researchers have ongoing relationships are necessary to be able to track the hidden and changing subpopulations of users and shifting trends in methamphetamine use. Since methamphetamine is not a typical "street" drug, new approaches are needed to gain access to users by expanding on established forms of street outreach and harm-reduction activities [50].

Conclusion

The multi-site study has used a strategy for linking intensive, qualitative studies, while allowing for the flexibility needed to access diverse user groups and collect meaningful social and public health information within particular settings. As the study suggests, the use of an ethno-epidemiological approach is well suited to identifying and monitoring emerging drug trends. The findings of the study also show that approach to be capable of providing crucial information for designing public health interventions for subpopulations within local and wider social and geographical settings.

References

1. B. Remberg, "Stimulant abuse: from amphetamine to ecstasy", *World Drug Report*, United Nations International Drug Control Programme (Oxford, Oxford University Press, 1997).

2. Community Epidemiology Work Group, *Epidemiologic Trends in Drug Use, Volume I* (Rockville, Maryland, National Institute on Drug Abuse, United States Department of Health and Human Services, 2000).
3. E. M. Gorman, P. Morgan and E. Y. Lambert, "Qualitative research considerations and other issues in the study of methamphetamine use among men who have sex with other men". In *Qualitative Methods in Drug Abuse and HIV Research*, National Institute on Drug Abuse, Research Monograph Series No. 157, E. Lambert, R. Ashbury and R. H. Needle, eds. (Rockville, Maryland, United States Department of Health and Human Services, 1995), pp. 156-181.
4. United States of America, Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, *Treatment Episode Data* (Rockville, Maryland, 1999).
5. M. D. Anglin and others, "History of the methamphetamine problem", *Journal of Psychoactive Drugs*, vol. 32, No. 2 (2000), pp. 137-141.
6. N. H. Harris and others, "Risk factors for HIV infection among injection drug users: results of blinded surveys in a drug treatment center, King County, Washington, 1988-1991", *Journal of Acquired Immune Deficiency Syndrome*, No. 6, 1993, pp. 1275-1282.
7. C. J. Reback and C. E. Grella, "HIV risk behaviours of gay and bisexual male methamphetamine users contacted through street outreach", *Journal of Drug Issues*, vol. 29, No. 1 (1999), pp. 155-166.
8. S. Pennell and others, "Meth matters: report on methamphetamine users in five western cities", National Institute of Justice Research Report (Washington, D.C., 1999).
9. United States of America, Office of National Drug Control Policy, *Methamphetamine Trends in Five Western States and Hawaii* (Washington, D.C., 1997).
10. M. A. Miller, "History and epidemiology of amphetamine abuse in the United States", *Amphetamine Misuse: International Perspectives on Current Trends*, H. Klee, ed. (Amsterdam, Harwood Press, 1997).
11. R. Anderson and N. Flynn, "The methamphetamine connection in northern California", *Amphetamine Misuse: International Perspectives on Current Trends*, H. Klee, ed. (Amsterdam, Harwood Press, 1997).
12. P. Morgan and J. E. Beck, "The legacy and paradox: hidden contexts of methamphetamine use in the United States", *Amphetamine Misuse: International Perspectives on Current Trends*, H. Klee, ed. (Amsterdam, Harwood Press, 1997).
13. M. J. Rotheram-Borus and others, "Going nowhere fast: methamphetamine use and HIV infection", *The Context of Drug Use among Drug Users and Their Sexual Partners*, R. Battjes, Z. Sloboda and W. C. Grace, eds. (National Institute on Drug Abuse, Research Monograph Series, No. 143, Rockville, Maryland, 1994), pp. 155-183.
14. E. M. Gorman and others, "Speed, sex, gay men and HIV: community and ecological perspectives", *Medical Anthropological Quarterly*, vol. 11, No. 4 (1997), pp. 505-515.
15. R. Cabaj and E. M. Gorman, "Sexual minorities and substance abuse: a profile", *A Provider's Guide to Substance Abuse Treatment for Sexual Minorities* (Rockville, Maryland, Substance Abuse and Mental Health Services Administration, United States Department of Health and Human Services, 2000).
16. M. J. Rotheram-Borus, T. Mann and B. Chabon, "Amphetamine use and its correlates among youth living with HIV", *AIDS Education and Prevention*, vol. 11, No. 3 (1999), pp. 232-242.

17. M. H. Agar, "Preface", *National Institute on Drug Abuse: Methamphetamine Multisite Project* (Seattle, Washington, 1998).
18. M. Clatts, L. A. Goldsamt and P. Case, "Pilot ethnographic study of 'speed' use in New York City", *National Institute on Drug Abuse: Methamphetamine Multisite Project* (Seattle, Washington, 1998).
19. M. H. Agar, "Recasting the 'ethno' in epidemiology", *Medial Anthropology*, No. 16, 1996, pp. 391-401.
20. M. Clatts, D. Welle and L. A. Goldsamt, "Reconceptualizing the interaction of drug and sex risk among MSM speed users: toward an ethno-epidemiology", *AIDS and Behavior*, vol. 5, No. 2 (2001), pp. 115-130.
21. R. Battjes, Z. Sloboda and W. C. Grace, eds., *The Context of Drug Use among Drug Users and Their Sexual Partners*, National Institute on Drug Abuse, Research Monograph Series, No. 143 (Rockville, Maryland, 1994).
22. T. W. Boyd, "Drug trafficking in north Texas and Oklahoma", *Epidemiological Trends in Drug Abuse, Community Epidemiology Work Group, Volume II* (Rockville, Maryland, National Institute on Drug Abuse, United States Department of Health and Human Services, 1996).
23. G. R. King and E. H. G. Ellinwood, "Amphetamines and other substances", *Substance Abuse: a Comprehensive Textbook*, J. H. Lowinson (Baltimore, Maryland, William and Wilken, 1996).
24. B. T. Burton, "Heavy metal and organic contaminants associated with illicit methamphetamine production", *Methamphetamine Abuse: Epidemiologic Issues and Implications*, National Institute on Drug Abuse, Research Monograph Series, No. 115, M. A. Miller and N. J. Kozel, eds. (Rockville, Maryland, 1991).
25. T. Ernst and others, "Evidence for long-term neurotoxicity associated with methamphetamine abuse: a 1H MRS study", *Neurology*, vol. 54, No. 6 (2000), pp. 1344-1349.
26. L. Seiden, "Neurotoxicity of methamphetamine: mechanisms of action and issues related to aging", *Methamphetamine Abuse: Epidemiologic Issues and Implications*, National Institute on Drug Abuse, Research Monograph Series, No. 115, M. A. Miller and N. J. Kozel, eds. (Rockville, Maryland, 1991).
27. E. M. Gorman, "A tale of two epidemics: HIV and stimulant use", *FOCUS: a Guide to AIDS Research Counseling*, vol. 13, No. 4 (1998), pp. 1-3.
28. E. M. Gorman, "The Crystal Palace in the Emerald City: ethnographic perspectives on methamphetamine use in Greater Seattle", *National Institute on Drug Abuse: Methamphetamine Multisite Project* (Seattle, Washington, 1998).
29. E. M. Gorman and R. Carroll, "The interface of substance abuse and HIV: considerations regarding methamphetamines and other 'party drugs' for nursing practice and research", *Journal of Nursing and AIDS Care*, vol. 11, No. 2 (2000), pp. 51-62.
30. L. Wermuth, "Poverty and methamphetamine abuse: a study of AIDS risk behaviors among rural California women", *Journal of Poverty*, vol. 3, No. 1 (1990), pp. 25-45.
31. E. M. Gorman and others, "Female methamphetamine users: ecological and community social work perspectives", *Journal of Social Work Practice and the Addictions*, forthcoming.
32. R. T. Trotter and J. J. Schensul, "Methods in applied anthropology", *Handbook of Methods in Cultural Anthropology*, H. R. Bernard, ed. (Walnut Creek, California, Altamira Press, 1998).

33. P. Willis, *Common Culture: Symbolic Work at Play in the Everyday Cultures of the Young* (Boulder, Colorado, Westview Press, 1990).
34. C. E. Sterk, K. W. Elifson and L. Jenkins, "Old timers and newcomers: an ethnographic snapshot of the methamphetamine scene and its public health consequences in metropolitan Atlanta", *National Institute on Drug Abuse: Methamphetamine Multisite Project* (Seattle, Washington, 1998).
35. S. Koester and C. Christianson, "Speed: a preliminary ethnographic report from Denver", *National Institute on Drug Abuse: Methamphetamine Multisite Project* (Seattle, Washington, 1998).
36. D. R. Gibson, R. Andreson and N. Flynn, "Epidemiology and public health implications of methamphetamine use in the Sacramento region", *National Institute on Drug Abuse: Methamphetamine Multisite Project* (Seattle, Washington, 1998).
37. E. M. Gorman and others, "Club drug and poly substance use among men who have sex with other men: implications for outreach, prevention and education", *Journal of Gay and Lesbian Social Services*, forthcoming.
38. S. M. Scrimshaw and E. Hurtado, *Rapid Assessment Procedures for Nutrition and Primary Health Care: Anthropological Approaches to Improving Programme Effectiveness* (Los Angeles, California, UCLA Latin American Center, 1987).
39. N. Tashima and others, "The community identification (CID) process: a discovery model", *Qualitative Health Research*, vol. 6, No. 1 (1996), pp. 23-48.
40. H. B. Schwartzman, *The Meeting: Gatherings in Organizations and Communities* (New York, Plenum Press, 1989).
41. M. H. Agar, *Speaking of Ethnography* (Newbury Park, California, Sage Publications, 1986).
42. P. Bourdieu, *Outline of a Theory of Practice* (Cambridge, Cambridge University Press, 1979).
43. D. Hebidge, *Subculture: the Meaning of Style* (New York, Routledge Press, 1979).
44. J. C. Johnson, "Selecting ethnographic informants", *Qualitative Research Methods, Volume 22* (Newbury Park, California, Sage Publications, 1990).
45. J. K. Watters and P. Biernacki, "Targeted sampling: options for the study of hidden populations", *Social Problems*, vol. 36, No. 4 (1989), pp. 416-430.
46. R. B. Rothenberg, "Commentary: sampling in social networks", *Connections*, vol. 18, No. 1 (1995), pp. 105-111.
47. H. R. Bernard, *Research Methods in Anthropology* (Walnut Creek, California, Altamira Press, 1995).
48. J. Spradley, *The Ethnographic Interview* (New York and Orlando, Florida, Harcourt Brace Jovanovich College Publishers, 1979).
49. M. H. Agar, *The Professional Stranger: an Informal Introduction to Ethnography* (New York, Academic Press, 1980).
50. W. W. Wiebel and others, "Outreach to IDUs not in treatment", *Handbook on Risk of AIDS among Injection Drug Users and Sexual Partners*, B. Brown and G. Beschner, eds. (Westport, Connecticut, Greenwood Press, 1993).

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

Human immunodeficiency virus (HIV) infection rates and heroin trafficking: fearful symmetries

C. BEYRER

Director, Johns Hopkins University Fogarty AIDS International Training and Research Program, Baltimore, Maryland, United States of America

ABSTRACT

There is mounting evidence that the spread of the human immunodeficiency virus (HIV) is associated with heroin trafficking routes. The relationship between the two is best illustrated by the routes leading from the two primary regions for the illicit opium poppy cultivation and heroin manufacture: the Golden Triangle of South-East Asia and the Golden Crescent of Central Asia. The producers in the Golden Triangle are the Lao People's Democratic Republic and Myanmar, and those in the Golden Crescent are Afghanistan and Pakistan. Together, those States accounted for perhaps 80-90 per cent of the world heroin supplies in 2002. HIV outbreaks resulting from unsafe injection practices among injecting drug users (IDU) in trafficking zones have been documented in Myanmar itself, in Belarus, China, India, Indonesia, the Islamic Republic of Iran, Malaysia, Pakistan, the Russian Federation, Tajikistan, Thailand, Ukraine, Uzbekistan, Viet Nam and in several States in Eastern Europe. Heroin trafficking in those States has led to serial epidemics: first of heroin use, then of injection, then of blood-borne pathogens, including hepatitis C and HIV. Ethnic and trade relationships in heroin trafficking zones appear to facilitate such epidemics, as does drug testing by petty traders in market nodes. Policy responses, or the lack thereof, have increased the vulnerability of users and their communities, as have the limited drug treatment options available in those zones. While "supply-side" approaches, including interdiction and policing, are likely to continue, those HIV epidemics will require improved drug treatment, access to HIV prevention services, including harm reduction, and new approaches to the prevention of HIV in areas where heroin trafficking occurs.

Keywords: heroin; trafficking; HIV infection; IDUs; Afghanistan; Burma; Lao People's Democratic Republic.

Introduction

There is growing evidence that the epidemic spread of human immunodeficiency virus (HIV) infection occurs in tight and complex relationship to heroin-trafficking

routes [1-3]. The relationship is most clearly illustrated by the routes leading from the two primary regions for illicit opium poppy cultivation and heroin manufacture: the Golden Triangle of South-East Asia and the Golden Crescent of Central Asia. The mechanisms that lead to those HIV outbreaks are just beginning to be understood, as are the special vulnerabilities of communities in trafficking zones. The principal heroin producers in the Golden Triangle are the Lao People's Democratic Republic and Myanmar and those in the Golden Crescent are Afghanistan and Pakistan. Together, those States accounted for perhaps 80-90 per cent of the world heroin supplies in 2002 [4]. While those four countries generate most of the world's heroin, the HIV epidemics resulting from unsafe heroin injection practices have largely been seen in the neighbouring countries or in countries of destination further afield. The spread of HIV is well documented in countries in and around the Golden Triangle, such as in Myanmar itself, in China, India, Malaysia, Thailand, Viet Nam and, more recently, in Indonesia [5-10]. For countries in and around the Golden Crescent, the data are only now emerging and the HIV epidemics are much newer: the epidemic spread of HIV, hepatitis C or both appears to be under way in Belarus, India, the Islamic Republic of Iran, Pakistan, the Russian Federation, Tajikistan, Ukraine, Uzbekistan and several States in Eastern Europe [7 and 11-15]. In virtually all the studies that have investigated HIV and hepatitis C virus infection among injecting drug users (IDU) in those regions, hepatitis C virus has been shown to be far commoner. The prevalence of hepatitis C virus among IDU generally reaches 90 per cent or higher, which is a function of the very high transmissibility of this agent through parenteral exposure. However, it is the subtyping of HIV that has proven more useful, thus far, as an investigative tool to understand trafficking.

The third important area for illicit poppy cultivation is the New World, in Colombia, Mexico and Peru. While significant for the markets of North and South America, it will not be discussed in the present article as it is less well understood in terms of interactions between trafficking and HIV spread in the production zones.

Licit poppy cultivation for pharmaceutical opiate derivatives such as morphine, codeine and demerol is centred in Tasmania, Australia, which accounts for about 50 per cent of licit world production, and in India and Turkey and has not been associated with heroin manufacture or trafficking, or with the spread of blood-borne infections. Opiate derivatives remain an important class of analgesics and are widely and generally safely used, with minimal public health effects. It is the illicit nature of criminal production and distribution and the rapid uptake of heroin use, injection and the unsafe use of equipment by young people in vulnerable communities along the trafficking routes and in destination markets further afield that has led to the fearful symmetry of heroin trafficking and the spread of HIV.

The present article seeks to clarify the degree to which the existing data demonstrate the relationship between heroin trafficking and HIV and to establish how much is known about the mechanisms that spread HIV. It also seeks to describe the tools for understanding and responding to that relationship that are

available to researchers, policy makers, clinicians and others. Finally, the article attempts to explain why societies from Ukraine to Viet Nam have been so vulnerable to the interactions between heroin trafficking and the spread of HIV and to examine what can be done to reduce the harm resulting from them.

