EMCDDA SCIENTIFIC MONOGRAPH SERIES No 6

Modelling drug use: methods to quantify and understand hidden processes

E.M.C.D.D.A.
European Monitoring Centre for Drugs and Drug Addiction

EMCDDA SCIENTIFIC MONOGRAPH SERIES No 6

Modelling drug use: methods to quantify and understand hidden processes

PLANNING GROUP
Lucas Wiessing
Christine Godfrey
Richard Hartnoll
Matthew Sutton
Philip Young

SCIENTIFIC EDITORS
Christine Godfrey
Lucas Wiessing
Richard Hartnoll

EDITORS
Frances Sharp
Rachel Neaman

E.M.C.D.D.A.
European Monitoring Centre for Drugs and Drug Addiction

Detailed information on the EMCDDA and its work can be found on its website **(http://www.emcdda.org)**

A great deal of additional information on the European Union is available on the Internet. It can be accessed through the Europa server (http://europa.eu.int).

Cataloguing data can be found at the end of this publication.

Luxembourg: Office for Official Publications of the European Communities, 2001

ISBN 92-9168-056-7

Printed in the United Kingdom

CONTENTS

FOREWORD
Georges Estievenart 7

GENERAL INTRODUCTION
Lucas Wiessing and Richard Hartnoll 11

PART I: WHAT IS MODELLING AND HOW CAN IT BE USED?

INTRODUCTION 17

CHAPTER 1
Epidemiology of drug use at macro level: indicators, models and policy-making
Lucas Wiessing, Richard Hartnoll and Carla Rossi 19

CHAPTER 2
What are dynamic models and how can they be used?
Christine Godfrey, Steve Parrott, Matthew Sutton, Vanessa Waby, Philip Young with Catherine Comiskey, Martin Frischer, Gordon Hay, Matthew Hickman, Johannes Jager, James Kahan, Ludwig Kraus, Roberto Mollica, Maarten Postma, Roland Simon, Colin Taylor and Lucas Wiessing 33

PART II: MODELLING DRUG USE WITH AVAILABLE DATA

INTRODUCTION 53

CHAPTER 3
Modelling drug misuse in Europe using geographic information systems
Martin Frischer and Heath Heatlie 55

CHAPTER 4
Using available data with dynamic models to estimate the extent of problem drug use
Carla Rossi 67

CHAPTER 5
Multiple-indicator models
Roberto Mollica and Fabio Mariani 91

PART III: MODELLING TECHNIQUES AND THEIR USES

INTRODUCTION 103

CHAPTER 6
The method of back calculation as a means of estimating incidence
Catherine Comiskey and Gordon Hay 105

CHAPTER 7
Compartmental modelling and stochastic dynamic systems
Gordon Hay and Catherine Comiskey 117

CHAPTER 8
Dynamic models in drug epidemiology
Colin Taylor and Philip Young 133

CHAPTER 9
Structural equations and path analyses: using multiple indicators of drug prevalence
Colin Taylor and Philip Young 149

PART IV: MODELLING THE HEALTH CONSEQUENCES OF DRUG USE

INTRODUCTION 167

CHAPTER 10
Modelling infectious diseases and other health consequences: predicting future morbidity consequences and spread of HIV, hepatitis B and hepatitis C
Mirjam Kretzschmar 169

CHAPTER 11
Developing multinational scenario analyses of health impacts of drug use
Johannes C. Jager, Maarten J. Postma and Peter W. Achterberg 183

PART V: ECONOMIC MODELS

INTRODUCTION 203

CHAPTER 12
Modelling the health-care costs of drug-use-related disease
Maarten J. Postma, Keith Tolley and Johannes C. Jager 205

CHAPTER 13
Economic market models
Matthew Sutton
221

GENERAL CONCLUSIONS
Richard Hartnoll and Lucas Wiessing 233

CONTRIBUTORS 239

FOREWORD

It is with great pleasure that I present this scientific monograph on the dynamic modelling of drug use. This publication provides a rich overview of the application of mathematical and statistical techniques in the field of drug use and it clearly points to the further potential for using models for analysing and understanding drug data. The monograph reveals a growing interest on the European and international stages for adding such complementary quantitative approaches to the already wide array of tools in the social and life sciences. In a field that is quickly moving and changing appearance, within the no-less dynamic European context, I believe that dynamic modelling will help decision-makers and scientists to better understand hidden processes of drug use, epidemiological trends and their relationships, and thus will contribute to sound decision-making on policies and interventions.

This monograph reflects the efforts and expertise of many people. Above all, I want to thank the experts who committed their time and work to discuss the future of modelling drug use at two meetings in York, UK, in May and November 1997, and at a seminar in Lisbon, Portugal, in May 1998. In particular I would like to thank: Azzedine Boumghar; Sandeep Chawla; Catherine Comiskey; Almudena De Silva Rivera; Martin Donoghoe; Martin Frischer; Christine Godfrey; Richard Hartnoll; Gordon Hay; Mathew Hickman; Johannes Jager; James Kahan; Ludwig Kraus; Mirjam Kretzschmar; Leif Lenke; Ed Leuw; Roberto Mollica; Steve Parrot; Maarten Postma; Peter Reuter; Carla Rossi; Gianpaolo Scalia Tomba; Roland Simon; Matthew Sutton; Colin Taylor; Gernot Tragler; Lucas Wiessing; and Philip Young. Special thanks are due to: Vanessa Waby; Monica Blum; Sonia Collaço; and Manuela Gomes for their organisational help.

Georges Estievenart
Executive Director
EMCDDA

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Lucas Wiessing and Richard Hartnoll

One of the central tasks of the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) is to analyse and make sense of data on drug use and its consequences in the European Union. This is a basic prerequisite for developing rational interventions and policies in response to drug problems. In this monograph, the potential role of modelling in this task is explored. Modelling is a way to simplify and understand complex processes or structures and is usually performed using mathematical and statistical techniques. It can increase the value of available data, for example by analysing key dimensions of large or complicated data sets. Modelling can also be used to estimate missing data, to forecast trends and to carry out virtual experiments on the likely outcomes of different scenarios. Drug use and problems involve complex information, and modelling may also help to enhance the interpretation of available but scarce data and to increase the understanding of unobservable processes and relationships. It can serve especially to describe the dynamics of drug use at a more abstract or aggregated (local, national, international) level.

It is this high level of abstraction that may make modelling somewhat difficult to 'sell', especially to those professionals who are nearer to the individual drug user, as the results of modelling studies seldom relate to individual experiences. It is exactly this abstraction, however, that gives modelling a strength of its own, by detecting common features within available data sets or by clarifying the relationships between more abstract concepts such as 'law enforcement', 'price' and 'demand for drugs'. This often results in more generalisable knowledge than can be achieved through direct observation, although of course data collected through direct observation form an indispensable basis for building any useful models.

Several problems arise when trying to analyse data in a way that is relevant to policy questions. Where there is a large amount of data covering different aspects of a complex situation it is hard to identify and analyse the key dimensions and relationships within the data. In other situations, despite a lack of data, there are a number of alternative scenarios of how different policy options may affect drug problems, and policy decisions will anyway be taken even if not well grounded on empirical knowledge. Modelling techniques, such as scenario analysis, can then at least give some formal guidance.

To explore the potential for using modelling approaches in the drugs field, the EMCDDA organised two meetings of modellers and a larger seminar with other European drug experts during 1997 and 1998. This monograph contains a number of reviews of different aspects of modelling work in the drugs field written by the participating experts. It aims to provide a broad overview of quantitative approaches to understanding drug data as a knowledge base and a first step towards more focused modelling projects.

Structure of the monograph

Part I of the monograph is introductory in nature. In Chapter 1, the use of modelling is discussed for macro-level, rather than individually based, epidemiology, and models are presented as complementary tools to direct empirical observation of data or indicators. A distinction is made between statistical models that are inductive and rely heavily on data, and mathematical models that are more deductive and serve to explore potential relationships in a more theoretical manner. It is stressed that the interpretation of modelling results is not straightforward and should take careful account of the validity of the model and its assumptions. A number of broader issues are discussed that either can be investigated with models or should be taken into account before modelling can be useful, notably a common understanding of policy objectives and case definitions. In Chapter 2, an overview is provided of the discussions and conclusions of the modelling experts who contributed to the monograph. The chapter starts with an overview of different policy questions. Varying ways of categorising models are then discussed, and the chapter concludes with an account of discussions about different priorities for future developments in modelling across Europe based on the seminar held in Lisbon in May 1998.

Three examples are given in Part II of models that integrate available data sources or indicators at macro level to describe and predict the spread of problem drug use in time and space. In Chapter 3, a geographic model is presented designed to predict the spread of drug use at macro level from larger cities to smaller towns. It is based on a small number of parameters such as those used in models of the spread of infectious diseases (length of epidemic cycle, probability of infection as a function of distance) and linked to a geographic information system (GIS) to enable the results to be visualised in maps. In Chapter 4, the use of data on trends is discussed in an interactive way, allowing for interdependencies between production processes of the observed data. The importance is stressed of using simple aggregate models for studying policy options rather than complex models of individual spread. An overview is provided of classical and new mathematical models that use routine data, and examples of their application, to estimate the extent of problem drug use in different countries. Chapter 5 follows a similar approach to estimate trends and interdependencies in indicators based on routine drug data. The model is simultaneously validated by new data while estimating and forecasting variations in problems and need for services over time. Illustrative examples are set out for describing how such a model could be used to assess the effectiveness and efficiency of new drug laws and therapeutic approaches.

The chapters in Part III take as their starting point available standard modelling techniques rather than available drug data sources or indicators. In Chapter 6, there is discussion of a method that has been very useful in generating predictions of AIDS and HIV, the back-calculation method. Based on discussions in the expert meetings held in York, it is proposed to use back calculation to estimate incidence or rates of initiation of heroin use from routine drug-treatment data. Knowing incidence rates would enable more direct evaluation of activities that aim to prevent drug users from starting to use heroin than is possible with prevalence estimates. In Chapter 7, compartmental models are discussed. As an example, a classical compartmental model for the spread of measles is

presented, illustrating the basic concepts of the infectious spread of a phenomenon which can be described quite satisfactorily with only a few parameters. The use of compartmental models is recommended for analysing drug careers, as has been done for several diseases, as well as for studying the processes that lead young people to initiate drug use. In Chapter 8, different dynamic models are discussed to estimate prevalence of drug use by describing the unobservable underlying processes. A series of basic issues are highlighted, among others the time frame and geographic boundaries of a prevalence-estimation exercise, the types and amounts of drugs used and the relation between definitions of drug use and the use of data sources. System-dynamic models are also presented, which have mostly been used to describe macro-scale drug phenomena and national policy scenarios, such as drug imports, drug treatments and law enforcement. They are relatively well suited for macro-level analysis as they model flows as aggregate measures rather than keeping track of individual persons or events. In Chapter 9, the use of factor analysis is discussed to confirm or explore underlying latent variables from observed variables. Four classes of observable variables are presented: frequency of use; amount used; amount of disruptive use (e.g. use at school or at work); and subjective perception of problems with use. It is argued that working with composite latent variables is often nearer to reality than interpreting use of each drug separately, as most drug users are poly-drug users and there are different and combined ways of using (e.g. moderate or heavy alcohol use in combination with other drugs). Structural equation models and linear structural relationship (LISREL) models that allow more flexibility but also introduce more complexity are also reviewed. The authors conclude that these models are difficult to manage and of lesser practical value, although they are potentially very useful to fit underlying theoretical concepts into models of drug prevalence.

Part IV centres around modelling health consequences of drug use rather than types of models. In Chapter 10, the main disease parameters needed to model infectious diseases in injecting drug users are discussed. Concepts from modelling studies are introduced, such as 'core groups' and 'mixing patterns', that have proved crucial to understanding the spread of disease and that might prove equally important in understanding the spread of drug use. A picture of the historical development of sexual-transmission models is given and several classic models of infectious diseases among drug users are reviewed. In Chapter 11, a structured overview of how to develop multinational scenario analyses is given together with an example for AIDS. In this approach, different disciplines, including mathematical modelling, are integrated to develop insight and a basic description of a problem, as well as to forecast the most probable future course and alternative scenarios including the estimated effect of interventions. It is argued that scenario analysis is well suited to analysing drug problems, as these are serious and important, complex, multinational, multidisciplinary and in need of (quantified) policy decisions. Scenario analyses are closely related to cost-effectiveness analyses for the allocation of health-care resources and both types of analysis can be performed jointly.

Part V of the monograph addresses economic aspects and modelling, starting with a focus on estimating drug-related health-care costs (Chapter 12). In general, a costing model aims to assess current and to project future health-care needs and financial requirements. The main step is to link an epidemiological sub-model which describes

the future development of disease to specifically formatted information on health-care needs (e.g. lifetime patient costs). The approach can be incidence- or prevalence-based. The former links estimated lifetime health-care needs and their costs (a staged-disease model of natural history is necessary) to epidemiological estimates of incidence in a certain year. This gives the potential future costs that can still be averted by prevention. The latter links current health-care needs and costs per patient (estimates of the prevalent patient mix are needed according to the stage of disease) to estimates of prevalence. This gives the health-care needs and costs of current disease with its immediate budgetary implications. It is recommended to extend the extensive existing work on HIV/AIDS to hepatitis as a first step, and subsequently to include other relevant drug-related diseases and consequences. In Chapter 13, the potential for modelling economic drug markets is described. It is argued that studies at individual level have the greatest potential, but the macro level is what policy-makers are most interested in and where most data are available. Studies at macro level (e.g. effect of enforcement), however, have not resulted in relevant outcomes so far, most probably due to a too-high level of aggregation. For a good markets study, it is recommended that sufficient resources be available for collecting high-quality data on consumption, prices and purities.

It is hoped that this monograph, and the newly formed European network of modellers behind it, will prove to be significant in stimulating the application of quantitative techniques to drug data. As good and comparable data on drugs are still rather scarce in Europe, progress is not expected to be very rapid. However, with increasing work on improving data sources and growing understanding of the basic knowledge needed for sound policy-making, mathematical and statistical modelling form an indispensable part of the set of tools that are available to understand and prevent the problems caused by the use of drugs.

What is modelling and how can it be used?

INTRODUCTION

Modelling can be a valuable complementary tool for policy decisions and can be especially useful at levels of larger aggregation (national, international). It is often, however, not well understood, both because non-modellers may have little understanding or interest in the mathematical presentations of modellers, and because modellers may have little understanding or interest in the practical consequences of their work and its relation to other disciplines, often being more interested in theoretical mathematical advances. Part I aims to discuss the potential for modelling in drug epidemiology in a manner accessible to non-modellers.

In Chapter 1, the authors discuss modelling approaches together with empirical data collection or in other words 'indicators' of drug use and problems. The emphasis is on the use of modelling and indicators as complementary tools for developing evidence-based policy decisions. Working at the macro level adds further complexity as data quality is often low. This affects both the use of indicators, which permit only limited interpretation and need to be complemented with studies, and modelling, of which the validity of results often crucially depends on the quality of data. Some general aspects of modelling are discussed, such as the differences in interpreting results from data-driven and more theoretical models. Models and indicators can generate knowledge for decision-making, but some general problems remain. These include lack of clarity in case definition and in the objectives of drug policies, particularly the distinction between non-problem and problem drug use. Lack of knowledge of the basic mechanisms of the spread and progression of drug use and problem drug use hinders and challenges modelling. Legal and moral dimensions also strongly hamper the formulation of clear scientific theory in several ways, of which the most obvious is their influence on basic data availability and quality.

Chapter 2 presents the discussions held at two meetings in York, UK, in 1997 and at a larger seminar in Lisbon, Portugal, in May 1998. A central issue when discussing the use of modelling is to start from the questions that should be answered rather than from the methods and techniques used. Questions by policy-makers tend to be simple and straightforward ('how many drug users are there?', 'what works?') but are not always easily answered. To be policy-relevant, questions and study objectives should in general relate to interventions or policy options, for example, what is the cost-effectiveness of one intervention compared to another? However, basic descriptive data must be available and may sometimes be generated with other types of models. Models have two broad types of application: the more quantitative estimations; and the more qualitative understanding of dynamic processes and relationships.

CHAPTER 1

EPIDEMIOLOGY OF DRUG USE AT MACRO LEVEL: INDICATORS, MODELS AND POLICY-MAKING

Lucas Wiessing, Richard Hartnoll and Carla Rossi

Indicators and models may help in studying and understanding the epidemiology of drug use at macro level. Indicators of drug use are mostly based on available routine data and give indirect information on prevalence and trends over time. They can be used to cover large geographic areas with a limited budget; however, they are usually limited in quality and scope. Modelling is based on mathematical theory and can be used to integrate data from different indicators and other sources. It can be used to estimate prevalence and incidence or to increase understanding of drug processes by simulating experiments that are not possible in real life. The interpretation of modelling results, however, crucially depends on sufficient understanding of their basic assumptions and limitations. A whole continuum of modelling may exist, depending on the availability and use of data, from empirical analyses with high applicability to real life to theoretical hypothesis-generating exercises based on many assumptions.

Compared to other areas, drug-use epidemiology has still made little use of modelling, partly because good-quality data are still scarce, but work on improving indicators and other data is in progress. In the drugs field, the application of evidence-based approaches is still hampered by the lack of a common understanding of priorities for study and intervention, a lack of knowledge of mechanisms of spread and progression to drug problems, and legal and moral discourses. Indicators and models are complementary and increasingly important tools for epidemiology at macro level and may help to clarify some of these issues. Although quality of data and thus possible inferences are often much more limited than in in-depth empirical studies, macro-level assessments are necessary to guide national and international policy decisions.

One problem of drug epidemiology at national or international (macro) level is the sheer scale of activities undertaken. While data on drug use can be collected directly within a city or small region, no funding is usually available for massive studies covering large areas or a number of countries. Even if it were possible, setting up such a study would not be straightforward, given important background differences between countries and regions and the many stakeholders involved. Data available at macro level are often aggregated rather than individual-level data, based on routine sources, with all their limitations of analysis and interpretation. The question is therefore (how) can valid inferences still be made at macro level? This chapter briefly introduces indicators

and modelling, two different but complementary tools that are receiving increasing attention for policy-making at macro level, and points to some still unresolved problems in the drugs field that interfere with such evidence-based approaches.