Injection drug use and the spread of HIV

In 2000, an international research team published the findings of its investigations of the causes of HIV infections along the four heroin trafficking routes leading from the Lao People's Democratic Republic and Myanmar to China, India, Thailand and Viet Nam [2, 16]. The findings had been obtained through the use of molecular epidemiology, Land Remote Sensing Satellite (Landsat) technology, qualitative research methods and epidemiologic review. By using deoxyribonucleic acid (DNA) fingerprinting technology, the team was able to show that HIV viruses from the blood of infected IDU could help trace heroin routes. Heroin users and petty traders helped the research team to understand how heroin use spreads in communities and pinpointed the key roads, villages, towns and cities through which heroin from the Golden Triangle was moving, leaving the legacies of addiction and acquired immunodeficiency syndrome (AIDS). Since the publication of those findings, several other groups have independently investigated those zones, confirming the general hypotheses made by the team and documenting the further spread of infections and their impact on communities in China, Myanmar and Viet Nam [3, 6, 17]. The fearful symmetry of heroin trafficking and the spread of HIV can perhaps best be illustrated by the HIV epidemic in Yunnan Province of China, which is east of Myanmar and the first destination for overland exports of heroin from Myanmar to the rest of China. The farthest province from China's booming coastal cities, Yunnan has the highest HIV infection rate in China [9]. The outbreak began in ethnic minority communities in three mountain districts along the China-Myanmar border in the early 1990s, notably among the Kachin and Wa ethnic groups [18]. As heroin spread among the young people of Yunnan and a rapid transition to injecting took place, there was a predictable rise in HIV infection. Equally predictable was the subsequent spread of infection to the non-IDU sexual partners, wives and children of the largely young adult male IDU population. Yunnan now has the most mature HIV epidemic in China. Intense security and interdiction activities (supported partly by the United States of America) have failed to control the heroin trade. But it could be argued that at the current stage of the spread of the HIV epidemic in Yunnan, the virus has moved well beyond the IDU population. A similar situation obtains in Manipur State, north-east India, on the western border of Myanmar: the early outbreak of HIV infection among IDU there has led to a disseminated HIV epidemic and the highest HIV prevalence by province in India [19]. The present article will attempt to establish whether or not such situations are typical.

Observation of the rate of HIV and AIDS infections in Belarus, China, Kazakhstan, the Russian Federation, Tajikistan and Ukraine or further east in Indonesia, the Islamic Republic of Iran, Malaysia and Viet Nam reveals an

epidemiologic picture of HIV in 2002 that is striking. What these diverse States have in common is that in each, the majority of reported HIV infections and AIDS cases in 2001 were attributed not to sexual transmission, which is the predominant mode in Africa, but to parenteral infection, exposure through needle-sharing behaviours among IDU [10]. While the numbers of IDU in any one of those States may not be large on a population basis, those States represent enormous young populations, many of which have rapidly rising substance abuse rates. One example is Viet Nam, a country with a population of over 78 million, where IDU accounted for 88 per cent of all reported HIV infections in 2000 and where heroin trafficking from the growth zones of the Lao People's Democratic Republic has led to a dramatic increase in use among young people [3].

In other regions and countries, the majority of HIV infections is not among IDU, but IDU spread has played an important role in HIV epidemics. HIV infections among IDU were reported to the World Health Organization by 52 countries in 1992; the number of reporting countries grew to 114 by the year 2000, underscoring the global nature of IDU risk and spread [20]. IDU-related outbreaks have also played a key role in the dynamics of the spread of HIV, notably in introducing HIV into populations, as shown by the early epidemiology of HIV in Thailand and Viet Nam, where IDU were the first group in which HIV outbreaks occurred [3, 21]. Such outbreaks have also played an important role in the dissemination and dispersion of novel HIV-1 subtypes. One example is the recent explosive spread of the HIV-1 subtype A virus in the Russian Federation and Ukraine and the outbreak of a B/C recombinant HIV strain that has been documented among IDU in southern and western China [1, 22].

HIV infection rates in South and South-East Asia

IDU played a crucial role in the spread of HIV in South and South-East Asia in every country with a significant HIV epidemic except, perhaps, Cambodia [2]. States experiencing epidemics that were initially IDU-related or are predominantly IDU-related include China, India, Malaysia, Myanmar, Thailand and Viet Nam [2]. IDU were the first group in which the epidemic spread of HIV was detected in China, Indonesia, Malaysia, and Myanmar [10] and in all those outbreaks trafficking in heroin from the Lao People's Democratic Republic or Myanmar, or both, was involved [2].

The HIV epidemic in Thailand is among the best documented in Asia. The epidemic spread of HIV in that country was first detected among IDU in Bangkok in 1988 [21]. It was an explosive outbreak with clear links to incarcerated IDU and occurred initially among low-income, ethnic Thai, male urban residents. However, HIV spread rapidly among Thai IDU nationwide and within a year HIV rates of 20-40 per cent were the norm. The IDU-related HIV epidemic was followed by a heterosexual outbreak of HIV that was larger and involved many more people. However, while the rate of infection of heterosexuals and other groups at risk declined after the period 1995-1996, that of IDU did not [23].

Similarly, in Malaysia and Viet Nam, IDU were the first group in which the epidemic spread of HIV was detected. However, unlike in Thailand, IDU have remained the predominant risk group affected by HIV in Malaysia and Viet Nam, accounting for roughly 60-70 per cent of cumulative infections in those two countries by 2000 [20]. The heroin in Malaysia appears to be transported by sea along that country's long coast on the Andaman Sea [4]. The Lao People's Democratic Republic was identified as the likely source of this heroin and molecular-typing data confirm that HIV infections have spread north from this zone in Viet Nam into the neighbouring Guangxi Province of China [5]. China's emerging epidemic remains overwhelmingly due to needle-sharing among IDU and the three most HIV-affected provinces of China (in order of prevalence, Yunnan, Xinjiang and Guangxi) have all experienced IDU-related outbreaks along the major heroin routes [1, 9].

Myanmar is a somewhat different case. It is the region's major producer and a major consumer of heroin [8]. Trafficking within Myanmar from the heroin manufacturing areas in the Shan and Wa hills in the far north-east of Myanmar has led to a national outbreak of HIV. The United Nations International Drug Control Programme (UNDCP) and the Ministry of Health of Myanmar identified drug use rates among township adults of 2-25 per cent in 1995, one of the highest rates worldwide [24]. IDU were the first group in which HIV was identified, in 1989, and remain at extraordinary risk, with HIV prevalence rates of 60 to 95 per cent nationwide [25]. It is unclear how heroin is moved within Myanmar, but it is known that a major city in the trade is Mandalay, the largest city in northern Myanmar. Petty traders from Manipur State in India travel inland across Myanmar from India to buy high-grade No. 4 heroin in Mandalay and trucks carrying heroin to India leave from Mandalay as well [2]. HIV rates among IDU in Mandalay have been consistently in the 60-80 per cent range since at least 1995. Mandalay is in a region that is too hot and at too low an altitude for poppy cultivation.

The outbreaks of HIV infection among IDU in South-East Asia have several features in common. First, they have been explosive: HIV prevalence among Bangkok IDU went from 2 to 40 per cent in six months in 1989. Second, they have been transnational: both China and India have had their highest prevalence zones along their borders with Myanmar (Yunnan Province and Manipur State, respectively). Third, they have led to further spread among non-injecting populations, initially the sexual partners of IDU, as has been documented in China, India and Thailand. Fourth, they have proven difficult to control owing to government policies on injection drug use, the status of drug treatment in the affected States and the limited HIV prevention measures targeting IDU.

The Golden Crescent: HIV infection rates and Central Asian heroin

Much less is known about heroin and HIV epidemics in the Golden Crescent than is known about those in South and South-East Asia. For most States affected by heroin from Afghanistan and Pakistan, HIV spread is a more recent event and

many have little data or research capacity. However, the little that is known suggests another region of fearful symmetry.

Several tools can be used to measure poppy production, but arguably the most accurate is Landsat satellite technology, which measures crop densities [4]. United States intelligence agencies have used Landsat to assess poppy cultivation, estimate opium base harvests and calculate heroin yields (10 kilograms of opium base gives roughly 1 kilogram of refined heroin). In 1996, after the establishment of Taliban rule, the estimated yield in Afghanistan was 200 tons [4]. By 1999, Afghanistan was manufacturing 450 tons of heroin per year and had become the world's largest single manufacturer in a multibillion-dollar industry. Poppy growing appeared to cease in 2000 after the leader at that time, Mullah Muhammed Omar, delivered an edict on the subject, but stockpiled heroin reserves held by producers and traffickers apparently ensured that the supply was maintained despite the ban on poppy cultivation. In 2002, the new Administration and its allies acknowledged that a reduction in poppy cultivation could only be achieved through a long-term process of agricultural reform and development and the extension of government control across the vast rural areas of the country. For the short to medium term at least, Afghanistan will remain a significant producer. The second largest grower in the region, Pakistan, produces about 20 tons of heroin a year, roughly equal to the production of the Lao People's Democratic Republic and mostly in the remote tribal zones along the Afghan border in the North-West Frontier Province [4]. Those areas are only marginally under federal control, very underdeveloped and likely to remain dependent on poppy cultivation for some years.

The HIV-related repercussions of heroin exports from Afghanistan and Pakistan are only now beginning to be understood, as nascent HIV epidemics take hold in a region for which data have been sparse, but that was thought to have been relatively spared from HIV. The Islamic Republic of Iran and Pakistan appear to be two of the major overland routes for the trafficking in heroin from Afghanistan [7]. While HIV prevalence is low in both States, Pakistan had an estimated 3 million heroin addicts in 2000 and has suffered great social harm as a consequence [11]. The Islamic Republic of Iran led the world in 1999 in narcotics seizures by volume [4]. The country also has an enormous epidemic of heroin use among its young people [7]. The Iranian Government is deeply concerned about this and it was a primary source of tension and border conflict while the Taliban regime was in control in Afghanistan. A nascent epidemic of HIV among Iranian IDU appears to have begun in the period 2000-2001, with recent reports of very high rates of HIV infection among incarcerated IDU in Tehran (up to 67 per cent in one facility) [7].

The countries most affected by HIV are the Russian Federation and the two former Soviet republics of Belarus and Ukraine [12, 14]. In its *Report on the Global HIV/AIDS Pandemic, 2000*, the Joint United Nations Programme on HIV/AIDS (UNAIDS) identified those three States as having the fastest-growing HIV epidemics in the world [26]. More than 75 per cent of all infections in the Russian Federation and its neighbours in 2000 were owing to injecting drug use

[12, 13, 27]. The far east of the Russian Federation has been particularly affected. The Irkutsk region of Siberia, around Lake Baikal, has the highest rate of HIV infection in the Russian Federation after Moscow and, again, more than 80 per cent of the HIV infections reported in Irkutsk have been among IDU [28]. Kazakhstan, too, has seen a recent outbreak of drug use and HIV infection, although it is unclear whether the source of the trafficking route is China to the east or Afghanistan to the south.

Poppy cultivation and politics

While the HIV epidemic is a new challenge to the Russian Federation, the trafficking connections of the Golden Crescent are not new, having existed during the long Afghanistan war with the former Union of Soviet Socialist Republics (USSR), when poppy cultivation by the Mujahidin was tolerated by the West because the anti-Soviet forces had no other exports comparable to heroin in terms of value and ease of transport [29]. The consequent high rates of heroin use and addiction among Soviet forces engaged in the Afghan conflict were a predictable outcome and helped undermine support for the war among those troops, their families and Soviet citizens.

The poppy farmers of Afghanistan are largely subsistence farmers who sell opium as a cash crop to supplement minimal incomes. As in the Golden Triangle region, the real profits of heroin come not from farming, but from trafficking and it is among the trafficking networks that real revenues accrue [29]. But part of the legacy of war is the local expertise in poppy cultivation and the production and sale of narcotics. The Afghan war, which the Soviets lost, appears to have brought heroin first to the dispirited troops and then to Moscow. Trafficking links may therefore be a legacy of the long struggle of the people of Afghanistan, although that remains speculation.

Another heroin-related epidemic is currently being experienced by China. The Xinjiang Uighur Autonomous Region is the only Muslim majority region of China. Xinjiang shares borders with Afghanistan, Kazakhstan and the Russian Federation (Siberia) and is linked to the rest of China by the Silk Road. It also has China's second highest rate of HIV infection by province, after Yunnan in the far south [20]. More than 78 per cent of the HIV infections in Xinjiang are owing to injection drug use involving heroin. Tragically, more than 90 per cent of IDU in the two largest cities in Xinjiang are ethnic Uighurs, which means that the HIV infections in this large province are largely among young Muslims [9].

Belarus, China, the Islamic Republic of Iran, Kazakhstan, Pakistan, the Russian Federation and Ukraine are all experiencing outbreaks of heroin use among their young people and all now appear to have HIV epidemics related to that use. Heroin exports from the Golden Crescent are at the root of these complex new problems. These are regional challenges, but they point to a global problem that ties the Golden Crescent to the Golden Triangle: illicit heroin revenues. In 2000, Afghanistan was the world's poorest State, on paper, and Myanmar was designated a "least developed country" by the United Nations. Afghanistan was

almost entirely dependent on donor aid in 2002 and had essentially no foreign reserves, a bankrupt treasury and limited licit exports. The economics of the drug trafficking networks based in the Golden Crescent are not known, but it is known that taxes on poppy farmers and protection money from drug traffickers were among the main sources of revenue for both the Taliban and the Northern Alliance before the interim Government led by Hamid Karzai came to power. In both Afghanistan and Myanmar, heroin has enabled the purchase of weapons on the black market and funded militias, insurgencies and crime [4]. Afghanistan has the potential to grow other crops, including grain and orchard fruits, but those would require irrigation systems and access to markets, which remains a huge challenge for much of the country. In Myanmar, the poppy-growing regions have been at war with the central Government virtually since the departure of the British after the Second World War [30]. In order to achieve a reduction in the opium supply from those regions, viable alternative economies for the rural poor will need to be established, which will take time, and sustained donor investment, and the nurturing of stable, functioning civil societies will be necessary. If Afghanistan were once more to descend into civil strife and warlordism, it is likely that heroin production would revive. Indeed, as in Myanmar, it is in the interests of the narcotics cartels and the corrupt leaders that they have supported that civil society should fail; a chilling reality, given the wealth and power that heroin revenues have already generated.

Mechanisms on the ground

Before discussing policy responses to the interactions of heroin and HIV, it would be useful to consider some mechanisms of those interactions which bear on the spread of HIV and how best to curtail it. One obvious feature shared by all of the primary trafficking zones out of the Triangle and the Crescent is geographic: overland heroin is moved almost exclusively across remote border regions, generally mountain and forest zones adjacent to the hills where poppy will grow. The illegal and clandestine nature of this industry demands such remote areas. Indeed, as a former director of UNDCP has pointed out, there are very few regions remote and lawless enough to support a major heroin industry. It is surely no coincidence that it is Myanmar, secretive, closed and ruled by junta; the isolated Lao People's Democratic Republic under its Communist Party; Afghanistan with its decades of strife; and the tribal zones of Pakistan that became the world's leaders in heroin production: these are almost the only places in the world closed enough to sustain the heroin industry.

A second shared feature is ethnicity. Those areas generally have in common populations who are ethnic minorities or tribal groups, or both [30]. In South-East Asia they are virtually all hill-dwellers, whereas the majority populations of the Lao People's Democratic Republic, Myanmar, Thailand and Viet Nam are lowlanders and rice cultivators [30]. One especially important factor is that those groups tend to straddle national borders. So in Yunnan Province of China, both sides of the China-Myanmar border lands are farmed by ethnic Kachin and Wa, and family, language, and trade links long pre-date heroin trafficking. Another example is the

ethnic Manipuris of Manipur State in India, who are Tibeto-Burmans like the majority population of Myanmar, not ethnic Indians, and can move easily into Myanmar to access the heroin markets in Mandalay [31].

For HIV to spread along trafficking routes, local people have to use the drugs. Qualitative work in China, India, Myanmar and Viet Nam has suggested a direct mechanism for the "exchange" of HIV-1 subtypes (known as clades or strains). This mechanism relies on the fact that many petty traders in the region are also users, who support their own habits by purchasing and selling small amounts of heroin. In at least four States, it is known that those petty traders typically self-test heroin purity by injecting themselves. Since travelling across those zones with injecting equipment is an obvious sign of intent to use drugs, they rarely have their own equipment. On the border between China and Viet Nam, for example, traders typically cross the mountains from China, stay the night with their contacts in Viet Nam and share drugs and equipment before making purchases. The very low genetic diversity of strains in this region suggests rapid spread of only one viral subtype, a molecular feature favoured by this kind of direct spread. Major traffickers moving heroin in consignments weighing hundreds of kilograms or more have very different ways to move their product, including trucking, sea and air routes. But it is likely that the spread of HIV in overland regions occurs on a more local person-to-person basis. A recent report from Yunnan Province found that 75.9 per cent of a large series of IDU in south-eastern Yunnan were of Han Chinese ethnicity which led the authors to conclude that the epidemic in Yunnan was no longer confined to non-Han ethnic minorities [32].