Indicators

Epidemiologists aim at describing the spread of disease in a population to provide evidence for public-health-oriented interventions and policies. Basic epidemiological measures are prevalence (all existing cases at a certain moment in time) and incidence (all newly occurring cases in a certain time period). To understand causal factors that lead to disease, a risk-factor analysis is usually performed by comparing individual cases with non-cases. In drug epidemiology it is usually difficult to determine prevalence and incidence, due to the 'hidden' nature of drug use. Normal sources of data such as medical services may be incomplete while standard tools such as household surveys may give biased results, due to social stigma and low social integration of heavy drug users. Working at macro level adds to these problems the difficulty of collecting comparable, quality data from a large number of sources and countries. As a result, additional methods are necessary to estimate or at least give some idea of prevalence, incidence and trends in drug use over time.

In domains like public health and the economy, the concept of indicators has been developed for large-scale data collection and for difficult-to-measure phenomena. Rather than aiming at exact prevalence and incidence figures, indicators are based on available or routine data on disease, risk factors or consequences, and may provide indirect information on the exact magnitude of and trends in the disease. Public-health indicators rely on data sources such as mortality registers, population censuses, routine health-service records, epidemiological surveillance data, sample surveys or disease registers. The idea is to select from these sources those variables that fulfil certain quality requirements (such as validity, reliability, etc.) to measure health and changes in health (WHO, 1999).

In the field of drug use, work on indicators was started in London in the early 1970s, taken to European multi-city level by the Pompidou Group of the Council of Europe, further developed at European multinational level by the EMCDDA and is currently being implemented at global level by the United Nations International Drug Control Programme (UNDCP) (Hartnoll, 1991; Hartnoll et al., 1989; EMCDDA, 2000). There is not much choice in available data on drug use, and pragmatic reasons have played an important role in the choice of indicators. Currently five so-called 'key indicators' are being implemented in the European Union. These are 'general population and school surveys', 'estimates of problem drug use', 'data from drug-treatment services', 'drug-related deaths' and 'drug-related infections' (HIV and hepatitis B and C). Standards have been developed and current work is focused on collecting the data and overcoming practical problems in data quality and comparability. Other indicators are at an earlier stage of development, such as indicators of social problems and crime, data from hospital and emergency rooms and data from the youth and dance scenes.

How to interpret drug-indicator data is not always straightforward. Sometimes it is simply assumed that the combined information from a set of indirect indicators reflects trends in the prevalence of problem drug use. However, indicator data may show fluctuations unrelated to prevalence, for example, due to changes in heroin supply or decreases in risk behaviour. Therefore indicators such as drug-treatment data, drug-related mortality and drug-related infectious diseases are also very important for their own sake and for evaluating interventions directed at each of the specific sub-populations at risk, rather than simply for following trends in 'problem drug use'.

Indicator data are often also of limited quality, due to their large coverage and routine nature, and it is dangerous to rely solely on indicators. Most routine data are not suitable for giving rapid insight into changing trends in problem use. For instance, the average latency time of five to eight years between first heroin use and first demand for treatment implies that treatment indicators are little suited to monitoring drug use in the non-treated population (Rossi, 1999; Wiessing et al., submitted). To interpret trends such as an increase in drug-related deaths at national level, and to take appropriate measures, indicators should be complemented and validated by smaller in-depth studies. Regularly repeated, small-scale studies are also important for gaining insight into possible (changing) causes or risk factors and confounding factors of the observed trends, as in such studies much more background information can usually be gathered. Routine indicators and one-off studies might together give a more complete picture of a phenomenon such as problem drug use. Indicators provide lower-quality data but at low costs and with high geographical coverage, while local studies provide high-quality data from a more limited (often high-risk) area.

Dynamic models

Dynamic modelling originated from biology and infectious-diseases epidemiology (Anderson and May, 1992). Many of its more recent advances were also made in the infectious-diseases field, especially within AIDS research (Pasqualucci et al., 1998). Over the last 15 years, modelling developments have closely paralleled the increasing knowledge of the etiology and transmission of AIDS and other sexually transmitted diseases. During that period, modelling studies moved from very simple AIDS data extrapolations to highly complex transmission models with large sets of parameters. Modelling and empirical research on AIDS have also mutually informed each other on important concepts. While patient characteristics pointed to the existence of specific groups at increased risk, such as homosexual men and injecting drug users, dynamic modelling clarified the potential role of core groups (with high-risk behaviour and/or assortative – 'like-with-like' – mixing patterns and consequently high infection levels) in the overall transmission dynamics within a susceptible population (Kretzschmar et al., 1996). While the natural history of HIV infection was becoming clearer from cohort studies, modelling provided a means of clarification and simplification (e.g. by determining a number of discrete disease stages) and a basis for policy decisions (e.g. on treatment needs per stage) (Longini et al., 1989). The effect of possibly important biological parameters which could not easily be studied empirically were often postulated

from dynamic modelling work. For example, a simulation model could show that the very high level of infectiousness during the first weeks of HIV infection may determine most new infections in the population, putting the effectiveness of HIV screening as a prevention tool into question (Daar et al., 1991; Kretzschmar and Wiessing, 1998).

For epidemiology at macro level, dynamic modelling can be a valuable complementary tool to following trends in indicators and other direct data analysis. Besides the more usual inductive or empirical methods of data collection and interpretation, where new insights follow from observation, dynamic modelling may often be nearer to a deductive approach, where new insights follow from theory. The use of the outcomes in real-life situations is sometimes quite limited and validation of the model assumptions to data can be very difficult. However, the importance of most types of dynamic modelling lies in the generation of theory that may provide a framework for practical decisions.

It is not easy to define dynamic models, and to understand how they are used it may be helpful to compare them to the more familiar statistical models used for empirical data analysis, although even this difference is not clearcut. Most models consist of a set of mathematical equations, often in the form of a computer program, that describe a process (e.g. the spread of drug use) in a very simplified manner, and that can be used to understand the behaviour of that process. There may be a whole continuum of models between the extremes of pure empirical data analysis with statistical models and using mathematical or economic theory or dynamic models. Statistical modelling techniques, such as different types of regression analysis (generalised equations), factor, cluster and path analysis, often rely much more on data than, for example, many compartmental or system-dynamic models, which are often more theoretical. Dynamic models may rely more heavily on sometimes unproven assumptions, and may often be based on 'thin' data the quality of which is not certain, which makes it much more difficult to keep track of their validity.

Figure 1: Some differences between statistical and dynamic modelling

Statistical analysis/modelling	**Dynamic/mathematical modelling**
Inductive	Deductive
Data-driven, fewer assumptions	Theory-driven, more assumptions
←	→
Estimations, assessments	**'Mind experiments', 'what if'**
Hypothesis testing	Hypothesis generating
Static analysis over time	Dynamic, feedback loops

Note: This scheme obviously does not fit all models. There are mathematical models that are very data-driven and statistical models that are based on very little data. Also some statistical analysis is not static over time (e.g. time-series analysis, survival analysis) and deterministic mathematical models can be used for estimations. Many models, however, appear to fit this scheme.

Both extremes of the continuum are useful in, for example, drug epidemiology, but their use may be very different. Rigorous statistical data analysis, such as regression analysis based on individual data, is mostly not possible at macro level, but statistical models can often still suggest useful answers when there are gaps in data of a certain type. This can, for example, happen by imputing prevalence data (extra- or interpolation) from other moments in time or locations where more data are available, or by using other data sources and, if possible, adjusting for potential biases. This makes it possible to generate a result where no data were available, although again the interpretation must be undertaken with care. There is a clear need for this type of modelling at macro level, where good data often do not exist, at least in the drugs field.

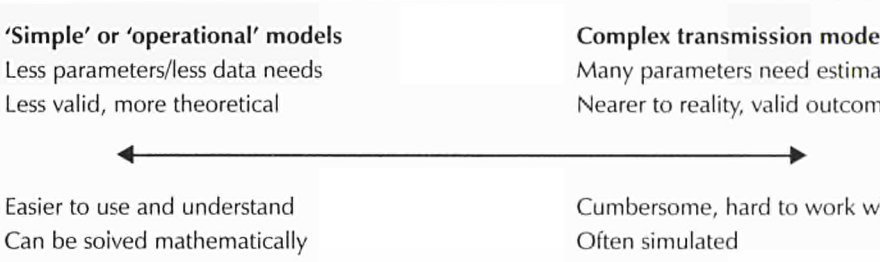
At the other end of the continuum, dynamic models can help in understanding a phenomenon, even if almost totally based on assumptions and with little or no data input. If the model is assumed to describe a process in real life sufficiently well, the behaviour of this process can be studied under different circumstances by varying parameter values and seeing the variation in outcomes. This semi-experimental situation is sometimes called 'what if' modelling, in other words, the question is what could happen under what circumstances if it is assumed that the process under study is well captured by the model. Modelling can thus provide a tool to simulate experiments that are not possible in real life for practical or ethical reasons.

It is important to understand the differences between the various types of modelling and to interpret the results accordingly, either as 'scientific facts', 'theoretical hypotheses' or something in between, depending on the certainty of parameters and the validity of the basic model chosen. Modelling results may involve many unproven assumptions, may be based on data of unknown quality and the true scope for bias may be much larger than, for example, suggested by calculated confidence intervals. To understand the results and correctly interpret them it is indispensable to have at least some understanding of the assumptions behind the model. The limitations of the model should always be clearly stated and presented together with the results. However, even for specialists this is often difficult, and non-specialists may have no way of distinguishing valid results with important direct implications from academically interesting theory with little direct practical meaning. To prevent this type of problem it is important for modellers and other scientists to work together in multi-disciplinary teams. This is, however, only possible if the different disciplines have some understanding of each other's approaches and are prepared to acquire at least some knowledge of the other field. If such a collaboration exists, the modeller may generate 'meaning' from scarce data, and the topic expert (e.g. epidemiologist) can help interpret the validity of the results against his or her knowledge of real-life situations and data-quality issues.

Within dynamic modelling different approaches are again possible. The starting point is often a relatively simple model that reflects the main elements and their relationships. However, the results are often criticised for not taking into account a possibly important variable. This may then subsequently be incorporated into the model and the model becomes more complex. This process may repeat itself resulting in highly complex models. Although these may resemble reality better than the simple models, it is often hard to keep track of what exactly happens and in practice they may not lead to better

or more valid results. As was the case in the AIDS field, the process of increasing complexity often directly reflects the state of the art of existing knowledge in the field. While some modellers attempt to incorporate as many parameters in their models as possible to increase the validity of their results, others prefer to work with simple models that can be easily understood. Simple models that succeed in giving new insights are often more influential (e.g. Kaplan and O'Keefe, 1989). Especially in the case of more complex models, sensitivity analysis is an important step. A sensitivity analysis looks at the influence of the different parameters on the modelling results by taking different, often extreme, values for these parameters. The model might then be found not to be sensitive to some of the parameters but very sensitive to others, which suggests how the model needs to be improved. The results of a complex modelling study can even be analysed applying statistical sampling techniques to the modelling results, as, for example, did Blower et al. (1991) in a study of HIV spread in New York.

Figure 2: Two basic approaches to dynamic modelling



An important question during the interpretation of modelling results is their intended use. Depending on the consequences, modelling results can be accepted either more or less conservatively. The same outcomes can be valid enough for publication in a high-level mathematical journal where the emphasis is on the methodological advances, while they may not be found sufficiently rigorous to be published in a high-level applied-science (e.g. public-health) journal which focuses on the external validity of the results and their practical consequences. Even for applied use, however, the question is often whether it is better to have less valid results than no results at all, which often leads to accepting outcomes that may be less valid than would be desirable from a strictly scientific point of view.

An example of the latter situation in drug epidemiology is prevalence estimation. Although sophisticated statistical models exist for local-level prevalence estimation of problem drug use, such as multi-sample capture–recapture with loglinear regression parameter estimation (Frischer et al., 1993), results are often not very clear due to data-quality problems and the often wide confidence intervals, which may still severely underestimate uncertainty. At national level, geographic heterogeneity and low data availability make it often impossible to use these methods. Simple multipliers are being used with no possibility of calculating confidence intervals (EMCDDA, 1999) but at best of giving some qualitative indication of a plausibility range (Uhl and Seidler, 2000).

Still it is very important to have prevalence estimates at local and national level, which form a basis for a large variety of policy choices, and these low-quality estimates are in general accepted as the state of the art. Other examples can be found in forecasts of future developments and in scenario analyses, which by definition are unreliable and may prove wrong in time, but are still of such high value for policy-makers that they are widely used.

Policy needs and problems using indicators and models

Epidemiology is mostly an applied or problem-driven science, where much emphasis is placed on the policy relevance of new knowledge and studies. It may therefore be useful here to discuss briefly what are the possible needs of policy-makers. This is especially important for epidemiology at macro level, where the link with the needs of individual drug users may be much less obvious than, for instance, in specific treatment or outreach settings.

The needs of a policy-maker may seem remarkably clear in general terms: to describe and understand a problem and follow trends over time; to design proper interventions; and to evaluate the results of interventions. In the case of drug use and its consequences, indicators and models may contribute to both a description and an understanding of the problem, while indicators are specifically suited to follow trends over time. Models may even help in designing and choosing proper interventions, for example, by assessing the cost-effectiveness of alternative options, although the basic knowledge of 'what works' (efficacy) needs to be assessed in carefully designed studies. Reality, however, seems harder than practice, and there is still little consensus among and between scientists, service providers and policy-makers about the most effective and relevant interventions, with opinions ranging from a total 'war on drugs' at one extreme to complete legalisation on the other. What is it that makes the drugs field seem so much more complex than would initially be assumed? The answer may partly be purely scientific/methodological, and partly lie in the interplay between science and policy. Some of the more methodological problems may also be a consequence of working at a high level of aggregation. Other problems relate to the lack of a clear case definition of a 'drug user', to lack of knowledge of the basic mechanisms of the spread of drug use, and lack of knowledge of progression into drug problems.

First, using indicators at a macro scale, it is often not possible to use a clear case definition of addiction or dependence such as the *Diagnostic and statistical manual* (DSM) IV, or to use disease codes such as the international classification of diseases (ICD) due to the quality and limitations of routine data. Indicators such as survey data may give good-quality information on lighter forms of drug use, but usually not on heavier patterns of drug use which cause most drug problems. Modelling techniques for prevalence estimation of problem drug use, such as capture–recapture, often rely on non-specialised data sources which have little information on drug use and thus do not permit using clear case definitions. A pragmatic solution may in that case be found by defining problem drug users as drug users that are in contact or in need of contact with health or social services, which, for example, in most European countries is mostly limited to

frequent heroin or amphetamine users. However, it is hard if not impossible to obtain more detail on the prevalence and patterns of problem drug use (e.g. breakdowns by type of drug) from these techniques and it is important that they are complemented by (expensive) local studies based on, for example, out-of-treatment recruitment of drug users.

Second, much more knowledge is needed of the basic mechanisms of the spread of drug use and problem drug use or addiction. As long as no general consensus exists among scientists on what risk factors are the most influential, from the wide range of social, psychosocial and biological factors identified in different studies (Rhodes et al., 1999), no good interventions can be expected. This lack of common understanding of a basic mechanism for the epidemiology of drug use also makes it difficult to reach consensus on results from modelling studies, contrary to more established fields as, for example, infectious-diseases epidemiology or the economy. Models of spread are often based on weak assumptions but can help to suggest the underlying processes. For instance, there is evidence that drug use often spreads as an infectious disease, in other words that the rate of new cases depends on the number of existing cases and the number of susceptible cases (Mackintosh and Stewart, 1979; De Alarcón, 1969). This may be understood in terms of each new drug user having a constant rate of initiating others. However, the role of supply factors (price of drugs, availability) and, unlike in the case of infectious diseases, the intentional marketing activities of drug dealers, are not well understood. Also, infectiousness (the probability of 'infecting' another person) may not be as constant as in a biological context and may depend on a range of unknown social factors. Unlike an infectious agent, the spread of drug use may be influenced by mass media, for example through developments in the music and fashion industries which may have contributed to the recent spread of ecstasy and other new synthetic drugs.

Third, more work is still needed to clarify the natural history of progression to problem drug use. It may be less important to study what leads young people to experiment with illegal drugs, as this mechanism seems no different from experimenting with alcohol or tobacco at earlier ages (Kandel and Faust, 1975; Bauman and Phongsavan, 1999). It is, however, crucial to understand why some young people progress into frequent use of heroin while others never go beyond experimenting with cannabis. It may well be that this is mainly due to social factors, such as 'meeting the wrong friends'. It may also be that some people are genetically prone to addiction (Nestler, 2000) or continue using in order to 'self-medicate' negative moods and mental problems (Leshner, 1999). An infectious-disease type of modelling might again help in predicting the spread of heavy use and, for example, in estimating the cost-effectiveness of interventions. Concepts could then be used from infectious-diseases epidemiology, such as 'infecteds', 'susceptibles', 'basic reproductive rate', 'herd immunity'. Key issues would be to distinguish 'susceptibles' from 'immunes' or to find a type of 'vaccination' that could turn 'susceptibles' into 'immunes', for example by detecting and treating depressive symptoms in young adults. If the epidemiology of illegal drug use would more resemble chronic-diseases epidemiology, models would focus on demographic or social developments that would lead to higher or lower prevalence/incidence of the predisposing genetic or mental-health factors in the general population. It would then

be important to identify those with a genetic or mental risk profile at an early stage and to find ways to protect these persons from addiction to any substance or behaviour (including, e.g., gambling) later on in life.

Other problems in the study of drug use may be less purely scientific or methodological but rather illustrate the complex interface between science and policy. Possibly the clearest example is the illegal status of drug use. There are few epidemiological areas in which disease is viewed as illegal. This legal and often moral aspect complicates public-health-oriented work on drug problems in different ways.