A fourth mechanism is also likely to be important, though somewhat variable. Along at least some of the major trafficking routes, services for truckers have developed, which, in addition to fuel, food and lodging, often include sex services. In South-East Asia, those sex services generally consist of roadside brothels, karaoke parlours, bars and so forth. In Central Asia they may be less apparent, but still available, or may have young male sex workers, as in the trucking industry in Pakistan. Those border zone sex service venues can overlap with drug trafficking and provide another mechanism by which HIV could spread where heroin and other contraband are moved. On the borders between China and Myanmar and between Myanmar and Thailand, women and girls are trafficked on the same routes, and indeed by some of the same trafficking networks, as heroin [33].

The interaction of heroin trafficking and sex industry-related HIV risks can also be found in the special economic zone of Pingxiang on the highway and train crossing from Viet Nam to China [1]. Pingxiang was one of the first Chinese cities to experience a rapid HIV epidemic among IDU and molecular work has confirmed the cross-border nature of this epidemic [5]. But Pingxiang also has a booming sex trade on the Chinese side of the zone. As many as 19 separate brothels in a four-street radius in the trucking zone in Pingxiang were counted in 2000, each with 10-30 women and girls working in it. HIV rates had remained low among those women until 2001, though there now appears to be increasing prevalence [34]. In settings like Pingxiang, sex workers and their clients in border and trafficking zones may be key "bridge" populations from IDU to wider networks of people at sexual risk.

Policy responses

IDU outbreaks associated with heroin trafficking have proven difficult to prevent or control. In the major production zones and in the wider affected regions, treatment and prevention programmes for drug use were limited before the spread of HIV [35]. That remains largely the case: across the whole of Asia the only place where evidence-based heroin treatment and methadone maintenance therapy are available on demand to drug users is the Hong Kong Special Administrative Region of China. This is tragic, given that there is a large and growing international evidence base for success in prevention of HIV infection and other blood-borne diseases among IDU [36]. While the majority of published reports have been from the developed world, principally Australia, North America and Western Europe, there have been several reports of pilot projects and successful programmes in Asia, including reports from India, Nepal, Thailand and Viet Nam. Much of that work has focused on harm reduction and needle and syringe exchange programmes, the basic tools of most reported interventions. In 1998, the *Journal of Substance Use and Misuse* published its "Bibliography on Syringe-Exchange References", which included several hundred published reports on those interventions and the debates that they generated [37].

Successful needle exchange programmes have been conducted in Australia, the Netherlands and the United Kingdom of Great Britain and Northern Ireland. In the largest analysis published, the incidence of HIV increased by about 6 per cent per year in 52 cities without needle exchange programmes and decreased by 5.8 per cent per year in 29 cities with needle exchange programmes [38]. The New York City needle exchange programmes have been studied in prospective cohorts: lower rates of incident HIV infection were documented among IDU using needle exchange programmes (1.4-1.6 per cent per year) than among those who did not attend needle exchange programmes (5.3 per cent per year, 95 per cent confidence interval: 2.4 to 11.5) [39]. Long-term methadone maintenance therapy has been shown to reduce HIV risk behaviours, in particular needle use, and there is strong evidence that methadone maintenance therapy prevents HIV infection among IDU.

Where harm reduction and methadone maintenance therapy are available, as they were to many IDU in the United States HIV Network for Prevention vaccine preparedness studies, sero-incidence can be low [40]. In that study, HIV incidence among homosexual and bisexual men between 1995 and 1997 was measured at 1.55 per 100 person years, while among male IDU the rate was 0.38 per 100 person years.

The Thai Working Group on HIV/AIDS Projections has recently published projected scenarios for the Thai epidemic [41]. They found that a decline in needle-sharing from 20 to 10 per cent among Thai IDU would avert 21,774 new infections by 2006 and 81,761 infections by 2020. That would constitute the single largest number of infections averted for any one intervention strategy. By 2006, roughly 3,800 of the expected 22,000 infections nationwide would be averted by that intervention alone. Viet Nam has reported on the feasibility of needle

exchange programmes and on pilot needle exchange programmes in the cities of Hanoi and Ho Chi Minh [42]. While they did not measure impact, they were able to conclude that needle exchange programmes were feasible, but that they required acceptance by the community and by the police to be sustained. Needle exchange programmes have also been implemented in India, notably in New Delhi and Manipur State, where high rates of IDU behaviour are common.

Taken together, those studies all support the contention that harm reduction and needle exchange programmes are effective prevention tools, and that they might have an effect on heroin-related epidemics in trafficking zones. Yet those approaches have been little used in the fight against HIV/AIDS.

It is difficult to imagine a public health tool with reasonable evidence of efficacy that has generated as much debate as have prevention programmes for IDU. A review of the literature suggests three principal problems with the implementation of harm reduction approaches and needle exchange programmes [43, 44]. First, they have repeatedly been seen as condoning or facilitating injecting drug use, making them politically unpopular outside the prevention community. Second, they have faced legal, security and policy challenges, since they require "safe" domains of interaction with active IDU. A third challenge, where needle exchange programmes have been implemented, is the coverage rates of needle exchange programmes for IDU populations. One exception to the lack of programmes in affected areas has been the rapid implementation of needle exchange programmes in the Russian Federation and the Commonwealth of Independent States, established in partnership with the Open Society Institute of the George Soros Foundation, Médecins Sans Frontières, the ministries of health of many affected nations and the United States Agency for International Development. The Open Society Institute has supported the establishment of more than 140 needle exchange programmes in the Russian Federation alone and that effort has the potential to reach levels of coverage which might control HIV among IDU.

Conclusions

Individuals, communities and countries that have the misfortune to be on major heroin-trafficking routes faced multiple epidemics in 2002. Those epidemics began with heroin use, heroin injection and then HIV infections. While the clear long-term goal for all the States involved is to be free of drug trafficking, the realities of the current political and development situations of the major producers, most notably Afghanistan and Myanmar, suggest that narcotics-based economies will be with the world for some time to come. In the short to medium term, a public-health-based approach would be to minimize the health impacts of heroin trafficking by working with affected communities. Such approaches could include reducing heroin addiction through improved treatment and support for IDU and reducing the spread of HIV among those who continue to inject by expanding harm reduction and needle and exchange programmes. The prevention of spread beyond IDU suggests that that may be critically important to the prevention of wider epidemics of HIV/AIDS. A clear priority for further research and

programmes are the front-line Central Asian States and around the Golden Crescent: the Islamic Republic of Iran, Kazakhstan, Kyrgyzstan, Pakistan, Tajikistan, Turkmenistan and Uzbekistan. Those must be considered States that are at very high risk of an explosive spread of HIV in the coming years and that could benefit from the programmatic and research experience that have elucidated the heroin and HIV interactions of the Golden Triangle.

References

1. X. F. Yu and others, "Emerging HIV infections with distinct subtypes of HIV-1 infection among injection drug users from geographically separate locations in Guangxi Province, China", *Journal of Acquired Immune Deficiency Syndromes*, vol. 22, No. 2 (1999), pp. 180-188.
2. C. Beyrer and others, "Overland heroin trafficking routes and HIV-1 spread in South and Southeast Asia", *AIDS*, vol. 14, 2000, pp. 75-83.
3. V. M. Quan and others, "HIV in Vietnam: the evolving epidemic and the prevention response", *Journal of Acquired Immune Deficiency Syndromes*, vol. 25, No. 4 (2000), pp. 360-369.
4. United States of America, Department of State, Bureau of International Narcotics and Law Enforcement Affairs, *International Narcotics Control Strategy Report* (Washington, D.C., 2002).
5. K. Kato and others, "Genetic similarity of HIV type 1 subtype E in a recent outbreak among injecting drug users in northern Vietnam to strains in Guangxi Province of southern China", *AIDS Research and Human Retroviruses*, vol. 15, No. 13 (1999), pp. 1157-1168.
6. S. Piyasirisilp and others, "A recent outbreak of HIV-1 infection in southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC recombinant", *Journal of Virology*, vol. 74, No. 23 (2000), pp. 11286-11295.
7. Monitoring the AIDS Pandemic, *The Status and Trends of HIV/AIDS/STI Epidemics in Asia and the Pacific*, report presented at the Sixth International Congress on AIDS in Asia and the Pacific, Melbourne, 4 October 2001.
8. C. Chelala and C. Beyrer, "Drug use and HIV/AIDS in Burma", *Lancet*, vol. 25, No. 354 (1999), p. 1119.
9. Y. Shao and others, "Genetic recombination of HIV-1 strains identified in China", XIIth International AIDS Conference, Geneva, 1998.
10. N. Crofts, G. Reid and P. Deany, "Injecting drug use and HIV infection in Asia: the Asian Harm Reduction Network", *AIDS*, vol. 12, Suppl. B (1998), pp. S69-78.
11. J. S. Gillis and M. H. Mubbashar, "Risk factors for drug abuse in Pakistan: a replication", *Psychological Reports*, vol. 76, No. 1 (1995), pp. 99-108.
12. K. L. Dehne and others, "The HIV/AIDS epidemic in eastern Europe: recent patterns and trends and their implications for policy-making", *AIDS*, vol. 13, No. 7 (1999), pp. 741-749.

13. A. Bobkov and others, "A sudden epidemic of HIV type 1 among injecting drug users in the former Soviet Union: identification of subtype A, subtype B, and novel gagA/envB recombinants", *AIDS Research and Human Retroviruses*, vol. 14, No. 8 (1998), pp. 669-676.
14. T. Rhodes and others, "HIV infection associated with drug injecting in the Newly Independent States, eastern Europe: the social and economic context of epidemics, *Addiction*, vol. 94, No. 9 (1999), pp. 1323-1336.
15. T. Khanina and others, "A homogeneous subtype A HIV-1 strain caused new local outbreaks among injecting drug users in Russia", XIIIth International AIDS Conference, Durban, 2000.
16. C. Beyrer and others, "Patterns of HIV spread associated with drug trafficking", Proceedings of the Global Research Network, Third Annual Meeting, Durban, August 2000.
17. K. Motomura and others, "Emergence of new forms of human immunodeficiency virus type 1 intersubtype recombinants in central Myanmar", *AIDS Research and Human Retroviruses*, vol. 16, No. 17 (2000), pp. 1831-1843.
18. Z. Wu and others, "Risk factors for intravenous drug use and sharing equipment among young male drug users in Longchuan County, south-west China", *AIDS*, vol. 10, No. 9 (1996), pp. 1017-1024.
19. S. Panda and others, "Interface between drug use and sex work in Manipur", *National Medical Journal of India*, vol. 14, No. 4 (2001), pp. 209-211.
20. Joint United Nations Programme on HIV/AIDS, *Report on the Global AIDS Epidemic, June 2000* (Geneva, 2000).
21. B. Weniger and others, "The epidemiology of HIV infection and AIDS in Thailand", *AIDS*, vol. 5, Suppl. 2 (1991), pp. 571-585.
22. V. Lukashov and others, "Molecular characterization of a growing HIV-1 epidemic among IDUs in the former Soviet Union", XIIIth International AIDS Conference, Durban, 2000.
23. Thailand, Ministry of Health, *National HIV/AIDS Sentinel Surveillance Reports, 2000* (Bangkok, 2001).
24. B. Thaung, K. M. Gyee and B. Kywe, "Rapid assessment study of drug abuse in Myanmar", Ninth International Conference on AIDS, Vancouver, 1996.
25. Myanmar, Ministry of Health, *National HIV/AIDS Sentinel Surveillance Reports, 1998, 1999* (Yangon, 2000).
26. Joint United Nations Programme on HIV/AIDS, *Report on the Global HIV/AIDS Epidemic, 2000* (Geneva, 2000).
27. D. Burrows, "Estimating the costs of implementing widespread needle and syringe exchange provision: examples from the Russian Federation and Ukraine", paper presented at the Twelfth International Conference on the Reduction of Drug Related Harm, New Delhi, 1-5 April 2001.
28. Russian Federation, Ministry of Health, *National HIV/AIDS Sentinel Surveillance Reports, 2001* (Moscow, 2002).
29. J. Elliot, *An Unexpected Light: Travels in Afghanistan* (New York, Picador, 1999).
30. C. Beyrer, *War in the Blood: Sex, Politics and AIDS in Southeast Asia* (London, Zed Books, 1998).

31. Khomdon Singh Lisam, personal communication with the author, Imphal, Manipur State, 1998.
32. C. Zhang and others, "High prevalence of HIV-1 and hepatitis C virus coinfection among injection drug users in the southeastern region of Yunnan, China", *Journal of Acquired Immune Deficiency Syndromes*, vol. 29, No. 2 (2002), pp. 191-196.
33. C. Beyrer, "Shan women and girls and the sex industry in South-east Asia: political causes and human rights implications", *Social Science and Medicine*, vol. 53 (2001), pp. 543-550.
34. Unpublished data, courtesy of Dr. Chen Jie, Guangxi Provincial Centre for Infectious Disease Prevention and Control, China.
35. V. Poshyachinda, "Drug injecting and HIV infection among the population of drug abusers in Asia", *Bulletin on Narcotics*, United Nations publication, vol. 45, No. 1 (1993), pp. 77-90.
36. J. Gray, "Operating needle exchange programmes in the hills of Thailand", *AIDS Care*, vol. 7, No. 4 (1995), pp. 489-499.
37. D. C. Des Jarlais and S. R. Friedman, "Fifteen years of research on preventing HIV infection among injecting drug users: what we have learned, what we have not learned, what we have done, what we have not done", *Public Health Reports*, vol. 113, Suppl. 1 (1998), pp. 182-188.
38. "Bibliography on syringe-exchange references", *Substance Use and Misuse*, vol. 33, No. 5 (1998), pp. 1231-1248.
39. S. F. Hurley, D. J. Jolley and J. M. Kaldor, "Effectiveness of needle-exchange programmes for prevention of HIV infection", *Lancet*, vol. 349, No. 9068 (1997), pp. 1797-1800.
40. D. C. Des Jarlais and others, "HIV incidence among injecting drug users in New York City syringe-exchange programmes", *Lancet*, vol. 348, No. 9033 (1996), pp. 987-991.
41. G. Seage and others, "Are US populations appropriate for trials of HIV vaccine?", *American Journal of Epidemiology*, vol. 153, No. 7 (2001), pp. 619-627.
42. Thai Working Group on HIV/AIDS Projections, *Projections of HIV/AIDS in Thailand, 2000-2020* (Thailand, Ministry of Public Health, AIDS Division, Bangkok, 2001).
43. V. M. Quan, A. Chung and A. S. Abdul-Quader, "The feasibility of a syringe-needle-exchange program in Vietnam", *Substance Use and Misuse*, vol. 33, No. 5 (1998), pp. 1055-1067.
44. P. Lurie and E. Drucker, "An opportunity lost: HIV infections associated with lack of a national needle-exchange programme in the U.S.A.", *Lancet*, vol. 349, No. 9052 (1997), pp. 604-608.
45. J. L. Sorensen and A. L. Copeland, "Drug abuse treatment as an HIV prevention strategy: a review", *Drug and Alcohol Dependence*, vol. 59, No. 1 (2000), pp. 17-31.

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

The need for a global understanding of epidemiological data to inform human immunodeficiency virus (HIV) prevention among injecting drug users

K. L. DEHNE
Joint United Nations Programme on HIV/AIDS, Vienna

M. ADELEKAN
United Nations Office on Drugs and Crime, Vienna

A. CHATTERJEE
Joint United Nations Programme on HIV/AIDS, Bangkok

G. WEILER
World Health Organization, Geneva, Switzerland

ABSTRACT

Since the 1980s, the injecting of illicit drugs, especially opiates, cocaine and amphetamines, has become a worldwide epidemic, affecting more than 130 countries. Because many injecting drug users (IDUs) share injecting equipment with other IDUs, they are at very high risk of contracting human immunodeficiency virus (HIV) and other blood-borne infections. Using case examples from Asia, Eastern Europe and Africa, we show that only a fraction of the data required to monitor HIV prevention among IDUs is effectively collected.