First, methodologically sound studies on interventions are very difficult to set up, not only due to the difficult type and illegal habit of patients, but also because some of the potential interventions themselves are illegal. This complicates collecting hard evidence on policy options and the debate often continues in a vicious circle. This is obvious regarding a controversial intervention such as controlled heroin provision (Van den Brink et al., 1999; Uchtenhagen et al., 1999), but even substitution treatment, which has been well accepted in the scientific field (Des Jarlais and Hubbard, 1999), cannot always be provided easily everywhere. Patients are often treated within law-enforcement environments (prisons), which do not always allow a confidential doctor–patient relationship. Patients may also be in and out of treatment and prisons, continuously interrupting therapy, making positive effects difficult to sustain and thereby increasing the risk of adverse consequences such as HIV infection or overdose (Metzger et al., 1998; Turnbull, 1997; Wiessing et al., 1996; van Ameijden et al., 1999). Moreover, the legal restrictions and stigma surrounding drug use in most countries lead to difficulties in obtaining reliable routine data or indicators (Hughes and Rieche, 1995). Drug users will try to keep away from any registries as much as possible, and even doctors might give a drug-related death a less stigmatised code on the death certificate in order to protect the relatives of the deceased. Legal restrictions might even affect scientific work in a more general manner, by influencing funding priorities or pressuring scientists to select 'politically correct' results.

Second, the legal and moral perspectives often lead to much confusion about case definition in policy debates. Rather than talking about 'illegal drug use' in general, as is often done, it is important clearly to distinguish problem drug use from non-problem drug use. Even if in many countries (usually non-problem) cannabis use does not differ from (often problem) heroin use in legal terms, public-health consequences are vastly different. Many forms of drug use, even if illegal, do not lead to more problems for the individual or society than moderate smoking or alcohol use (Nicholson et al., 1999; Marwick, 1999). It is problem drug use, such as injecting or frequent use of heroin, crack and other hard drugs, that incurs most costs to society in the form of infections, deaths, addiction and crime (EMCDDA, 2000; Amaro, 1999). A policy-maker in general tries to minimise the social costs of drug use and these should therefore be directly targeted by interventions. Drug policies that do not distinguish between different modes of drug use may not be efficient and may even be counter-productive. Differentiating between lighter and heavier forms of drug use in policy objectives, using case definitions based on clear diagnostic rather than legal criteria, could greatly help progress in the drugs field.

Third, the lack of clarity in the definition of the problem and the lack of (possibilities to gather) scientific evidence have resulted in a lack of common objectives to reduce drug problems. While no one discusses whether HIV infection or cancer should be prevented or not, this consensus does not exist with regard to drug use. It is important to reach evidence-based agreement on general objectives and on an optimal balance of intervention priorities to reach these objectives (Hartnoll, 1990). Some scientists believe that reducing all drug use (if at all possible by law enforcement) will lead to less problem use as well, while others think that this is not possible and that law enforcement actually increases drug problems, having no or only small effects on the prevalence of non-problem use (Uhl, 2000). Both approaches, however, imply totally different interventions, legal and preventive in one case and public-health-oriented in the other. It is interesting to note that usually little effect is found of changes in law enforcement on levels of problem drug use (MacCoun and Reuter, 1997; Lenke et al., 1996), and that countries with totally different policies may show very similar prevalence (EMCDDA, 2000). It may therefore be more useful to aim at preventing problem drug use and its secondary consequences, such as overdose and HIV transmission (Stimson, 1996; Wodak, 1995), as do most policies regarding alcohol abuse, rather than to attempt to eliminate all forms of drug use altogether.

Where are we?

As discussed above, five key indicators are currently being implemented in the European Union and others are still in development, but data quality and comparability issues have not yet been solved and data collection has only recently started. Many models of drug use exist but consensus on a basic mechanism to monitor the spread and progression of problem drug use is still lacking, while legal issues around drug use severely complicate the data collection and monitoring of problem drug use, and the development and evaluation of effective interventions.

Improving the quality of problem drug-use monitoring is of great importance. While non-problem use can be followed through general population and school surveys, prevalence estimation of problem use with modelling approaches is still in its infancy. Prevalence and incidence are fundamental measures of the spread of problem drug use and the effects of policy interventions. While prevalence among sub-groups can provide information on exit rates from the drug-user population (e.g. through treatment), detailed information on incidence is important for assessing entry rates and evaluating preventive measures.

At present, data quality does not permit the precise measurement of prevalence and incidence in most countries, even if important methodological advances are being made. It is necessary to obtain more commitment for quality data collection for such statistical estimation modelling as well as for developing dynamic models to investigate policy choices. Legal barriers (data-protection laws) have to be removed to facilitate scientific work on all potentially useful interventions, but also to improve data availability.

More multi-disciplinary work is necessary to clarify and describe the spread of problem drug use in space and time. Economic models need to be further developed to better estimate the costs to society of different consequences of drug use and to estimate the cost-effectiveness of policy options. The role of supply factors (drug-markets research) and how these interact with demand in general, but especially with progression to problem drug use, need to be clarified. Understanding of the spread of drug-related infectious diseases is more advanced than other areas of drug use. While practical work is needed on HIV prevention (such as increasing the coverage of interventions) prevention of hepatitis C infection still needs to be improved by better understanding the transmission among injecting drug users. As effectiveness of interventions cannot easily be assessed at macro level, the monitoring of disease should at least include monitoring coverage of the target population by those interventions (Habicht et al., 1999).

In conclusion, indicators and models form important and complementary tools for the epidemiology of drug use at macro level. Although quality of data and thus possible inferences are much weaker than in local studies, macro-level assessments are indispensable for guiding policy decisions. However, interpretation of results at macro level should be even more careful than in general, which conflicts with the fact that policy decisions have to be taken. There is still a range of problems specific to the field of drugs that needs to be solved before effective measures can be widely adopted.

Acknowledgements

This chapter benefited from the EU-funded European network to develop policy-relevant models and socio-economic analyses of drug use, consequences and interventions (TSER/DGXII project ERB 4141 PL980030). The authors are grateful for the critical and constructive comments of Dr Alfred Uhl.

References

Amaro, H. (1999). 'An expensive policy: the impact of inadequate funding for substance abuse treatment'. (Editorial). *American Journal of Public Health* 89: 657–9.

Anderson, R. M., May, R. M. (1992). *Infectious diseases of humans – dynamics and control*. Oxford: Oxford University Press.

Bauman, A., Phongsavan, P. (1999). 'Epidemiology of substance abuse in adolescence: prevalence, trends and policy implications'. *Drug and Alcohol Dependence* 55: 187–207.

Blower, S. M., Hartel, D., Dowlatabadi, H., Anderson, R. M., May, R. M. (1991). 'Drugs, sex and HIV: a mathematical model for New York City'. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 331 (1260): 171–87.

Daar, E. S., Moudgil, T., Meyer, R. D., Ho, D. D. (1991). 'Transient high levels of viremia in patients with primary human immunodeficiency virus type 1 infection'. *The New England Journal of Medicine* 324 (14): 961–4.

De Alarcón, R. (1969). 'The spread of heroin abuse in a community'. *Bulletin on Narcotics* 21: 17–22.

Des Jarlais, D. C., Hubbard, R. (1999). 'Treatment for drug dependence'. *Proceedings of the Association of American Physicians* 111: 126–30.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (2000). *Annual report on the state of the drugs problem in the European Union*. Lisbon: European Monitoring Centre for Drugs and Drug Addiction.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (1999). 'Study to obtain comparable national estimates of problem drug use prevalence for all EU Member States'. (EMCDDA report CT.97.EP.04). Lisbon: European Monitoring Centre for Drugs and Drug Addiction.

Frischer, M., Leyland, A., Cormack, R., Goldberg, D. J., Bloor, M., Green, S. T., Taylor, A., Covell, R., McKeganey, N., Platt, S. (1993). 'Estimating the population prevalence of injecting drug use and infection with human immunodeficiency virus among injecting drug users in Glasgow, Scotland'. *American Journal of Epidemiology* 138 (3): 170–81.

Habicht, J. P., Victora, C. G., Vaughan, J. P. (1999). 'Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact'. *International Journal of Epidemiology* 28: 10–18.

Hartnoll, R. (1991). 'Epidemiological approaches to drug misuse in Britain'. *Journal of Addictive Diseases* 11 (1): 33–45.

Hartnoll, R. (1990). 'A framework for evaluating activities to reduce the demand for drugs in Europe'. *Gaceta Sanitaria* 4 (21): 239–51.

Hartnoll, R., Avico, U., Ingold, F. R., Lange, K., Lenke, L., O'Hare, A., de Roij-Motshagen, A. (1989). 'A multi-city study of drug misuse in Europe'. *Bulletin on Narcotics* 41: 1–2, 3–27.

Hughes, P. H., Rieche, O. (1995). 'Heroin epidemics revisited'. *Epidemiologic Reviews* 17: 66–73.

Kandel, D., Faust, R. (1975). 'Sequence and stages in patterns of adolescent drug use'. *Archives of General Psychiatry* 32: 923–32.

Kaplan, E. H., O'Keefe, E. (1989). 'Needles that kill: modelling HIV transmission via shared drug injection equipment in shooting galleries'. *Review of Infectious Diseases* 11: 289–98.

Kretzschmar, M., Wiessing, L. G. (1998). 'Modelling the spread of HIV in social networks of injecting drug users'. *AIDS* 12 (7): 801–11.

Kretzschmar, M., van Duynhoven, Y. T., Severijnen, A. J. (1996). 'Modelling prevention strategies for gonorrhea and chlamydia using stochastic network simulations'. *American Journal of Epidemiology* 144 (3): 306–17

Lenke, L., Olsson, O. (1996). 'Sweden: zero tolerance wins the argument?', in Dorn, N., Jepsen, J., Savona, E. (eds) *European drug policies and enforcement*. London: Macmillan Press.

Leshner, A. I. (1999). 'Science-based views of drug addiction and its treatment'. *Journal of the American Medical Association* 282: 1314–16.

Longini, I. M. Jr, Clark, W. S., Byers, R. H., Ward, J. W., Darrow, W. W., Lemp, G. F., Hethcote, H. W. (1989). 'Statistical analysis of the stages of HIV infection using a Markov model'. *Statistics in Medicine* 8 (7): 831–43.

MacCoun, R., Reuter, P. (1997). 'Interpreting Dutch cannabis policy: reasoning by analogy in the legalization debate'. *Science* 278: 47–52.

Mackintosh, D. R., Stewart, G. T. (1979). 'A mathematical model of a heroin epidemic: implications for control policies'. *Journal of Epidemiology and Community Health* 33: 299–304.

Marwick, C. (1999). 'Illicit drug users not idle; report says 70 % go to work'. *Journal of the American Medical Association* 282: 1320–1.

Metzger, D. S., Navaline, H., Woody, G. E. (1998). 'Drug abuse treatment as AIDS prevention'. *Public Health Reports* 113 (suppl. 1): 97–106.

Nestler, E. J. (2000). 'Genes and addiction'. *Nature Genetics* 26: 277–81.

Nicholson, T., White, J., Duncan, D. F. (1999). 'A survey of adult recreational drug use via the world wide web: the Drugnet study'. *Journal of Psychoactive Drugs* 31: 415–22.

Pasqualucci, C., Ravà, L., Rossi, C., Schinaia, G. (1998). 'Estimating the size of the HIV/AIDS epidemic: complementary use of the Empirical Bayesian Back-Calculation and the Mover–Stayer model for gathering the largest amount of information'. *Simulation* 71: 213–27.

Rhodes, T., Lilly, R., Fernández, C., Giorgino, E., Kemmesis, U. E., Ossebaard, H. C., Lalam, N., Faasen, I., Spannow, K. E. (1999). 'Risk factors associated with drug use: the importance of "risk environment"'. (EMCDDA report CT.97.EP.12). Lisbon: European Monitoring Centre for Drugs and Drug Addiction.

Rossi, C. (1999). 'Monitoring drug control strategies: hidden phenomena, observable events, observable times'. *International Journal of Drug Policy* 10: 131–44.

Stimson, G. V. (1996). 'Has the United Kingdom averted an epidemic of HIV-1 infection among drug injectors?' *Addiction* 91 (8): 1085–8.

Turnbull, P. (1997). 'Prisons. Vancouver conference review'. *AIDS Care* 9: 92–5.

Uchtenhagen, A., Dobler-Mikola, A., Steffen, T., Gutzwiller, F., Blättler, R., Pfeifer, S. (1999). *Prescription of narcotics for heroin addicts. Main results of the Swiss National Cohort Study*. Basel: Karger.

Uhl, A. (2000). 'Non-problematic use – problematic use: a paradox', in Springer, A., Uhl, A. (eds) *Illicit drugs: patterns of use – patterns of response*. Proceedings of the 10th Annual European Society on Social Drug Research (ESSD) Conference on drug use and drug policy in Europe. Innsbruck: Studienverlag.

Uhl, A., Seidler, D. (2000). *Prevalence estimate of problematic opiate consumption in Austria*. Vienna: LBISucht.

van Ameijden, E. J., Langendam, M. W., Coutinho, R. A. (1999). 'Dose-effect relationship between overdose mortality and prescribed methadone dosage in low-threshold maintenance programs'. *Addictive Behaviors* 24: 559–63.

van den Brink, W., Hendriks, V. M., van Ree, J. M. (1999). 'Medical co-prescription of heroin to chronic, treatment-resistant methadone patients in the Netherlands'. *Journal of Drug Issues* 29: 587–608.

Wiessing, L. G., Calvan, M. G., Ravà, L., van Ameijden, E., Buster, M., Hickman, M., Ribeiro, J., Rossi, C. (submitted). 'Analysis of the latency period from onset of heroin use to first treatment; comparisons among four European sites, possible implications for treatment services and for understanding the natural history of problematic heroin use'.

Wiessing, L. G., Toet, J., Houweling, H., Koedijk, P. M., van den Akker, R., Sprenger, M. J. W. (1996). 'Prevalence and risk factors for HIV infection among drug users in Rotterdam, the Netherlands'. XIth International Conference on Aids, Vancouver, 7–11 July 1996. Abstract Tu.C. 2548.

Wodak, A. (1995). 'Harm reduction: Australia as a case study'. *Bulletin of the New York Academy of Medicine* 72: 339–47.

World Health Organization (WHO) (1999). *Teaching health statistics*. Geneva: WHO.

CHAPTER 2

WHAT ARE DYNAMIC MODELS AND HOW CAN THEY BE USED?

Christine Godfrey, Steve Parrott, Matthew Sutton, Vanessa Waby, Philip Young with Catherine Comiskey, Martin Frischer, Gordon Hay, Matthew Hickman, Johannes Jager, James Kahan, Ludwig Kraus, Roberto Mollica, Maarten Postma, Gianpaolo Scalia Tomba, Roland Simon, Colin Taylor and Lucas Wiessing

Drug use and problems vary across the countries of Europe, within each country and over time. Through the work of the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) and other European bodies, an increasing amount of data (both qualitative and quantitative) is being collated on drug use and problems. Knowledge of the current situation is important for policy-makers, but there is also a need to plan for the future. What sense can be made of the data available? Can trends in use or problems be predicted? Can simulations be made of different policy instruments to examine their potential consequences?

Researchers use a number of tools to investigate these questions. A model is defined as a simplified description of a system or phenomena in order to assist in calculations and predictions. As may be expected, however, researchers from different areas of study have developed different models and used different statistical and mathematical techniques in order to attempt to put some of their models into practice. Which approach yields the most useful result? While some models may seem a gross simplification of a complex problem, others are so intricate and technical, denying the general reader's ability to understand the logic used, that the model becomes a 'black box'.

During 1997, a group of experts drawn from across Europe met twice in York to consider the feasibility of developing dynamic models of drug use and related problems. Review papers were commissioned from this expert group and the extended and revised reviews form the main contents of this monograph (see full report, Godfrey et al., 1997). A further seminar, which included the original group of modelling experts, drug researchers and other drug specialists, was held in Lisbon in 1998. The aim of this seminar was to explore which policy questions could be addressed by which type of models and how the results of such modelling could be disseminated to policy-makers across Europe (Godfrey et al., 1998).

This chapter contains a synthesis of the discussions and conclusions from these three different meetings. The potential policy questions are considered in the first section. The various types of models and a summary of their advantages and disadvantages is

given in the second section. How modelling techniques could be used to address policy questions and the potential for further research is explored in the third section. The chapter concludes with a discussion on the role of dynamic modelling.

What are the policy questions?

Drug issues span a vast array of policy areas at different levels and are of enormous international concern with some policies being addressed at this level. There are also national, local and community issues, which often cross cut traditional policy divisions. Policy-makers include not just politicians, but also bureaucrats. Developing relationships between researchers and the large number of different decision-makers may be problematic, but this does not necessarily imply that research cannot be fed into the policy-making process.

The questions posed by policy-makers of all sorts tend to be easy to comprehend and beyond dispute. They centre on issues such as: how many people use drugs? How much do they consume? What are the consequences of this use? Which policies work? What are the costs of these policies? These questions can obviously be extended and one important role of modelling is to outline the dimensions of drug use and problems and their interactions. Part of the researcher's role is to demonstrate the impact of drug use and problems not just on physical consequences such as disease rates and health-care systems, but also on social impacts. This extends to the effects on the community, the economy and the judicial system.

Many policy concerns are of a comparable nature. Why is drug use in the policy-maker's area higher or lower than in another? As more national and regional data become available policy-makers may be tempted to make invalid and inappropriate conclusions about the efficacy of their policies. Modelling data by formulating and testing hypotheses is one means of informing the policy debate. For example, research may be able to help explain why some countries have lower prevalence of cannabis use than others. Further collection of data, particularly over time, raises other policy-driven questions, particularly why do use and patterns of drug use change? Why does some drug use become problematic? Can the results of policy changes be observed?

While these questions are simple, the answers have proved difficult to find. Partly, research has been hampered by the lack of data. Modelling techniques have an important part to play in enabling researchers to make the best use of the limited available data. Bodies such as the EMCDDA, which are playing a major role in collating comparable data, are clearly important. Partly, there is a lack of research and researchers, with many modellers working in isolation and without major funding. However, these are not the only issues.