In Asia, considerable work has been done to estimate the size of drug-injecting populations and measure risk behaviours. A few HIV prevalence surveys have also been carried out. In Eastern Europe and the countries of the former Union of Soviet Socialist Republics, the focus of attention has been on the monitoring of service reach and the establishing of infection rates among those seeking services, while reliable estimates of the number of IDUs at risk and, therefore, the proportion of IDUs reached by services, are lacking. In Africa, the main purpose of specific data collection has been to establish the existence of a significant drug-injecting problem.

For a comprehensive understanding of the HIV epidemics among IDUs and of the efforts to prevent them, however, all three types of data, on the size and pattern of drug injecting, on service provision and on programme impact, including on risk behaviours and HIV prevalence and incidence, need to be systematically collected. In particular, the monitoring of the coverage of treatment and prevention services for

IDUs has so far been neglected. In order to improve data collection for HIV prevention among IDUs, therefore, an effort is required that goes beyond individual geographical locations, sectors and disciplines.

Keywords: injecting drug use; HIV prevention; population size estimation; behavioural surveillance; response monitoring; coverage; HIV prevalence.

Introduction

Since the 1980s, the injecting of illicit drugs, especially opiates, cocaine and amphetamines, has become a worldwide epidemic, affecting perhaps 10 million people in more than 130 countries, most of them young people between the ages of 15 and 30 [1]. For most cities, countries and regions worldwide, no reliable estimates of the number of injecting drug users (IDUs) exist, however, and those figures relating to the extent and pattern of drug injecting and their trends that are available are often the "guesstimates" of ministry of interior or ministry of health officials rather than the results of research or systematic surveillance. The present article aims to provide examples of recent initiatives to improve the accuracy of IDU estimates in several world regions.

Because many drug injectors share injecting equipment with other IDUs, they are at very high risk of contracting HIV and other blood-borne infections. For example, very rapid increases in HIV prevalence among IDUs, from less than 5 per cent to over 30 or 50 per cent within one to three years, have been reported in cities in Belarus, India, Myanmar, the Russian Federation, Thailand, Ukraine, the United Kingdom of Great Britain and Northern Ireland and the United States of America, among others [2]. Significant epidemics among drug users leading to several thousand infections among IDUs have also been reported from countries and regions such as China, Malaysia, Southern Europe and parts of Latin America, where IDUs account for a large proportion of all HIV cases [2], although the real extent of those epidemics is rarely clear. Both the extent to which drug-using populations engage in risk behaviours, including sharing non-sterile injecting equipment and practising unsafe sex, and the levels of IDUs infected with HIV are of obvious relevance to HIV prevention efforts.

Furthermore, there is evidence from several sites that HIV epidemics among drug users can be prevented or stabilized through the implementation of appropriate public health policies and early interventions and therefore data on programmes and service provision are also needed [3]. These public health interventions and the policies that allow for their implementation mainly relate to two types of service: drug dependency treatment and HIV prevention services, including outreach and the provision of information and safe injecting equipment for those not in treatment. A recent review of the impact of drug dependency treatment on HIV risk behaviours and infection rates showed that all standard opiate dependency treatment modalities have the potential to reduce street heroin injecting, the sharing of non-sterile injecting equipment and consequently the

transmission of HIV [4]. Retention in treatment and treatment progress (quality of the treatment), rather than simply time spent in treatment, were found to be predictors of risk reduction, and substitution treatment such as methadone maintenance was more effective than detoxification and psychosocial counselling alone [4]. There is also significant evidence that service packages comprised of (peer) outreach, risk-reduction counselling and needle exchange programmes can be effective in modifying risk behaviours and preventing HIV infections [5].

Three types of data therefore appear to be required in order to be able to monitor progress in HIV prevention programming comprehensively: *(a)* data on the size and geographical location of drug-injecting populations, including on the types and patterns of drugs injected (regular drug users' injecting frequency differs according to whether opiates, cocaine or amphetamines are being injected); *(b)* data on service provision (for example, treatment demand and needle exchange programme enrolment); and *(c)* data on the impact or outcome of prevention (and care) programmes, including on risk behaviours (for example, rates of equipment-sharing), HIV prevalence and incidence.

As the following case examples from Asia, Eastern Europe and Africa show, only a fraction of those data tend to be routinely and consistently collected in the more than 130 countries known to have a drug-injecting problem and, even where most of these data are collected, they are not necessarily compiled and analysed across services, sectors and cities or utilized for joint planning and programming.

Efforts to estimate the size of drug-injecting populations and to measure HIV risk behaviours among injecting drug users in Asia

In Asia, considerable work has been done on estimating the size of drug-injecting populations and measuring risk behaviours. Some service statistics, such as drug dependency treatment, are also available. Various national- or state-level agencies have been involved in various types of data collection. For example, public health and social welfare agencies have used traditional surveys and indirect methods of estimation. In the absence of good secondary data, street-based surveys of drug users have been carried out and extrapolation techniques have been used. Law enforcement-focused drug control agencies in Asia have used police surveillance in communities as data sources and as a starting point for the estimation of IDU population size. However, the statistics available do not always distinguish between injectors and non-injectors. Establishing the successful coordination of data collection efforts and a common understanding of the methodology and purpose across sectors remains a challenge in most countries of Asia.

Table 1 lists some of the methods used to estimate the size of the IDU populations in Asian countries, the appropriateness of which was reviewed by various agencies and experts at a workshop held recently in Indonesia. The Global Assessment Programme on Drug Abuse of the United Nations International Drug Control Programme is currently developing a methodological tool kit to support

the application of indirect estimation techniques [6]. The tool kit provides an overview of the different estimation procedures and provides practical guidelines for their implementation, including for their application in developing countries.

Table 1. Methods of estimating the size of the drug-using and other vulnerable populations in Asia

Method	*Country*
National household survey	India
Multiplication and capture-recapture from treatment centre and other data	Iran (Islamic Republic of), Pakistan, Thailand
National drug-testing register (for example, compulsory drug screening for driving-, marriage- or trade-licence applicants)	Iran (Islamic Republic of)
Street-based enumeration and extrapolation at the city and state levels	Bangladesh, India
Police surveillance in communities	China, Myanmar

Sources: United Nations International Drug Control Programme, Regional Office for South Asia, New Delhi; national drug control agencies in China and Myanmar; T. D. Mastro and others, "Estimating the number of HIV-infected injection drug users in Bangkok: a capture-recapture method", *American Journal of Public Health*, vol. 84, 1994, pp. 1094-1099; E. Razzaghi and others, *A Rapid Assessment of Drug Abuse in Iran* (United Nations International Drug Control Programme, Tehran, 1999); Co-operative for American Remittances to Europe (CARE), Bangladesh, *Drug Injecting and Potential for Continuing Spread of AIDS: A Baseline Assessment in Dhaka City* (May 1998); S. Sarkar and others, "Rapid spread of HIV among injecting drug users in north-eastern states of India", *Bulletin on Narcotics*, vol. XLV, No. 1 (1993), pp. 91-105.

City estimates are usually more readily available than state-, province- or national-level estimates, and national estimates developed from incomplete lower-level data are unreliable. Nevertheless, the results of all those studies combine with less systematic data collection efforts by ministries of health and drug control agencies to suggest that the number of drug injectors living in South and East Asian countries may be large, ranging from around 20,000 in Nepal and 25,000 in Bangladesh to 70,000-200,000 in the Islamic Republic of Iran, 200,000 in Malaysia, 150,000-300,000 in Myanmar and 600,000 to 1 million in China [6].

Studies of HIV risk behaviours among drug-injecting populations have also multiplied in Asia in recent years. Several countries have now included IDUs as a risk group in their national HIV behavioural surveillance (for example, Bangladesh [7], India [8], Indonesia and Viet Nam). In many of those national behavioural surveillance studies, sampling has considerably varied over time, however, from purposive to more random and representative. Risk behaviours have generally been described as frequent, with needle or syringe sharing rates among IDUs ranging from 30 to over 90 per cent, depending on the indicator and reference period used. However, variable data quality, the absence of trend data and the lack of

Table 2. Selected risk indicators among 515 male drug injectors in urban Dhaka

Indicator definition	*Time reference*	*Indicator*
Mean age at first injection		20
Mean length of years injecting		12.3
Most common drug injected		Tidijesic
Percentage using previously used needle or syringe	Last week	81.4
Percentage using previously used needle or syringe	Last month	87.4
Percentage passing on needle and syringe	Last week	76.1
Percentage passing on needle and syringe	Last month	79.0
Percentage having professional injection	Last week	3.5
Percentage having professional injection	Last month	5.2
Percentage sharing needles or syringes active or passive	Last week	93.4
Percentage sharing needles or syringes active or passive	Last month	94.2
Mean number of injections	Last week	14.6
Mean number of injections	Last month	59.5
Mean number of people sharing	Last sharing event	3
Percentage buying sex from any commercial partner	Last month	32.6
Percentage buying sex from any commercial partner	Last year	57.9
Percentage using condom at last paid sex	Last time	21.1
Percentage selling sex to any commercial partner	Last year	5.2
Percentage having participated in needle or syringe exchange	Ever	30.9

Source: *Family Health International*, Bangkok, 2002.

comparability owing to definitional problems over what constitutes "sharing" have been common problems.

From some sites, such as Dhaka, the results of several rounds of cross-sectional behavioural surveys among drug injectors are available, as well as data on a variety of risk indicators (see table 2).

As mentioned above, HIV prevalence surveys have also been carried out among IDUs in Asian countries. For example, in Mandalay, Myanmar, in 2000, 63 per cent of IDUs were found to be already infected and in Thailand recent surveys have shown infection rates of approximately 50 per cent. In both China and India, HIV prevalence rates have varied from less than 5 per cent to 80 per cent in different sites [9].

By contrast, there is little systematic collection of data on HIV prevention service delivery for IDUs in the Asian region, partially because there are so few services. In most Asian cities, only a small minority of IDUs has access to either needle and/or syringe exchange or quality drug-dependency treatment. Exceptions are Bangkok [10] and Hong Kong Special Administrative Region of China [11], where a significant proportion of IDUs are enrolled in methadone maintenance programmes.

Table 3. HIV prevalence among injecting drug users in Eastern Europe, selected studies

Place	*Year*	*Characteristics*	*Sample size*	*HIV prevalence (percentage)*
Ukraine				
Poltava	2000	Needle and/or syringe exchange	259	42.1
Odessa	2001	Needle and/or syringe exchange	250	68.0
Kharkiv	2000	Needle and/or syringe exchange	250	17.6
Kryvyi Rig	1999	Needle and/or syringe exchange	249	28.1
Russian Federation				
St. Petersburg	2001	Needle and/or syringe exchange	252	35.7
St. Petersburg	2000	Survey of prison inmates	9 727	46.0
Novosibirsk	2000	IDUs seeking care	239	5.9
Rostov-na-Donu	2001	Needle and/or syringe exchange	255	33.3
Belarus				
Svetlogorsk	2000	IDUs seeking care	250	74.0
Minsk	2000	IDUs seeking care	224	22.3
Mogilev	2000	IDUs seeking care	224	1.8
Vitebsk	2000	IDUs seeking care	154	0.0
Kazakhstan	1999/2000	Registered IDUs	21 013	0.5
Temirtau	2000	Needle and/or syringe exchange	415	26.0

Sources: Ministries of Health of Belarus and Ukraine; K. L. Dehne, "HIV among IDUs and the extent of heterosexual spread in Eastern Europe", *Global Research Network Meeting on HIV Prevention in Drug-Using Populations, Fourth Annual General Meeting*, Melbourne, Australia, October 2001 (National Institute on Drug Abuse, Baltimore, Maryland, United States of America, 2001).

Programme indicator development in Eastern Europe

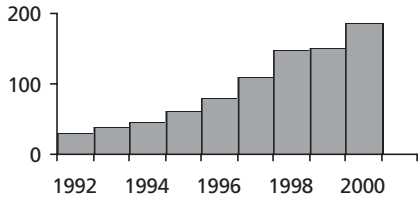
In Eastern Europe and in the countries of the former Union of Soviet Socialist Republics, the focus of attention has not been on the estimation of the number of drug injectors, but on the monitoring of service reach and the establishing of infection rates among those attending services. In some countries, routine HIV testing of illicit drug users seeking treatment and upon arrest also continues.

More than 150 outreach or needle exchange projects exist in the subregion, virtually all established since the first outbreaks of HIV occurred among IDUs in that region in the mid-1990s [12]. In some of those sites, HIV prevalence surveys have been conducted among service attenders or from the residual blood of discarded syringes, or both. The results of recent survey rounds are shown in table 3.

In recent years, the measurement of service reach has drawn increasing attention. The various outreach or needle and/or syringe exchange programmes serve from under one hundred to several thousand IDUs and some projects routinely monitor not only the absolute number of client visits, for example, during the previous month, but also retention (repeat visits by individual clients) and the number of needles and syringes exchanged [13]. A number of behavioural studies among opiate injectors are also under way.

Drug dependency treatment episodes of opiate users, mostly injectors, are also routinely monitored, and all drug users in treatment are kept on a register. In the Russian Federation, for example, the number of illicit drug users treated, mostly opiate-dependent injectors, increased from 45,000 in 1992 to 275,000 in 2000 (see figure I).

Figure I. Number of registered cases of "narcomania" in the Russian Federation per 100,000 population



Source: Russian Federation, Ministry of Health.

Note: The term "narcomania" is a Russian word meaning addiction. Those affected by narcomania are considered to comprise various types of drug abuser or illicit drug user, or both, including injectors, or all types of clients admitted for drug treatment. The proportion of injectors among those admitted is estimated at 60 to 70 per cent.

However, only a small fraction of all persons with a recent history of drug injecting are believed to be in treatment at any given time in Eastern European countries, implying an urgent need to arrive at more reliable estimates of the real number of IDUs.

The lack of reliable IDU estimates is currently the main concern of programme developers in the region. In 1999, a review of all the available data and estimates suggested that with up to 600,000 IDUs in Ukraine and between 1.5 and 3 million IDUs in the Russian Federation, amounting to more than 1 per cent of the population in each country, that region might be experiencing the highest rate of drug injecting among all world regions [14]. The very large number of registered HIV infections among IDUs, with more than 100,000 new cases in the Russian Federation alone in 2001, also suggests a major drug-injecting problem (V. Prokrosvkyi, Russian Federal AIDS Centre, personal communication).

However, many individual cities and administrative regions rely only on very rough estimates of the real number of IDUs derived from key informants' interviews and the rather erratic use of treatment demand multipliers. Most projects therefore have no clear idea of the number of IDUs in their catchment area, or the proportion of them reached by prevention services. The Ministry of Health of Ukraine, for instance, has estimated that between 5 and 15 per cent of all IDUs countrywide are reached by prevention services [15], but in reality estimates may vary even more widely, depending on which sites are used for extrapolation.

Several community-based studies are now under way, and the United Nations International Drug Control Programme is currently auditing all the available data and assessing data quality in the region.

Situation assessments in Africa

In Africa, where there are serious constraints on the collection of data specifically on drug use, owing in part to the shortage of qualified research personnel, basic equipment and funding. The emphasis has been on the local IDU situation and baseline assessments, rather than on the collection of countrywide or region-wide data on IDUs, HIV or prevention programmes. The priority has been to establish the existence of a significant drug-injecting problem in countries where there is already preliminary or anecdotal evidence of it.

Several studies have now confirmed the existence of illicit drug injecting in countries such as Kenya [16], Nigeria [17] and South Africa [18, 19] and significant rates of HIV risk behaviours among IDUs. Furthermore, questionnaires returned by Member States to the United Nations International Drug Control Programme indicate that drug abusers in several other countries in Africa do inject illicit drugs. Such data lack essential details, however, as most are not derived from systematic studies.

A study of 86 current injectors in hospitals in Johannesburg, South Africa, showed significant rates of medical sequelae of non-sterile injecting of synthetic opiates, including thrombo-phlebitis, right-side endocarditis and HIV [18]. Another report, from Gauteng and Cape Town, also in South Africa, showed an increasing proportion of heroin users among hospital patients, with 36 and 51 per cent respectively of them admitting at least incidental injecting modes of consumption. Unfortunately, the report did not provide the absolute numbers of patients involved, nor information on their HIV status [19].