One of the recurring themes of the discussions in this project was the link between researchers and policy-makers. In seeking funding researchers have been encouraged to stress the practical relevance of their work. However, the motivation and interests of researchers in a complex area may not be the same as those of the policy-maker. Researchers can become very frustrated if clear policy recommendations arising from research are not enacted. In contrast, policy-makers can become disillusioned if model

predictions prove false. Examples of the inaccuracy of the initial HIV/AIDS models in predicting the spread of HIV and consequent health-care resources from data were frequently quoted in discussions. Similarly, decision-makers can be discouraged from considering research if the model is not transparent and is masked in technicalities. There can be a long time lag between the policy-maker posing a question and researchers being able to provide an answer. Disputes between researchers can also alarm policy-makers but be one of the main driving forces behind the development of research techniques and results.

Continuing dialogue between researchers and policy-makers is clearly required. There is considerable variation across countries in how research is used in the policy-making process. In some countries there are bodies which act to synthesise research and bring researchers and decision-makers together. However, research also has to be timely and politically acceptable. The impact of research varies considerably, as was demonstrated by some of the case-study material used in the project. The first important aspect is the availability of data about policies. In the 1997 York seminar, Pierre Kopp presented analysis about the expenditure of the French government on different drug policies (Godfrey et al., 1998). These data are not readily available and as such they provide some background against which policy statements can be checked, for example, the balance between enforcement and treatment expenditures. Peter Reuter, at the same meeting, provided two examples of drug modelling work with different impacts on policy-making (Godfrey et al., 1998). In the first example, simple models were constructed to demonstrate the impact of various ways of restricting the supply of drugs. The policy background was a move to 'seal the borders' by using a military solution. This policy change was not fully supported by the military authorities and the model was able to demonstrate how the costs of smuggling varied by routes and the impact of alternative interdiction policies. The study, despite its limitations, did have an impact and this was also helped by general concern about the amount of expenditure on interdiction by the United States Congress.

The second example, quoted by Reuter, involved the work of Rydell and Everingham (1994). This was a simulation model of the cost-effectiveness of treatment compared to enforcement policies to reduce cocaine consumption. This work has been favourably reviewed by other researchers and frequently quoted. It failed, however, to have any political impact, with no change in the proportions spent on enforcement compared to treatment despite the clear conclusions that treatment was more cost-effective. There can only be speculation about why this study failed to have a major policy impact. It could be simply that the budgets for interdiction are set centrally while those for treatment are set locally and that there was therefore no clear policy decision-making body to act on the results. It may be that the research was not timely because the US federal budget at that time was not the focus of a strong political debate. It may be that such research findings take a lot longer to be acted upon with the need for more fundamental shifts of opinion as the research results are disseminated more broadly.

To summarise, researchers can contribute to the debate about drug use, problems and the effectiveness of alternative policy instruments. Modelling, with available data, is an important part of this research contribution. Policy questions and concerns are relatively

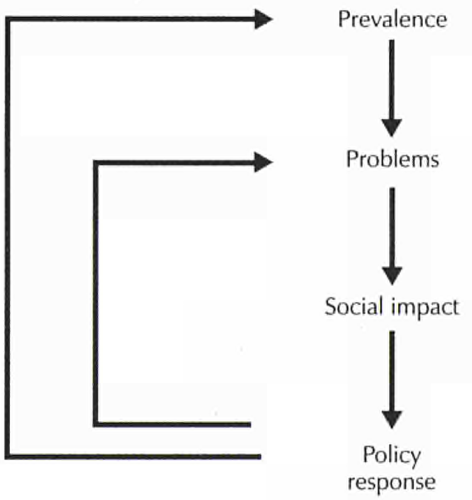
straightforward. However, drug problems themselves are complex and the lack of available data means that answers to these questions are not necessarily readily available. The next step is to explore the modelling techniques available and how they can be used to address policy questions.

Types of model

From the policy perspective, any model has to help illustrate the link between the prevalence of drug use, the problems created by drug use, the social impact of these problems and the policy response (see Figure 1). There are direct links between prevalence, problems and social impact which can influence policy responses. However, policies can be directed at the problems or at changing prevalence. Models of various sorts provide one approach to exploring the inter-relationships between different factors and explain some of the observed variations in drug use and problems over time and between countries. Further, the effects of policy measures can be explored using the results of such models within a 'what if' framework. In a dynamic framework the relationship between the prevalence and problems of drug use and policy impact may need to be traced through individuals' drug careers. At population level there is a need to examine the impact in specific time periods of the total of these individual effects. Different levels of geographical coverage of the models may be required to match the policy-making bodies across Europe.

The major aim of the project which forms the basis of this monograph was to explore the feasibility of using dynamic models as a means of interpreting data on drug use and problems across Europe. It was recognised that the tasks would draw on research from many academic disciplines. This, coupled with the technical nature of the models,

Figure 1: Framework for models



could lead to difficulties in finding common, agreed and accessible definitions and language. A subsidiary aim of the project was to uncover some of the potential confusions caused by the same terms being used in different techniques and to start to define some of the remaining disputed terms.

The objectives of the project included an attempt to provide some inventory of the types of tools which may be used in models attempting to specify relationships between different factors impacting on the constructs set out in Figure 1. This inventory of models could not be restricted to those techniques already in use because of the current relative paucity of specific drug-related research. However, models have been increasingly used to address related areas, such as communicable diseases, and other substances, such as alcohol and tobacco.

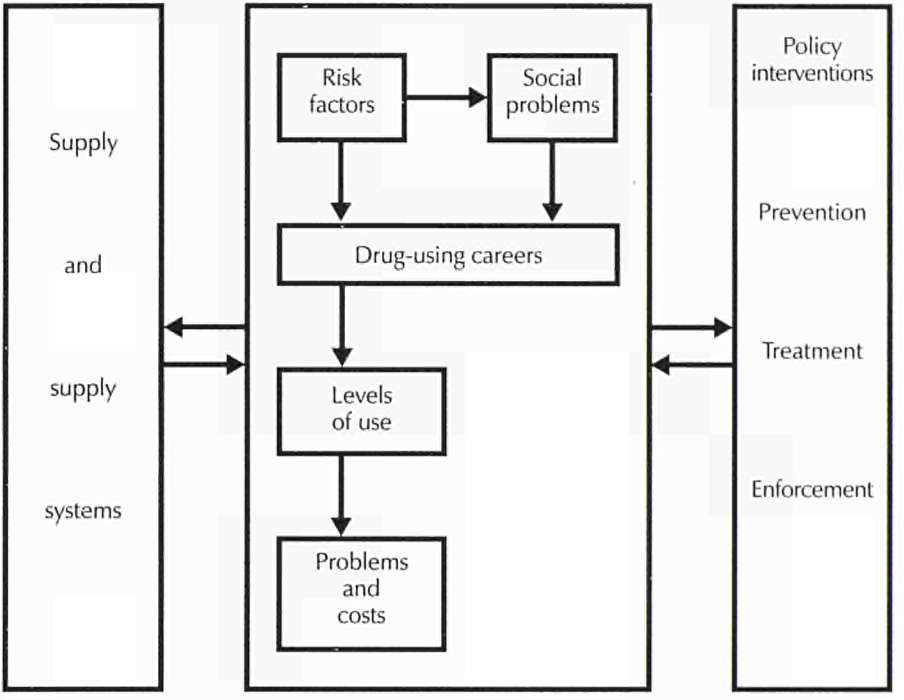
In discussion, it became clear that the term 'dynamic' could be interpreted in a wider sense than changes over time. Within Europe, the spread of drug use, drug supply and drug problems across geographical areas is of wide interest. To build models, a more fundamental understanding of the social processes which contribute to drug-using careers is also necessary. Many of the models have little behavioural foundation and it was felt that these aspects could not be neglected.

Two further areas were considered: data requirements and data availability for models; and policy use and acceptability. In considering dynamic models, one interest at European level would be to refine the use of existing national data. A second objective would be to examine what data are collected and whether these data could be improved or augmented. All models are built on a series of assumptions. The mathematics and statistics needed to construct the models can be difficult for non-specialists to follow. For policy acceptability, it may be important for the process of model construction to be visible and understood by a wide audience. This may conflict with criteria developed by researchers.

Considering the most feasible questions to be addressed, a model was proposed which extends and develops Figure 1. This schema, shown in Figure 2, has three main boxes: supply of drugs; policy interventions; and the main box describing drug use and problems. This main box illustrates the relationships which many models may seek to address. Obviously this schema does not explicitly cover different geographical areas, but it could be applied at different levels. An attempt was made to address different parts of this schema with a mixture of potential projects which are both feasible and of interest to policy-makers, as set out in the fourth section of this chapter.

In the next section, a summary of what is known based on the reviews contributed by the participants at the seminars and the potential for different techniques to address these inter-relationships is discussed before the priorities to fill the gaps in knowledge, which may be of special interest to decision-makers, are explored.

Figure 2: Schema for research and policy questions



Findings from review papers and case studies

In this section, a short summary of each review topic is presented based on the discussions at the second expert meeting. Where appropriate, references are made to the chapters below but the discussions at the expert meeting were based on earlier versions of these reviews. The discussion of each technique was framed around the following questions:

- What are the outputs of the model?
- To what extent are existing examples of these models applicable in Europe?
- What geographic areas do the models consider?
- What are the data demands of the model?
- Why have the models not been applied?
- What would be the added value of further research?