The Kenya (Nairobi and Mombasa) and Nigeria (Lagos) studies not only confirm the existence of significant drug injecting outside hospital patient populations, but also the difficulties in directly attributing HIV transmission to drug use in settings where HIV infections are frequent due to sexual transmission [16, 17]. In both studies, the researchers experienced no difficulties in recruiting a significant number of drug users from the streets and interviewing them about the extent of their injecting, the associated risk factors and the services available to them. In Lagos, 398 street users of heroin and cocaine were recruited through snowball sampling techniques from four local government areas, of whom 82 reported having ever injected and 54 currently injecting. The injecting risk factors identified included sharing and reusing injecting equipment and injecting in a dirty environment. The preliminary results of the Nairobi and Mombasa study seem to be similar to those of the Lagos study.

The HIV risk attributable to drug injecting was not clear in the Lagos study, and detailed results for the Kenya study are not yet available. Although the 10 per cent HIV prevalence rate of the sample of drug users in Lagos was higher than the

5 per cent prevalence rate of the general population, there was no significant difference between the HIV rates obtained for non-injectors (10 per cent) and ever-injectors (8.9 per cent). Furthermore, female drug abusers, irrespective of their injecting status, were significantly more likely to be HIV positive than their male counterparts, a factor possibly attributable to the fact that most of the females also engaged in commercial sex work.

Thus the few available studies all suggest that, although not as rare as hitherto perceived, drug injecting is probably much less frequent in sub-Saharan Africa than in other regions and that the drug-injecting epidemics are at an early stage. However, there is a risk that the extent of the problem may continue to be underestimated, even if in reality drug injecting and associated HIV risks increase. The reasons for this are the low level of awareness among drug policy makers, law enforcement agencies and the general public of the possible link between IDU and HIV/AIDS and the associated consequences in this region [16], the tendency among AIDS programme managers not specifically to consider drug users when determining HIV surveillance sites and protocols and the logistical constraints related to research and programme development mentioned above.

In North Africa and the Middle East, IDU/HIV data collection efforts have also mostly concentrated on situation assessments and individual studies. However, the results differ in that drug injecting may well be more frequent there than in sub-Saharan Africa, while sexual HIV transmission is much less widespread, leading to injecting drug use representing a major or even the most significant mode of HIV transmission in some of the countries. For example, recent reports indicate that IDUs constituted 91.7 per cent of the 4,439 HIV/AIDS cases registered among Libyan nationals up to the end of 2001 [20]. The percentages of IDUs among reported AIDS cases in Algeria, Bahrain and Tunisia were 18.4, 73 and 34 respectively [21].

In a recent report commissioned by the World Bank, all the available data reported on drug injecting and HIV in the North Africa and Middle East region were compiled [22].

Summary and conclusions

In the three regions of Asia, Eastern Europe and Africa, the collection of data on HIV/AIDS and drug use concentrated initially almost exclusively on the description of samples of IDUs and the prevalence and incidence rates of HIV/AIDS among them. Indeed, to date, reliable information on the real size of the IDU population is still frequently lacking, despite efforts by the United Nations International Drug Control Programme and others to collect such data. That finding is not particularly surprising, given that IDUs usually constitute a semi-hidden population that tends to avoid contact with law enforcement authorities and often with health authorities as well and thereby escapes direct estimation [23]. Very few good national-level estimates of the true numbers of IDUs in specific sites exist, although sometimes seemingly exaggerated figures provided by

Governments gain in plausibility when additional data sources become available, such as the large number of HIV infections reportedly due to IDU in the Russian Federation. Efforts to establish the level of HIV infection in the IDU population usually rely on ad hoc testing among IDU populations such as those attending services, receiving treatment in hospital, sampled through outreach, arrested for illicit use or living in the street, while systematic sentinel surveillance approaches still represent the exception rather than the rule [24].

The situation with regard to data collection for HIV prevention among IDUs in Latin America has not been described in the present article but is unlikely to differ significantly from that in other developing regions. Nor are the situations in Australia, North America and Western European countries described here, although surveillance activities in those regions are more systematic [25]. By contrast, most African countries appear to lack even the basic data needed to monitor HIV infection related to injecting drug use. Although HIV infection on the African continent remains largely of heterosexual origin, the role of injecting drug use may not have been sufficiently researched and has perhaps been underestimated.

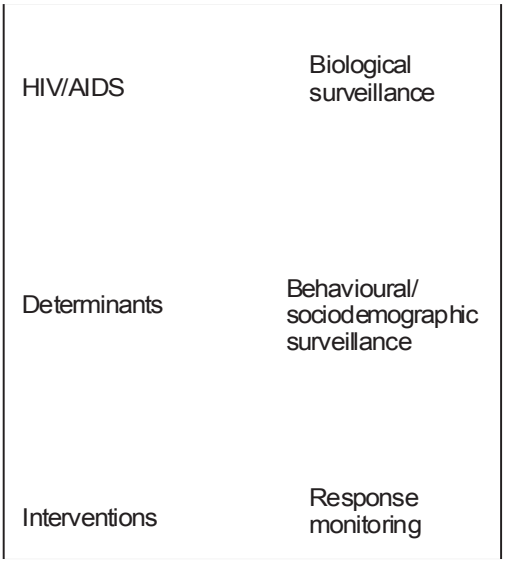
It is only recently that efforts have been intensified to supplement biological surveillance with behavioural surveillance. The Asian experience, in particular, shows that the concept of "second generation surveillance" [26] adds significantly to the understanding of the twin epidemics of drug use and HIV. As the example from Bangladesh illustrates, behavioural data can point to the presence of factors explaining changes in the epidemic or warn of the epidemic's future potential (table 2). Unfortunately, in most countries those efforts have been limited to a few sites.

While the interplay between biological and behavioural surveillance may generate useful insights into the course and determinants of the HIV epidemic, its suggestive power with regard to the nature of appropriate public health responses is, however, limited. The important question of how the response to drug use and HIV should be modified is unlikely to be answered in the absence of a description of the activities that are in place. Similarly, learning from success by examining how change has been achieved depends on an understanding of the responses employed. Therefore, monitoring the public health response to HIV/AIDS, in particular service delivery, represents a critical third component of drug monitoring and biological and behavioural surveillance systems, and it is the integration of all three components that can turn such systems into more powerful and relevant public health planning tools (see figure II).

To date, in all the regions reviewed, the collection of data on the availability of treatment and prevention services for IDUs has been, in the context of limited reach, low on the priority list. However, as the number of IDUs enrolled in needle-syringe exchange programmes in Eastern Europe and Asia increases, so does the demand for more precise information on the actual and desirable reach of existing programmes.

Ultimately, working towards a public health response to HIV among IDUs that integrates existing scattered projects into cohesive programmes on an appropriate scale is the objective of establishing comprehensive surveillance and monitoring

Figure II. Monitoring HIV prevention among injecting drug users



systems. Although the effectiveness of certain HIV prevention interventions is well established, such systems need to take a perspective that extends beyond a single intervention and is based on a full understanding of the appropriate mix and coverage of services at the city, province and country levels.

It is recognized that no single intervention can halt HIV transmission among IDUs and their non-injecting partners; the prevention of IDU-related HIV epidemics requires the provision of a range of services, including community-based information, education and communication; voluntary counselling and testing; the provision of sterile injecting equipment and condoms; and treatment and care for IDUs [27].

Each of those service areas can be described in a number of dimensions [28]: "provision of services" aims to capture the nature of services delivered and includes features such as the affiliation of services and the cost of services at the point of delivery; "service utilization" aims to quantify to what extent the services offered are being used by IDUs and highlight their accessibility and acceptability; and "service coverage" relates the observed utilization to the potential demand, in terms of the entire IDU population, specific geographical areas (such as rural versus urban areas) or IDUs living in specific settings (such as incarcerated IDUs).

The barriers that tend to prevent an effective response are poorly understood and may include outdated health care and HIV prevention concepts and structures, as well as a lack of understanding of the contextual and political environments in which policies and programmes are being developed [29].

In particular, many decision makers appear to fail to take into account the crucial concept of "service coverage". For example, Burrows reported that ministry of health officials in one Eastern European country had failed to take into account the fact that the coverage of types of programme that had proved effective elsewhere had been minimal and had repeatedly described them as "ineffective" or "of doubtful effectiveness" because the number of HIV infections had continued to rise in the cities employing those approaches [30]. Conveying the message to public health planners and decision makers that even the most effective intervention can only prevent infections among those reached is therefore one critical objective of integrated monitoring and surveillance systems.

The above example illustrates that the objective is not only to fill in the gaps in drug epidemiology and biological, behavioural and response data, but also to make strategic use of such data for HIV programming. To achieve that objective, a broad understanding of epidemiological data beyond individual geographical locations and across various public health disciplines will be required. Such a global understanding of epidemiological data could, and should, play a key role in driving and shaping the political agenda of all stakeholders, such that their goals are appropriate and realistic and their pledges for funding meet the identified needs.

References

1. Joint United Nations Programme on HIV/AIDS (UNAIDS*), Report on the Global HIV/AIDS Epidemic, June 2000* (Geneva).
2. A. Ball, "HIV prevention among injection drug users", *Global Research Network Meeting on HIV Prevention in Drug-Using Populations, Inaugural Meeting*, Geneva, 25-26 June 1998 (Baltimore, Maryland, United States of America, National Institute on Drug Abuse), pp. 13-18.
3. D. Des Jarlais and others, "The structure of stable sero-prevalence HIV-1 epidemics among injecting drug users", in *Drug Injecting and HIV*, G. Stimson, D. Des Jarlais and A. Ball, eds. (London, UCL Press, 1998), pp. 91-100.
4. M. Farrell and others, "Effectiveness of drug dependency treatment in prevention of HIV among IDUs*", Thirteenth International Conference on the Reduction of Drug-Related Harm*, Ljubljana, 3-7 March 2002.
5. S. F. Hurley, D. J. Jolley and J. M. Kaldor, "Effectiveness of needle-exchange programmes for prevention of HIV infection", *Lancet*, vol. 349, No. 9068 (1997), pp. 1797-1800.
6. A. Ball, "HIV prevention among drug users in Asia*", World Health Organization Western Pacific Regional Office HIV/AIDS/TB Programme Management Meeting*, Melbourne, Australia, 2-4 October 2001 (Manila, World Health Organization (WHO)).
7. People's Republic of Bangladesh, Ministry of Health and Family Welfare, Directorate General of Health Services, AIDS and STD Control Programme, "Report on the second national expanded HIV surveillance, 1999-2000" (September 2000).
8. National AIDS Control of India (NACO). Web site at www.naco.nic.in/vsnaco/indianscene/executive.htm (accessed 19 September 2002).

9. G. Reid and G. Costigan, *Revisiting "The Hidden Epidemic": A Situation Assessment of Drug Use in Asia in the Context of HIV/AIDS* (Burnet Institute and the Centre for Harm Reduction, Australia, 2002).

10. UNAIDS South East Asia Intercountry Team, "Best practice case study of the Bangkok Metropolitan Authority Methadone Treatment Programme", forthcoming.

11. Hong Kong Special Administrative Region of China, Security Bureau, Narcotics Division, "Report on review of Methadone Treatment Programme" (Hong Kong Special Administrative Region of China, 2000).

12. UNAIDS Eastern Europe IDU Task Force secretariat, "Inventory of Needle Exchange/Outreach Projects" (Vienna, 2002).

13. K. L. Dehne, "HIV among IDUs and the extent of heterosexual spread in Eastern Europe", *Global Research Network Meeting on HIV Prevention in Drug-Using Populations, Fourth Annual General Meeting*, Melbourne, October 2001 (National Institute on Drug Abuse, Baltimore, Maryland, United States of America, 2001).

14. K. L. Dehne and Y. Kobyshcha*, The HIV Epidemic in Central and Eastern Europe: Update 2000* (UNAIDS, Geneva, 2000).

15. Ukraine, Ministry of Health, *HIV/AIDS Surveillance in Ukraine, 1987-2000* (Kiev, Ukrainian AIDS Centre and UNAIDS, 2000).

16. M. Odek-Ogunde, "Preliminary report on the World Health Organization Drug Injecting Study Phase II in Nairobi, Kenya", paper presented at the *United Nations International Drug Control Programme and Joint United Nations Programme on HIV/AIDS Workshop on Drug Abuse and HIV/AIDS in Africa*, Sharm El Sheikh, Egypt, 5-9 November 2001.

17. M. L. Adelekan and others, "Injection drug use and associated health consequences in Lagos, Nigeria: findings from the World Health Organization Drug Injecting Study Phase II", *Third Annual Meeting Report, Global Research Network Meeting on HIV Prevention in Drug-Using Populations*, July 2000, Durban, South Africa (Department of Health and Human Services, United States of America, 2001), pp. 27-41.

18. P. G. Williams, S. M. Ansell and F. J. Milne, "Illicit intravenous drug use in Johannesburg: medical complications and prevalence of HIV infection", *South African Medical Journal*, vol. 87, No. 7 (1997), pp. 889-891.

19. South African Community Epidemiology Network on Drug Use, "Alcohol and drug abuse trends: July-December 2001", in *SACENDU Update*, C. Parry and others, eds. (Medical Research Council, South Africa, April 2002).

20. M. A. Sammud, "Libyan Arab Jamahiriya report", presented at the *Twelfth Inter-Country Meeting of National AIDS Programme Managers*, 23-26 April 2002, Beirut.

21. World Health Organization, "Progress report on acquired immunodeficiency syndrome (AIDS) in the Eastern Mediterranean region", *Forty-eighth Session of the Regional Committee for the Eastern Mediterranean*, 30 September-4 October 2001 (Alexandria, Egypt, WHO Regional Office).

22. C. Jenkins, and D. Robalino, "Overview of the HIV/AIDS situation in the Middle East and North Africa and Eastern Mediterranean region", paper presented at the *Conference on Meeting the Public Health Challenges in the 21st Century in the Middle East, North Africa and Eastern Mediterranean Region*, 17 June 2002 (World Bank, forthcoming).

23. D. Des Jarlais, K. L. Dehne and J. Casabona, "HIV surveillance among injecting drug users", *AIDS*, vol. 15, Suppl. 3, pp. S13-S22.
24. UNAIDS and WHO, Epidemiological fact sheets on HIV/AIDS and sexually transmitted infections (Geneva, 2000).
25. European Monitoring Centre for Drugs and Drug Addiction, *Annual Report on the State of the Drugs Problem in the European Union* (Lisbon, 2001). Available online at http://annualreport.emcdda.org/ (accessed 20 September 2002).
26. UNAIDS and WHO Working Group on Global HIV/AIDS and STI Surveillance, *Guidelines for Second Generation Surveillance for HIV: The Next Decade*. Available online at www.who.int/emc documents/aids_hiv/whocdscsredc20005c.html (accessed 20 September 2002).
27. United Nations, "Preventing the transmission of HIV among drug abusers: a position paper of the United Nations System", *Annex to the Report of the Eighth Session of the Administrative Committee on Coordination Subcommittee on Drug Control*, 28-29 September 2000. Available online at www.unaids.org/publications/documents/specific/injecting/Hraids.doc (accessed 20 September 2002).
28. G. Weiler, C. Archibald and J. Siushansian, "Global indicator database on HIV prevention in injecting drug users: a discussion paper", *Global Research Network Meeting on HIV Prevention in Drug-Using Populations, Fourth Annual General Meeting*, 11-12 October 2001, Melbourne.
29. T. Rhodes and others, "Drug injecting, rapid HIV spread and the 'risk environment': implications for assessments and response", AIDS, vol. 13, Suppl. A, pp. 259-269.
30. D. Burrows, personal communication (2002).

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

An ethical framework for drug epidemiology: identifying the issues

C. L. FRY*

Senior Research Fellow, Turning Point Alcohol and Drug Centre, Melbourne and Fellow, Department of Public Health, University of Melbourne, Parkville, Australia

W. HALL**

Professorial Research Fellow and Director, Office of Public Policy and Ethics, Institute for Molecular Bioscience, University of Queensland, St. Lucia, Australia

ABSTRACT

The Global Workshop on Drug Information Systems: Activities, Methods and Future Opportunities, held in Vienna from 3 to 5 December 2001, highlighted the need for an ethical framework for drug epidemiology. The present article suggests some first steps that could be taken towards developing such a framework by identifying some of the key issues for consideration. The scope of drug epidemiology is defined and attention is drawn to the current dearth of scholarship and lack of specific guidelines on the ethical issues raised by such research. The importance of ethics in drug epidemiology is explained and it is argued that a guiding framework would be helpful in promoting an understanding of some of the prominent ethical challenges in this field (for example, obtaining free and informed consent to participation, the use of inducements to recruit subjects, the protection of interviewees from violations of privacy and the risk of prosecution, and the safety of field research staff). The traditional principles of biomedical ethics are outlined and their limitations in enabling an understanding of ethical issues in drug epidemiology are considered. The utility of practical case-based approaches to ethical analysis is also discussed. The article concludes with broad recommendations for an ethical framework for drug epidemiology that can be refined in further discussion on those important issues.

Keywords: ethics; epidemiology; drug use; public health.

*Craig Fry is grateful to Libby Topp of the National Drug and Alcohol Research Centre in Sydney and to the United Nations International Drug Control Programme for the opportunity to attend the Global Workshop on Drug Information Systems: Activities, Methods and Future Opportunities, held in Vienna in December 2001, which inspired this work.

**Wayne Hall acknowledges the encouragement given to his thinking on the issues in this article by a request from the World Health Organization to provide a discussion paper for a meeting on the ethical issues raised by neuroscience research on addictions.

Introduction

The idea for the present article grew from the first author's participation in the Global Workshop on Drug Information Systems: Activities, Methods and Future Opportunities, organized by the United Nations International Drug Control Programme and held in Vienna from 3 to 5 December 2001. The purpose of the workshop was to discuss global trends in illicit drug use and to reflect on the utility of epidemiological methods in monitoring such trends. At the workshop some common practical challenges of drug epidemiology research were highlighted, such as issues of study design, subject recruitment, data collection and analysis, reporting, the secondary uses of data, many of which raised ethical challenges, or at least signalled issues that ought to be the subject of explicit ethical analysis.

At the workshop it became clear that there was currently no ethical framework to guide investigators in planning and conducting epidemiological studies of illicit drug use or to assist ethics committees and funding bodies in appraising research proposals and identifying ways to resolve those issues. The purpose of the present article, therefore, is to discuss some of the first steps that could be taken towards developing such a framework.

Public health, epidemiology and ethics

The mortality and morbidity associated with the misuse of alcohol and illicit drugs represent a significant challenge to public health [1, 2]. A key feature of the modern public health response to such drug problems has been the use of epidemiological research methods to define at-risk populations, identify opportunities for intervention and evaluate the effects of different policies.

A consensus has slowly emerged that epidemiological research raises a unique set of ethical challenges [3-5], including questions of data-sharing, confidentiality, privacy and the role of epidemiology in public health advocacy. Although a number of ethics guidelines have been proposed for epidemiologists, the thinking on the ethics of epidemiology is still in its infancy and little attention has been given to defining its core values [5].

More generally, there is a lack of ethics literacy in the public health field [6-8] and no agreed framework in public health for analysing ethical dilemmas [9]. There is a similar scarcity of critical discussion of either the ethical underpinnings of addiction research or how to deal with its day-to-day ethical challenges [10]. In addition, the increasing application of drug information systems and rapid assessment methods to the study of drug epidemiology has occurred in a theoretical vacuum [11]. A salient example of this lack of analysis has been the failure to develop an ethical framework for drug epidemiology.

The boundaries of epidemiological research on drug addiction are not sharply defined. Drug epidemiology includes surveys of patterns of licit and illicit drug use in the community that define populations at risk [12]. It also encompasses longitudinal studies of the personal and social factors that predict the course of drug

use [12, 13]. Drug epidemiology also includes studies of the prevalence and correlates of drug dependence in the general population using standardized diagnostic interviews [14]. Observational studies of treated populations are also employed to examine rates of mortality, morbidity and abstinence among drug-dependent persons [15].

Ethical challenges in drug epidemiology

In many developed countries, there are institutional research ethics committees that oversee the ethical conduct of human research. These committees typically apply broad ethical principles in advising on the ethical conduct of biomedical, clinical and social research. There are reasonable concerns about the applicability of such principles and standards to new and emerging specialized fields of research, such as drug epidemiology. General ethical principles often fail to provide specific guidance in dealing with the complexities and ambiguities of ethical challenges that arise in everyday practice [16]. There are also questions about how such standards and guidelines, which have been developed in a specific cultural context, may be applied in developing countries that may have either very different or no research traditions and may not have an established institutional ethics committee system.

Epidemiological research on drug addiction exemplifies many of these concerns. The present article includes an examination of some of the major challenges in drug epidemiology that also have important ethical aspects. Many of these ethical challenges remain unresolved, leaving open the possibility of serious ethical breaches.

Free and informed consent

The adequacy of informed consent is commonly assessed in terms of the level of information provided to participants about research procedures, risks, benefits and safeguards; the types of information delivery when literacy levels and preferred communication modes are considered; the opportunities for participants to voice concerns and ask questions; the extent to which consent is free from duress, undue influence or intimidation; and who provides consent in accordance with local traditions.

Free and informed consent to participate in epidemiological research does not present any special problems for adults who can understand the nature of their participation and can freely decide whether or not they wish to participate. It presents more of an ethical issue for epidemiological studies of adolescents; such studies are increasingly being conducted because adolescence is when drug use often begins [17]. The participation of adolescents in any form of research usually requires parental consent and adolescent assent [4]. Obtaining such consent can be cumbersome in school-based surveys of drug use, which are an efficient

way of doing surveys of drug use. Low response rates and the underrepresentation of minority groups have prompted researchers to use a method of "passive parental consent", in which parents are informed that a survey is to be done in a circular that invites them to object to their child's participation. It is then assumed that the absence of parental objection means that the child can be included in school surveys. This approach requires more ethical justification and discussion.

The issue of informed consent also arises in the case of research involving participants who may be intoxicated, or who may have a psychiatric condition requiring medication [18]. There has been little discussion in the addictions literature of the implications for consent, autonomy and voluntariness of recruiting intoxicated persons for drug research. Few records are kept as to the intoxicated state of research participants in drug epidemiology studies, though we note from experience that it is not unusual in illicit drug research for a small proportion of participants to be intoxicated to some degree at interview. The College on Problems of Drug Dependence based in the United States of America, has suggested that informed consent should not be obtained when prospective participants are intoxicated, in withdrawal or cognitively impaired [19]. However, it is unclear how these states may be reliably determined. Some important ethical questions to be explored are whether intoxication should be an absolute exclusion criterion, how participants who are intoxicated but lucid should be dealt with, how researchers should judge the extent of impairment and how reliable such judgements would be.

The payment of research participants may also raise issues of consent in studies of drug users. In Australia, for example, it has been common practice since 1984 for drug researchers to provide monetary payment to drug users who participate in research interviews. It has proved to be a successful way of recruiting illicit drug users for a variety of studies. It is also standard practice in drug research in the United States [19]. While the bioethics literature has explored the ethics of paying research participants [20-23], it has not yet considered the special issues that exist in relation to reimbursing drug users for research involvement. Critics of the practice are concerned that cash payments could serve as an inducement and be used by drug users to buy drugs [17]. Non-cash reimbursement has been suggested as more appropriate for this group. Advocates of cash payment claim that payment for research participation is an ethical practice, as it reflects the ethical principles of respect and dignity [20, 24], and that non-cash methods may reflect negative drug user stereotypes ("they only participate in research for the money") and a paternalistic view of the capacity and rights of users to make their own choices. That issue is controversial and remains unresolved in the drug research setting. Issues requiring further attention include the suitability of various ethical frameworks for understanding research participant motivation; the relative weight placed upon various motivators for determining research involvement; definitions of inducement; the quantification of the value of time and out-of-pocket expenses; the mechanisms for reimbursement; and analysis of the risks and harm that may arise from reimbursement practices.

Confidentiality, privacy and legal hazard

Protecting the privacy of participants and the confidentiality of the information that they provide is critical in drug epidemiological research. The use of certain drugs, such as cannabis, cocaine and heroin, is illegal, as is the use of alcohol by persons who are under the minimum legal drinking age. Drug use surveys may also ask about other illegal acts, such as driving while intoxicated, selling illegal drugs or engaging in theft, fraud or violence to finance drug use. If such data were linked to individuals and provided to the police, criminal charges could be brought. In the United States, certificates of confidentiality, which provide subjects with an assurance that this will not happen, can be obtained by researchers. The situation in most other countries is much less clear [25, 26].

Assuring the confidentiality of information is less of a problem when data are collected in a single cross-sectional interview and no information identifying subjects, such as their names, is obtained. Confidentiality may become much more of an issue in longitudinal studies, during which multiple contact details may be collected to allow individuals to be contacted again for additional interviews. Standard precautions are to store separately and securely names and identifiers and the survey data. This could become a major issue if biological samples from which deoxyribonucleic acid (DNA) samples can be derived (blood, for example) were taken. Such a DNA sample would provide a unique identifier for all individuals (except identical twins) that could, if linked with questionnaire or interview data, permit named persons to be linked to self-reported illegal acts. Special legal protection and research precautions would be necessary to protect privacy in such cases. The implications for drug epidemiology of recent changes to health privacy and data protection legislation in a number of jurisdictions will require careful monitoring [27].

Safety issues

Interviews with illicit drug users are often thought of as potentially dangerous [28]. In order to protect the confidentiality of participants when conducting face-to-face survey work with illicit drug users in the field, interviews are often conducted in settings that are out of the public view. Interviews may be held outside of normal business hours, at night, at the participant's residence, in settings where researcher safety cannot necessarily be guaranteed. In drug research, topics of enquiry are often sensitive and may cause feelings of anxiety and discomfort for participants. The issue of safety arises in connection with interview locations and timing; interview content; the level of support, backup and researcher training; the response to crises that require confidentiality to be broken; and the carrying of valuable personal and research items. Safety is a concern for the researcher, the participant and third parties. Although broad safety protocols for social research are starting to emerge [29-31], they may need to be adapted for use in drug epidemiology.

Other challenges

Another issue is the practice of integrating questions ("piggy-backing") on certain aspects of illicit drug abuse into studies designed for other purposes. This approach may be indicated where more in-depth research is neither feasible nor funded and where investigators are seeking to minimize the burden of similar studies on particular target groups. However, experience suggests that there are often trade-offs to such an approach [31]. Careful consideration should be given to issues such as the likely impact on reliability and validity of reports obtained from participants; researcher training; the potential for confusion created around informed consent and confidentiality assurances for qualitatively different studies; and the ethicality of participant payment if reimbursement is for multiple studies.

Ethical concerns also arise in connection with the use of the findings of drug epidemiological research to inform policy development and decision-making. Given the varying degrees of empirical uncertainty, the values that underpin the choice of research topics and the utilization of findings have ethical implications that should be explored [7].

Challenges for drug epidemiological research in developing countries

Recent discussion has highlighted the ethical challenges posed by conducting comparative epidemiological studies of drug use across different cultures [3, 4], particularly in developing countries that have no research tradition. This work, which is still in its infancy, needs to be given priority. In addition to the application of broad bioethical principles, a focus on the significant practical challenges that exist for drug epidemiology research in developing countries may highlight issues that ought to be the subject of ethical analysis and oversight. It cannot be assumed that the notions of informed consent, confidentiality and privacy that have arisen out of ethical debates in developed countries can be straightforwardly applied in all cultures and societies.

For example, there are aspects of informed consent that are particular to the conduct of drug epidemiology in developing countries. As it is a relatively recent development in research ethics, there are still many unanswered questions about the requirements of informed consent in these settings [32]. The relevance of certain issues, such as participant vulnerability, levels of awareness of and expectations about rights, communication difficulties, documentation issues and the rules of obtaining consent in hierarchical societies, is still contested [33].

Further, issues of race, culture and gender may impact upon researcher safety, in particular when the research context is developing countries [30]. Non-indigenous researchers may find it particularly difficult to conduct fieldwork in such settings. The issues may also have implications for research participants and third parties to the research, especially in small communities where the participation of a person in research is more difficult to disguise than it is in larger cities. Another issue is that of monitoring the conduct of drug epidemiological research

in developing countries. Consideration should be given to the special needs that exist in jurisdictions that lack the local institutional ethics committee infrastructure to support such ethical oversight.

In considering the ethics of research in developing countries, the need for drug epidemiologists to understand the social, economic and political context in which their work is conducted is clear [34]. This will require a commitment to including the views of all stakeholders and engaging in discussion about local ethics.

Ethics discussion in epidemiology and public health has emerged from philosophical ethics, bioethics and the writings of public health practitioners reflecting upon everyday ethical challenges in professional practice (for example, advocacy, coercion, scientific misconduct, privacy, conflicts of interest and the rights of vulnerable communities) [35]. The discussion below draws from each of these sources, beginning with a brief description of bioethical approaches to the analysis of ethical issues in biomedical research.

Ethical analysis

There are a bewildering array of competing ethical theories that seek to rationalize moral rules and allow decisions to be made on what conduct is right or good in problematical cases [36, 37]. These include principle-based deontological ethics; situational ethics; utilitarian ethics; consequentialist ethics; casuist case-based ethics; narrative ethics; feminist ethics; hermeneutical ethics; and virtue ethics [38]. All of these theories capture some aspects of ethical reasoning, but none commands universal assent among ethicists [37].

Over the past three decades, an influential set of moral principles has emerged in the Anglo-American analyses of the ethics of biomedical research [4, 37, 39]. "Principlism", as it has been called, appeals to the principles of autonomy, non-maleficence, beneficence and justice. Some variants of these principles have also been included in influential international statements of ethical principles for medical research, such as the World Medical Association Declaration of Helsinki and the statements of United Nations organizations [4]. These principles have their limitations, but they provide a useful beginning for ethical discussion.

The principles of biomedical ethics: autonomy, non-maleficence, beneficence and justice

Respecting autonomy means respecting and not interfering with the action of rational persons who have a capacity for autonomous action, that is, adults who are able to decide freely upon a course of action without influence, coercion or force [37]. In the context of biomedical research, the principle of respect for autonomy is usually taken to require informed consent to treatment or research participation, voluntariness in research participation and the maintenance of confidentiality and privacy of information provided to a researcher.

The principle of non-maleficence is simply the duty to do no harm [37]. Following the principle of non-maleficence requires refraining from causing harm or injury or from placing others at risk of harm or injury. In the biomedical research context, the principle of non-maleficence requires researchers to minimize the risks of research participation [4, 39].

Beauchamp and Childress have identified "positive beneficence" and "utility" as two elements of the principle of beneficence [37]. Positive beneficence requires researchers to perform actions that result in a benefit. Utility requires them to ensure that the benefits of their actions outweigh the burdens that they impose upon others. The principle of beneficence therefore requires that an action produces benefits and that its benefits outweigh its burdens. In the context of biomedical research, this means that the benefits of the research to society should outweigh its risks to participants and also that, in the case of individual participants, the benefits of participation exceed the risks.

Justice is probably the most controversial of the four moral principles. For the purpose of this discussion, "justice" refers to "distributive justice" rather than retributive (criminal) or rectificatory (compensatory) justice [37]. In bioethics, the principle of distributive justice has been central to debates about how to ensure equitable access to health care and reduce unequal health outcomes. In the case of research, the principle of distributive justice refers to the equitable distribution of the risks, as well as the benefits of research participation [4]. A fair and just research policy would aim to achieve a distribution of the benefits and burdens of research participation that is as fair and equitable as possible.

The limitations of biomedical ethics

Concerns have been raised as to whether biomedical ethics is an appropriate model for public health, given the tension between the individualistic orientation of bioethics and the societal focus of public health. Callahan and Jennings also note that, in a pluralistic society, numerous ethical perspectives coexist on matters of widespread interest and importance and that one or more might be appropriate for any particular ethical problem [7].

However, others have observed that the prescriptive use of key ethical principles may discourage consideration of alternative ethical perspectives, such as those emphasizing collective rather than individual responsibilities. Witkin has warned that ethical principles may therefore be seen as instruments of control rather than as moral guides for research [16]. An increasingly popular alternative approach is to view ethics as a discourse rather than as a system of rules and to encourage narrative accounts of actual ethical situations that, for example, drug epidemiological researchers face.

Moral positions are historical rather than timeless and subject to revision and augmentation. Witkin argues that by ignoring alternative ethical perspectives, people limit their capacity to assess the limits of their own belief systems and so engage in ethical discussions only within the boundaries of the "taken-for-granted" [16]. An ethical framework for drug epidemiology might usefully transcend

those boundaries by adopting a more proactive approach that breaks with the reactive research ethics traditions of the past [34].

In the absence of consensus on a universal theory of ethics, the ethical analyses of public policy cannot be a matter of deducing moral rulings from categorical imperatives or applying a utilitarian calculus to all the possible courses of action. Ethical analysis does not always achieve consensus, but the range of morally acceptable behaviour is often narrowed by ethical debate. A dialectical discovery process can identify common moral rules and shared justifications for morally acceptable courses of action. This has been described by Rawls as the method of "reflective equilibrium" [40]. It involves testing ethical principles (that may be derived from one or more ethical theories) against widely shared moral rules and judgements that have been called the "common morality" [37]. People aim to reduce the discrepancies between their moral principles and their understanding of the "common morality" and, by iterative adjustment, work towards an equilibrium between their principles and their shared moral rules and judgements.

A more recent variation of this form of pluralism has been called "pluralistic casuistry". Brody argues that, in contrast to the monistic ethical theories that attempt to reduce morality to a single value or set of principles, pluralistic casuistry reflects the reality of how people engage in moral reasoning [4]. Casuistry or case-based ethics is a method of practical ethical reasoning emphasizing the value of moral intuitions about particular cases over theories or principles [35]. Pluralistic casuistry recognizes that multiple moral values may coexist and are modifiable with reflection on more cases. Casuistry is also sympathetic to communitarian ethics, in which morality is also seen as contextual and where divergent ethical values of different communities are respected. A common theme in pluralistic approaches to ethical analysis is the key role of public discussion in achieving a balance between competing ethical values [41, 42].

Conclusions

Biomedical ethics is crucial to biomedical, clinical and social research efforts. It prescribes the boundaries of ethical research conduct by identifying core principles. However, those principles do not provide guidance in relation to the day-to-day ethical challenges that researchers encounter, in particular in specialized areas such as drug epidemiology. Furthermore, such guidelines quickly become dated, with advances in modern science occurring "much faster than either ethics, law or social and public policy" ([38], pp. 279-280).

Making decisions about what is ethical involves more than just following accepted prescriptions and principles [34]. The main virtue of ethical principles such as autonomy and beneficence is that they alert us to important ethical issues; they do not solve ethical problems. Such principles must be applied and tested in the analysis of specific cases by a process of open debate and discussion if they are to be interpreted at the practical or applied level. This approach to ethical analysis needs to inform discussions of ethical issues that arise in undertaking drug epidemiological research in developing countries.

A useful way to bridge the gap that exists between the principles of ethical research and the special challenges of drug epidemiology is through the development of an ethical framework specific to drug epidemiology. A casuistic social analysis of ethics also has utility if one is committed to sharing the decision-making in this developmental process and also interested in practical outcomes.

As drug epidemiological research becomes more global, the ability to consider the potential role of ethical systems that differ from those that have grown in the biomedical tradition will be critical for successful international collaboration. Taking a collaborative and open approach to ethics will also allow for particular ethical challenges to be viewed through many lenses, the result of which should be improved ethics problem-solving. Roberts and Reich have noted recently that for epidemiologists, an awareness of "alternative ethical arguments has become as important as knowing the advantages and disadvantages of different epidemiological techniques" ([9], p. 1059).

The ethical analysis of epidemiological research on drug use is an underdeveloped field, even in developed societies with a tradition of drug research and ethical protection of human participants in medical research. The authors hope to have demonstrated the need for drug researchers to address those issues in a more systematic way. The urgency of doing so is increased by recent efforts to expand epidemiological research on drug use to cultures and societies with little tradition of drug research and often less experience in the ethical oversight of human medical research. Given the role of international organizations such as the United Nations International Drug Control Programme and the World Health Organization in sponsoring and encouraging such research, such organizations may consider facilitating discussions on ethical issues from which an ethical framework may emerge. This suggestion should not discourage drug researchers in developed countries from initiating their own discussions of those issues.

The present article is intended as a first step towards developing an ethical framework for drug epidemiology. It is hoped that it will serve as a useful beginning for future debate on the ethical challenges involved in conducting drug epidemiology research and in doing so will assist in raising the profile of research ethics considerations for the addictions specialty in public health.

References

1. D. J. Collins and H. M. Lapsley, *The Social Costs of Alcohol and Drug Use* (Canberra, Australian Government Publishing Service, 1996).
2. C. J. Murray and A. D. Lopez, "Evidence-based health policy: lessons from the Global Burden of Disease Study", *Science*, vol. 274, 1996, pp. 1593-1594.
3. Council for International Organizations of Medical Sciences, *International Guidelines for Ethical Review of Epidemiological Studies* (Geneva, World Health Organization, 1991).
4. B. Brody, *The Ethics of Biomedical Research: an International Perspective* (Oxford, Oxford University Press, 1998).

5. S. S. Coughlin, "Ethics in epidemiology at the end of the twentieth century: ethics, values and mission statements", *Epidemiologic Review*, vol. 22, 2000, pp. 169-175.
6. N. Kass, "An ethics framework for public health", *American Journal of Public Health*, vol. 91, 2001, pp. 1776-1782.
7. D. Callahan and B. Jennings, "Ethics and public health: forging a strong relationship", *American Journal of Public Health*, vol. 92, 2002, pp. 169-176.
8. B. W. Levin and A. R. Fleischman, "Public health and bioethics: the benefits of collaboration", *American Journal of Public Health*, vol. 92, 2002, pp. 165-167.
9. M. J. Roberts and M. R. Reich, "Ethical analysis in public health", *The Lancet*, vol. 359, 2002, pp. 1055-1059.
10. C. Fry, "Raising the profile of human research ethics in addictions research: a key role for addictions journals", *Addiction*, vol. 97, 2002, pp. 229-230.
11. P. Griffiths and others, "Drug information systems, early warning and new drug trends: can drug trend monitoring systems become more sensitive to emerging trends in drug consumption?", *Substance Use and Misuse*, vol. 35, 2002, pp. 811-844.
12. J. G. Bachman and others, *Smoking, Drinking and Drug Use in Young Adulthood: the Impacts of New Freedoms and Responsibilities* (Mahwah, New Jersey, Lawrence Erlbaum, 1997).
13. D. B. Kandel and K. Chen, "Types of marijuana users by longitudinal course", *Journal of Studies on Alcohol*, vol. 61, 2000, pp. 367-378.
14. J. C. Anthony and J. Helzer, "Syndromes of drug abuse and dependence", *Psychiatric Disorders in America*, L. N. Robins and D. A. Regier, eds. (New York, Academic Press, 1991), pp. 116-154.
15. Y. I. Hser and others, "A 33-year follow up of narcotic addicts", *Archives of General Psychiatry*, vol. 58, 2001, pp. 503-508.
16. S. L. Witkin, "Ethics-R-Us", *Social Work*, vol. 45, 2000, pp. 197-200.
17. J. L. Brody and H. B. Waldron, "Ethical issues in research on the treatment of adolescent substance abuse disorders", *Addictive Behaviors*, vol. 25, 2002, pp. 217-228.
18. R. E. Tarter and others, "Cognitive capacity in female adolescent substance abusers", *Drug and Alcohol Dependence*, vol. 39, 1995, pp. 15-21.
19. College on Problems of Drug Dependence, "Human subject issues in drug abuse research", *Drug and Alcohol Dependence*, vol. 37, 1995, pp. 167-175.
20. C. Grady, "Money for research participation: does it jeopardize informed consent?", *American Journal of Bioethics*, vol. 1, 2001, pp. 41-44.
21. R. Macklin, "'Due' and 'undue' inducements: on paying money to research subjects", *A Review of Human Subjects Research*, vol. 3 (Institutional Review Board, 1981), pp. 1-6.
22. P. McNeill, "Paying people to participate in research: why not?", *Bioethics*, vol. 11, 1997, pp. 390-396.
23. M. Wilkinson and A. Moore, "Inducement in research", *Bioethics*, vol. 11, 1997, pp. 373-389.
24. A. J. Ritter, C. Fry and A. Swan, "The ethics of reimbursing injecting drug users for public health research interviews: what price are we prepared to pay?", *International Journal of Drug Policy* (forthcoming).

25. J. Fitzgerald and M. Hamilton, "Confidentiality, disseminated regulation and ethico-legal liabilities in research with hidden populations of illicit drug users", *Addiction*, vol. 92, 1996, pp. 1099-1107.
26. W. Loxley, D. Hawks and J. Bevan, "Protecting the interests of participants in research into illicit drug use: two case studies", *Addiction*, vol. 92, 1996, pp. 1081-1085.
27. D. A. Lawlor and T. Stone, "Public health and data protection: an inevitable collision or potential for a meeting of minds?", *International Journal of Epidemiology*, vol. 30, 2001, pp. 1221-1225.
28. S. Wright, H. Klee and P. Reid, "Interviewing illicit drug users: observations from the field", *Addiction Research*, vol. 6, 1998, pp. 517-535.
29. C. Day and others, "Interviewing/safety protocol for the National Drug and Alcohol Research Centre" (Sydney, National Drug and Alcohol Research Centre Safety Committee, 2001).
30. G. Craig, A. Corden and P. Thornton, "Safety in social research", *Social Research Update*, vol. 29, 2002, pp. 1-6.
31. World Health Organization, "Putting women first: ethical and safety recommendations for research on domestic violence against women", Department of Injuries and Violence Prevention, undated.
32. C. Ijsselmuiden and R. Faden, "Research and informed consent in Africa: another look", *Health and Human Rights: a Reader*, J. M. Mann and others, eds. (New York, Routledge, 1999), pp. 363-372.
33. S. Sánchez and others, "Informed consent procedures: responsibilities of researchers in developing countries", *Bioethics*, vol. 15, 2001, pp. 398-412.
34. S. R. Benatar and P. A. Singer, "A new look at international research ethics", *British Medical Journal*, vol. 321, 2000, pp. 824-826.
35. D. L. Weed and R. E. McKeown, "Ethics in epidemiology and public health: I. Technical terms", *Journal of Epidemiology and Community Health*, vol. 55, 2001, pp. 855-857.
36. J. Rachels, *The Elements of Moral Philosophy*, 2nd ed. (New York, McGraw-Hill, 1993).
37. T. L. Beauchamp and J. F. Childress, *Principles of Biomedical Ethics*, 5th ed. (Oxford, Oxford University Press, 2001).
38. M. Somerville, *The Ethical Canary: Science, Society and the Human Spirit* (Viking/Penguin Books, Australia, 2000).
39. A. R. Jonsen, *The Birth of Bioethics* (Oxford, Oxford University Press, 1999).
40. J. A. Rawls, *Theory of Justice* (Oxford, Oxford University Press, 1971).
41. S. Hampshire, "Morality and convention", *Utilitarianism and Beyond*, A. Sen and B. Williams, eds. (New York, Cambridge University Press, 1982).
42. M. Nussbaum, "Non-relative virtues: an Aristotelian approach", *The Quality of Life*, M. Nussbaum and A. Sen, eds. (New York, Oxford University Press, 1993).

Offprint from Bulletin on Narcotics, vol. LIV, Nos 1 and 2, 2002

Epidemiology and policy: the post-war context*

V. BERRIDGE

Professor of History, London School of Hygiene and Tropical Medicine, London

ABSTRACT

Chronic disease epidemiology emerged after the Second World War as the dominant public health research technology. It was important in the developing relationship between research and policy. The post-war years saw the rise of the randomized controlled trial and of the "evidence-based" movement, which also influenced the drugs field. Research collaboration between the United States of America and the United Kingdom of Great Britain and Northern Ireland was strong. A multitude of theories has been used to explain the relationship between research and policy. The present article outlines four broad tendencies: the evidence-based model, the journalistic view, sociology of scientific knowledge and science policy approaches. Four examples from the field of substance use underline the relationship: the discovery by Doll and Hill of the relationship between smoking and lung cancer; the Ledermann hypothesis that limiting alcohol consumption in a society reduces drinking problems and its impact on the alcohol field; the study by Hartnoll and Mitcheson of prescribing injectable heroin versus oral methadone; and the evaluation of needle exchange in 1987. Conclusions are drawn about why research had an impact and the forms of impact that can be identified. Quantitative methodologies, epidemiology above all, dominated research in the post-war period. There has been a complex process of mutual accommodation between policy makers and researchers, to which this public health technology has been central.

Keywords: epidemiology; policy; smoking; drugs; harm reduction; alcohol; history.

Introduction

The relationship between scientific research and public policy has been much discussed in recent years, in particular in the area of health policy and health services. In the United Kingdom of Great Britain and Northern Ireland, increasing concern at the Department of Health that policy should be based on "relevant" research was

*The present article is based on "The impact of qualitative research on policy-making: setting the scene", in Understanding and Responding to Drug Use: the Role of Qualitative Research, J. Fountain, ed., European Monitoring Centre for Drugs and Drug Addiction, Scientific Monograph Series No. 4 (Luxembourg, Office for Official Publications of the European Communities, 2000).

reflected in the launching of a National Health Service research and development initiative in 1990. A parallel concern in clinical medicine resulted in the growth of "evidence-based medicine", premised largely on the methodology of the randomized controlled trial, the "gold standard" of post-war medicine. The demand for "evidence-based" policy-making has grown ever louder and has become an international movement. Drug policy and research have shared in those evidence-based tendencies. However, the history both of drug policy and of the rise of research in the post-war years largely remains to be written. The topics of the articles in this special issue give an idea of the range of areas that need to be covered.

The present article aims to look at both the rise of epidemiology in general as a significant technical tool in the period after 1945 and the historical roots of the "evidence-based" notion of research and policy-making in the United Kingdom. It contains a survey of the range of theories that have been brought into play to analyse and describe that relationship and a description of four case studies chosen from substance use: smoking, alcohol and drugs. Finally, some conclusions are reached about the effect of evidence and epidemiology on policy in the last fifty years.

The rise of chronic disease epidemiology

"Vital statistics" have been important in studying patterns of disease since the early sixteenth century in many European countries and were central to the nineteenth century sanitarian movement. In the United Kingdom, the Registrar General's Reports were matched by independent enquiries by local societies such as the Epidemiological Society of London. Until the Second World War this tradition of enquiry allied to reform was essentially focused on the control of epidemic and infectious disease; thereafter, patterns of mortality and morbidity were changing from epidemic to chronic disease, with cancers and cardiovascular disease figuring more centrally. The immediate post-war years saw the rise of a new style of epidemiology that increasingly provided the dominant framework for the explanation of public health issues. Chronic disease epidemiology, with its use of probabilistic statistics and the concept of "risk", came to provide an autonomous mode of explanation. The British and American investigations of the relationship between smoking and lung cancer, which are discussed below, were milestones in the rise of modern epidemiology. The subsequent Framingham heart study in the United States of America provided epidemiology with a key role in the study of cardiovascular disease. This was a period in which risk factor epidemiology, with its technical tools like the odds ratio, case control and prospective studies, was established, not without controversy, as the dominant mode of investigation of post-war public health issues. It was the era of the science of chronic disease. Epidemiology established a hegemony equivalent to that of laboratory science and the probabilistic mode of statistical reasoning established its legitimacy in opposition to older genetic traditions within statistics. The culmination of these developments, at least as far as the United Kingdom was concerned, was marked by Bradford Hill's postulates for epidemiological causation enunciated in a lecture and published in 1965 [1, 2].

The history of research and policy

These technical developments took place during a period in which public health was reoriented to focus on the role of individual lifestyle and the individual's responsibility for his or her own health. Within government, there was an increasing emphasis on the use of research to inform policy. The idea that policy should be evidence-based emerged at that time. In the United Kingdom in the early 1970s, the Rothschild Report emphasized the role of government departments as clients or customers of the outside researchers from whom they commissioned research, which was to be developed on a customer-contractor basis. The danger of that approach, as some observers pointed out at the time, was that research was developed according to the view from the policy end rather than the more open-ended approach to funding that had previously been the norm. Although the history of those developments where drugs were concerned in the United Kingdom and elsewhere remains to be written, central to it would be the rise of the so-called drug problem in the 1960s and the perception by government and research councils, in particular the Medical Research Council, of the need for a better understanding of drug use to inform policy. The establishment of the Advisory Council on the Misuse of Drugs in the early 1970s, of the Addiction Research Unit at the Institute of Psychiatry and the use of notification as a research and surveillance tool all testified to the strong epidemiological dimension to drugs research [3]. Relationships between researchers in the United Kingdom and United States were close and visits to the United States by leading researchers from the United Kingdom had significant impact on ideas about research, in particular in the field of epidemiology.

Theories on the relationship between research and policy

A variety of different literatures and disciplinary traditions analysing the relationship between research and policy has developed, in particular since the 1980s [4]. In the present article, four broad tendencies are surveyed: the evidence-based medicine and health policy model, the journalistic view, sociology of scientific knowledge, and policy science and science policy approaches, all of which share areas of overlap.

The evidence-based medicine movement is much concerned with the concept of clinical effectiveness and is increasingly allied with the methodology of the randomized controlled trial; it is underpinned by positivist models of science and rational models of policy-making. Supporters of the movement believe that research, if properly funded and correctly positioned, can and should have an influence on policy, either directly or in some more diffuse way; what is important is to secure a working relationship between the two.

The journalistic view also sees the relationship between research and policy-making as desirable, but in a more partisan way. At times of crisis conspiracy theories abound, with key participants blamed for having been slow to act or

having failed to act. Such criticism is common not only in journalistic accounts, but also in academic analyses: recent examples are the crises of acquired immunodeficiency syndrome (AIDS) and bovine spongiform encephalopathy.

The third tendency, sociology of scientific knowledge, does not presuppose this rationality. Latour, for example, provides a model of the "actor networks" that sustain the research process [5]. The strength of any scientific claim is based on the resources, whether people, organizations, other disciplines or objects, from which its proponents are able to derive support. Much of this work is concerned with themes of emergence and resistance in science, an inward-looking perspective that does not take account of policy development.

This has been a concern of the policy science and science policy literature. Here, the policy dimension to the production of knowledge and its validation has been more central. Jasanoff, for example, has drawn attention to the differential impact of the "same science" in different national and policy contexts [6]. There is the "co-production" of knowledge, where government agencies also negotiate the meaning and the boundaries of science. The concept of the "policy community", linking scientific communities and government in various forms, is of importance here [7]. Theories of various types in this field place emphasis on the role of networks in policy, networks that facilitate the interchange of scientific orthodoxies between researchers and policy makers and that underpin a reciprocal relationship between the two.

Four examples of the relationship between research and policy for substance use

In this theoretical context four key paradigmatic pieces of research will be considered [8]. The research carried out by Richard Doll and Sir Austin Bradford Hill, which was first published in 1950, was a case-control and subsequently a prospective epidemiological study that demonstrated the connection between the rise in lung cancer and the habit of smoking. It ultimately led to a new policy agenda for smoking and for public health more generally, stressing price mechanisms, advertising controls, and the role of the mass media and health education. Through a focus on epidemiological modes of proof and on public health epidemiology rather than the genetic and hereditarian modes of biostatistics, it marked a decisive "paradigm shift" for both science and policy.

The Ledermann hypothesis had little influence when it was first advanced in the 1950s, but came into its own in the 1970s: Sully Ledermann was a French demographer and statistician who suggested that there was a relationship between average per capita levels of alcohol consumption and the general level of alcohol misuse in a population. In the 1950s, the disease view of alcoholism was dominant, while population approaches such as that which Ledermann proposed had connotations of temperance and morality rather than science. Only in the 1970s did the policy agenda change, when a coalition of doctors specializing in the treatment of alcoholism and alcohol problems, civil servants, the voluntary sector that provided treatment and support for alcoholics and those with alcohol

problems, the police and the legal profession rallied behind a call for the limitation of availability and harm based on the Ledermann hypothesis.

The study by Hartnoll and Mitcheson of prescribing injectable heroin versus oral methadone was carried out in the 1970s in a drug dependence unit in London and randomly allocated addicts to injectable heroin or oral methadone. Although the researchers themselves had not sought to change prescribing policy from heroin to oral methadone, towards what was viewed as a more "confrontational" treatment response, their research was widely credited with having done so. The study coincided with the introduction of oral methadone and there is some evidence that the change in policy took place before the research had been completed and the results disseminated. Here was research support for a change that was already under way.

Finally, the needle exchange evaluation carried out in 1987, which was hastily funded by the United Kingdom Department of Health to see if such "harm-reduction" approaches were an appropriate means of dealing with the threat of the spread of human immunodeficiency virus from drug users into the general population appeared to prove that syringe exchange "worked" and was a significant factor in unlocking government support for harm reduction and for increased funding for such services. Here was apparently a classic example of the influence of research on policy [1].

Why did research have an impact?

The following section consists of an attempt to deconstruct these historical episodes and identify any common variables or any specific factors that appear to have influenced particular situations. It should be noted that some research, such as that on needle exchange programmes, had a direct impact, while other research, such as the studies on smoking and alcohol, had an indirect impact, in that it changed the climate of opinion in which policy was formed.

Taking into account that difference, some common variables and some issues specific to particular cases can be identified. In all the case studies cited in the present article, the primary methodology was quantitative rather than qualitative and that reflects the dominant statistical and epidemiological paradigm of post-war research, a paradigm for which the research by Doll and Hill was significant. These case studies show that statistical and epidemiological methodologies were dominant in post-war public health research. However, methodology alone does not explain the reason for the policy impact of these studies. The case studies demonstrate the importance of the policy alliances that support science within government and reveal the ways in which science and research have provided coherence for policy communities. Within that relationship, the perspective of change over time is important. Science that had had little impact on policy in the conditions of the 1950s acquired the status of orthodoxy by the 1970s, when the policy situation itself had altered. The case studies on smoking and alcohol provide examples of that process. By the 1970s, the research on smoking was emblematic of a new public health constituency in the United Kingdom that

strongly advanced the policy case. Similarly, Ledermann, who had been ignored in the conditions of the 1950s, was the hero of the new alcohol lobby by the 1970s. The other case studies cited in the present article also illustrate these processes at work. The Hartnoll and Mitcheson study appealed to a clinic worker constituency which, with psychiatrists, was convinced of the need for new "active treatment" policies. As far as AIDS as a policy issue was concerned, for some while harm reduction had been the unspoken objective of a new and broader health policy community around drugs. The AIDS crisis and the funding and results of research, which were carefully managed, enabled deeply held political objections to be overcome.

Here is a symbiotic process in which the validation of science is not just an internal matter, as in Latour's model, but is rather a process of mutual accommodation with policy alliances and interests. For example, in the United Kingdom, links with government have been important, in particular links with civil servants and expert committees. In terms of government policy on smoking, the Chief Medical Officer, Sir George Godber, played a central role in the formation of the new public health agenda and later, in the era of AIDS and drugs, the civil servant responsible for drug policy, Dr. Dorothy Black, ensured that research results were presented in ways and situations that were acceptable to politicians. Committees have also played an important role. The reports on smoking published by the Royal College of Physicians, in particular those of 1962 and 1971 which had a considerable impact, gave independent authority to "scientific facts" and made them widely available through the media. In the debate over methadone versus heroin prescription surrounding the study by Hartnoll and Mitcheson, the committee of London drug psychiatrists that met at the Department of Health in the 1980s was important in determining prescribing policy. The reports on AIDS and Drug Misuse of the Advisory Council on the Misuse of Drugs, in particular the Part One report of 1988, provided important legitimacy for research results and the concept of harm reduction. These gatekeeping institutions or individuals have been important in the relationship between research and policy. In the post-war period, the media has played the role of mediator between research and policy: one example is the part played by the media in publicizing the early reports on smoking of the Royal College of Physicians and putting the "scientific facts" into the public and policy domain [1].

Some areas of policy have their own particular traditions. Health policy in the United Kingdom has been characterized by a tension between central direction and local self-determination. This is particularly clear in the field of drugs, where policy initiatives often emerge initially at the local level, as happened in the case of the Hartnoll and Mitcheson research, which was supported by a local, London-based, committee. Needle exchange also emerged first as a local initiative.

Crisis, too, is an important variable, well known to historians, who have analysed and debated its impact in, for example, war-time situations. The crisis of war has often led to more radical change in policy than would otherwise have been tolerated. The AIDS and drug policy case study of the introduction of harm-reduction cited above shows this clearly. This enabled what was essentially a

political change to be masked as a technical issue, as a matter of research results, made essential by the crisis of the day.

Clearly, the operative factors in different policy arenas may well be different. These case studies show that the alcohol and smoking arenas have been historically more diffuse than the arena for drugs, with a greater variety of interests involved, not least the industrial and legitimate economic interests within government [9]. The absence of these interests for drugs has given it a smaller and potentially more cohesive policy community and perhaps made the relationship between research and policy easier to establish. But this is to write from the British perspective. The nature of the structures and of the interrelationships with government will obviously vary in different national cultures. In some European countries such as the Netherlands, the relationship between researchers and policy-making civil servants has been close in recent years, while in others it has not been possible for research to be part of the process. In the United States, although there were moves in New York to establish needle exchange as a "controlled trial", that is, as a scientific procedure, these moves foundered on the opposition of local political interests, most notably African-American politicians, who condemned harm reduction as potential genocide for their constituency [10]. The fact that institutions and structures and local, federal or national political cultures vary has also to be taken into account.

The role of researchers in these processes has also changed over time. The role of the researcher as an active participant in the policy-making process has developed in the post-war period. Sir Austin Bradford Hill, in the 1950s, was firmly of the opinion that his role was to do scientific research and that it was the responsibility of the Chief Medical Officer and others to promote the policy options. He believed that too great an involvement in policy-making by scientists would undermine the independence and unbiased nature of science. Researchers such as Mitcheson took the view that their results were "hijacked" by policy makers to draw conclusions that they had not intended. Changes in research policy since the first smoking results were published in the 1950s have drawn researchers more closely into the policy-making process. However, they appear rarely to have been in control of the use made of their results.

Just four case studies are presented in the present article, but there are many others that would repay attention. Take, for example, the role of "community epidemiology", its development in the United States and its dissemination elsewhere. In the United Kingdom, community epidemiology was developed in the 1980s through the funding of local drug indicator studies by the Department of Health [11]. This research funding had a strong policy intent. It was part of a more general attempt to move drug services away from an exclusively London, psychiatrist-dominated, focus and to develop a broader range of services more accessible to users. Assessing the number of users that required drug services was a crucial dimension both of knowledge and of policy development. The expansion of these types of study into the European arena through the multi-city study on drug abuse of the Pompidou Group of the Council of Europe and the formation of the European Monitoring Centre for Drugs and Drug Addiction draws attention to

another aspect of the recent history, which is the growth of a strong European epidemiological dimension to the relationship between research and policy.

The "politics of evaluation" have been underlined more recently by the evaluation of the various heroin trials, in Switzerland in particular, and the controversies that have surrounded them [12, 13]. The "Cannabis Warning Scheme" introduced in July 2001 in the London borough of Lambeth, allowing those caught in possession of cannabis to be let off with a warning, was legitimated by widely quoted research demonstrating that police time could be saved and hence costs reduced [14, 15]. In those case studies, the policy dimension of research was of crucial importance. The use of research and the "value-free" scientific model underlines an attempt to depoliticize controversial policy issues and to present them to outsiders as simply technical questions to which research will provide the answer.

It is clear from the above discussion that the type of "rational" relationship between research and policy presupposed by the evidence-based medicine and policy model is rarely the case. Policy alliances both around science and in relation to policy are variable and the contexts of time, culture and country have to be taken into account. Overall, the case studies presented here demonstrate two things: one methodological and one that relates to policy processes. The dominant policy impact since the 1950s has been through quantitative methodologies, among which epidemiology has assumed increasing importance. Evaluation is also a policy tool. These methodologies are intricately bound up with policy processes and alliances. There is much in these examples to support the cynic's view that policy makers use research as a drunk uses a lamp-post, for support rather than for illumination. There is also, however, a more complex process of mutual accommodation between researchers and policy makers through various networks; this is Jasonoff's "co-production" of knowledge [6]. The role of epidemiology in that process became increasingly important in the post-Second World War era. However, the history of developments in terms of drugs and other substances in the last fifty years remains largely unexplored.

References

1. V. Berridge, "Science and policy: the case of post-war smoking policy" in *Ashes to Ashes: The History of Smoking and Health*, S. Lock, L. Reynolds and E. M. Tansey, eds. (Amsterdam, Rodopi, 1998).
2. L. Berlivet, "Association and causation: the debate on the scientific status of risk factor epidemiology and its impact on health policies", in *Networks of Knowledge: Health and the Relationship Between Science and Policy Since 1945*, V. Berridge, ed. (Amsterdam, Rodopi, forthcoming).
3. G. Edwards and C. Busch, *Drug Problems in Britain: A Review of Ten Years* (London, Academic Press, 1981).
4. V. Berridge and J. Stanton, "Science and policy: historical insights", *Social Science and Medicine*, special historical issue, Science Speaks to Policy, vol. 49, No. 9 (1999), pp. 1133-1138.

5. B. Latour, *Science in Action: How to Follow Scientists and Engineers Through Society* (Milton Keynes, Open University Press, 1987).
6. S. Jasanoff , "American exceptionalism and the political acknowledgement of risk", *Daedalus*, Fall 1990, pp. 61-81.
7. A. G. Jordan and J. J. Richardson, *British Politics and the Policy Process* (London, Allen and Unwin, 1987).
8. V. Berridge and B. Thom, "Research and policy: what determines the relationship?", *Policy Studies*, vol. 17, No. 1 (1996), pp. 23-34.
9. M. Read, *The Politics of Tobacco: Policy Networks and the Cigarette Industry* (Aldershot, Avebury, 1996).
10. W. Anderson, "The New York needle trial: the politics of public health in the age of AIDS", in *AIDS and Contemporary History*, V. Berridge and P. Strong, eds. (Cambridge, Cambridge University Press, 2002).
11. R. Hartnoll and others, "Estimating the prevalence of opioid dependence", *Lancet*, vol. 1, 1985, pp. 203-205.
12. S. L. Satel and E. Aeschbach, "The Swiss heroin trials: scientifically sound?", *Journal of Substance Abuse Treatment*, vol. 17, No. 4 (1999), pp. 331-335.
13. M. Farrell and W. Hall, "The Swiss heroin trials: testing alternative approaches", *British Medical Journal*, vol. 316, No. 7132 (1998), p. 639.
14. *Sunday Times*, "Is Tony Blair turning a blind eye to cannabis?", 2001.
15. K. Ahmed, "Cannabis users facing caution instead of arrest", *Observer*, 1 September 2002.

UNITED NATIONS
Office on Drugs and Crime

Vienna International Centre, P.O. Box 500, A-1400 Vienna, Austria
Tel: (+43-1) 26060-0, Fax: (+43-1) 26060-5866, www.unodc.org

كيفية الحصول على منشورات الأمم المتحدة

يمكن الحصول على منشورات الأمم المتحدة من المكتبات ودور التوزيع في جميع أنحاء العالم. استعلم عنها من المكتبة التي تتعامل معها أو اكتب إلى: الأمم المتحدة، قسم البيع في نيويورك أو في جنيف.

如何购取联合国出版物

联合国出版物在全世界各地的书店和经售处均有发售。 请向书店询问或写信到纽约或日内瓦的联合国销售组。

HOW TO OBTAIN UNITED NATIONS PUBLICATIONS

United Nations publications may be obtained from bookstores and distributors throughout the world. Consult your bookstore or write to: United Nations, Sales Section, New York or Geneva.

COMMENT SE PROCURER LES PUBLICATIONS DES NATIONS UNIES

Les publications des Nations Unies sont en vente dans les librairies et les agences dépositaires du monde entier. Informez-vous auprès de votre libraire ou adressez-vous à: Nations Unies, Section des ventes, New York ou Genève.

КАК ПОЛУЧИТЬ ИЗДАНИЯ ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ

Издания Организации Объединенных Наций можно купить в книжных магазинах и агентствах во всех районах мира. Наводите справки об изданиях в вашем книжном магазине или пишите по адресу: Организация Объединенных Наций, Секция по продаже изданий, Нью-Йорк или Женева.

CÓMO CONSEGUIR PUBLICACIONES DE LAS NACIONES UNIDAS

Las publicaciones de las Naciones Unidas están en venta en librerías y casas distribuidoras en todas partes del mundo. Consulte a su librero o diríjase a: Naciones Unidas, Sección de Ventas, Nueva York o Ginebra.

Printed in Austria
V-03-87283—October 2003—2,080
United Nations publication
Sales No. E.03.XI.17
ISBN 92-1-148170-8
ISSN 0007-523X