Geographic information systems

Geographic information can be seen in two parts. These geographic information system (GIS) models have been mainly used as a means of displaying spatially referenced data

(see Chapter 3). This can be important in itself in identifying routes of the spread of diseases or socially related behaviour such as drug use. However, such systems can also be augmented by linking the spatial framework with relational databases and epidemiological functions. This could be applied to examining, for example, deprivation and town size in relation to the spread of drug use or to drug-supply routes.

The technique was thought to have great potential at many levels and especially for the EMCDDA. The main advantage is the visibility of the output and its appeal to policy-makers. However, the process may need considerable initial investments and more research is needed to develop some of the underlying functions of the spread of drug use.

Dynamic models of data-production processes

Chapter 4 contains a number of illustrations of the application of different techniques based on available data in various countries and addressing a range of research questions. This chapter illustrates how modelling techniques may be successfully applied.

The advantage of the approach adopted in the original review paper was its pragmatism and the speed with which some results could be obtained. It had the advantage of making optimal use of available data to examine some important aspects of the impact of illicit drugs. However, one of the reasons the approach may not be widely adopted was that it was seen as over-simplistic. Collating data across Europe if comparability is required can also be very time-consuming.

The discussion also highlighted the importance of not using one single methodology or set of data to model. Rather, there was a need to apply several approaches to reduce the uncertainties of modelling drug use and problems.

Multiple-indicator models

Multiple-indicator models use time series of indicators and statistical techniques to explore the links or correlations between the different indicators (see Chapter 5). These indicators are likely to be in three major groups: medical and treatment indicators; social indicators; and legal indicators. The relationships are also likely to be in the form of time lags. For example, the numbers in treatment in the current period may be linked to indicators of drug use several years earlier.

The main advantage of this technique is that it uses available data. However, it would require time series of the 'same' indicator. These data may exist, but experience suggests that even collating data can be a time-consuming and resource-intensive process. However, the data-monitoring systems of different groups across Europe are addressing this problem.

One disadvantage of the technique is that the correlations found do not imply causality. Changes in the social processes, which generate the indicators, are likely to invalidate the model predictions. Only by examining individual processes can some understanding be gained of the factors influencing the trends in different patterns of drug use and problems.

Back-calculation methodology

This methodology has been applied and developed mainly to provide estimates of the total numbers of the population infected with HIV (see Chapter 6). The technique involves using the known or identified population of those with a disease with some 'incubation rate' to back calculate the total number of infected persons. The purpose of this technique is to generate estimates of the 'hidden' population, which is clearly of policy interest as well as providing potential input into other modelling techniques.

The back-calculation method has been applied extensively throughout Europe to the problem of estimating the numbers infected with HIV. It may have applications in other diseases linked to drug use. The expert group also discussed ways of applying the technique to measuring the hidden population of drug users. The proposal was that it may be possible to use data on the numbers presenting for treatment for drug problems with data on the time delay between drug initiation and treatment to estimate the hidden pool of users. This may be appropriate for some type of drug use, for example opiate use, but not for others. However, not all opiate users will present for treatment and hence using the back-calculation method could be seen as providing a minimum estimate. This proportion attending for treatment may also vary with the supply of treatment over time. It was not clear whether different behavioural aspects could be incorporated easily within an essentially mechanistic model, although different statistical techniques could be used to simulate the importance of different assumptions. Despite these issues, this technique appeared promising for addressing some questions.

The mathematical and statistical techniques used to perform the back calculation vary. While such technicalities may be difficult to describe in full, the concept of the model can be described in relatively simple terms. There is a pool of expertise across Europe in applying this model, although it may need refining on a country-by-country basis. One disadvantage of the model is its sensitivity to small changes in the input data.

Compartmental modelling and stochastic dynamic systems

The basis of these types of models is the division of the population into two or more groups (see Chapter 7). The models then explore how individuals move from one compartment to another. In some circumstances, the rate at which individuals transfer through the compartments is known and in other models different probability rates can be tested. The technique provides a means of combining data from different sources in a systematic way to offer insights into drug use and related problems. An example was given of how such models may be used to explore how individuals move through a drug-using career. It also provides a framework in which different scenarios can be explored.

These types of models have been used across Europe in the study of the spread of HIV. They have also been used to estimate the spread of HIV among injecting drug users. One possible disadvantage is that estimates depend heavily on the parameters employed and the realism of the structure of the underlying compartmental model. The models

(depending on the complexity) are demanding of data, but building conceptual frameworks may help refine existing data-collection methods.

Discussion of the models highlighted the basic need to understand the social process affecting an individual's drug-using career or the process of initiating drug use. This points to the need for multidisciplinary approaches to the development of models.

System-dynamic models

This term caused some confusion as it referred to a group of models that contained some of the techniques used elsewhere. The distinguishing feature of a system-dynamic model is the attempt to examine the flows between different parts of a system (see Chapter 8). Generally, some sort of feedback mechanism is generated between the different parts of the system, rather than being a form of progression as in some of the compartmental models. However, parallels were noted between system-dynamic models and the compartmental and scenario models described in other chapters.

System-dynamic models have also been applied to more aggregated data than compartmental models. Examples are provided in the research generated by RAND (Rydell and Everingham, 1994), where such models have been used to simulate policy changes such as switches of expenditure between enforcement and treatment. More aggregated models frequently use a combination of methods and techniques to generate data for their simulations. More individually based models may have applications in examining problem drug use, treatment take-up and consequences through time.

Discussion suggests that these types of model have the potential to address very important policy questions. However, some models are very complex and may not be sustainable given current data availability. The more complex the model, the more useful it could be. The critical question remains, however, how to pre-judge the benefits associated with extending the complexity of the model of each stage.

Structural equations and path analysis

The aim of these techniques is to examine the relationships between known variables to examine the underlying (latent) concepts (see Chapter 9). These are generally exploratory tools rather than ones leading to predictions, although they have been applied to explore aspects such as the concepts that may lead individuals through different pathways of drug use. The models generally use individually based data.

One advantage of the techniques discussed by the group was the ability to explore data to find previously unknown relationships. This may be a useful technique to apply to survey data. Unobservable constructs are used widely in some disciplines such as psychology, but may not be so acceptable to a broader audience. It is perhaps the 'black box' nature of the models which limits their acceptability to policy-makers. Such models may be of much wider use to the research community.

Modelling infectious diseases and other health consequences of drug use

This is clearly an important area of application of different modelling techniques (see Chapter 10). These models have the primary aim of examining the factors that influence the spread of different health consequences and predicting that spread. There is a body of work on the spread of HIV and the issue is how the models can be extended to other diseases. The expert group expressed particular concern about the spread of the hepatitis C virus (HCV) among the drug-using population.

The objective of the different models is to understand the quantitative relationships between the characteristics of a specific disease, the contact patterns in the population within which it is transmitted and the resulting incidence and prevalence. The models can also be used to simulate the effects of prevention programmes and provide data for modelling cost consequences.

In some of the examples presented, behavioural aspects, such as the dynamics of social networks, were combined with some of the more mechanistic models of disease spread. These behavioural models are more difficult to develop and require additional data, especially on drug-using careers and behaviours. Because of the differences between areas these models are generally applied at local level.

There may be advantages in starting exploratory work with simple models. This may be seen as a priority given the consequences to individuals of the spread of these diseases.

Scenario analysis

This was another term which caused some confusion. Chapter 11 focuses on one specific type of analysis and aims to set out a conceptual model for exploring the health impact of drug use at multinational level. As the main model is conceptual rather than mathematical, it is necessary to build conceptual boxes and draw out the main relationships between the boxes. In the model for the European Commission Concerted action on multinational AIDS scenarios, for example, the main 'boxes' were: policy; demography; determinants of behaviour; health status; health-care system and associated costs; and autonomous developments such as medical technology, macro-economic factors and socio-cultural factors. Some links can be seen between this area and compartmental models and system-dynamic models. The next step is collecting data and examining all the impact factors. Once the data have been collated this usually forces some simplification of the conceptual model. Different techniques are then combined to 'solve' the model and its consequences. The next step is then to set a reference scenario given certain conditions and simulate changes from the reference scenario by altering some of the model's policy parameters.

The advantage of this type of scenario analysis is that it explores the links between different policy developments and consequences. Many other techniques are used in constructing the models. It does, however, demand considerable resources. It could be seen as a useful framework within which to develop more detailed research work. As such it may provide a link between the models exploring the factors influencing drug use and those tracing consequences.

There are other more qualitative approaches to scenario analysis and these involve a different combination of approaches. Policy-makers may, for example, be asked about which states they would most desire in the future and what they would most wish to avoid. The modellers then attempt to trace back from these outcomes the processes which affect them. This allows insights into what steps have to be taken in the present to get to some desired outcome in future. This type of approach may well have applications in the drugs field and could provide useful insights into the policy-making process not covered by other more mathematical and statistically based models.

Modelling cost consequences

This is a form of economic modelling which takes the results of many other models to consider the resource impact of different patterns of drug use and their consequences (see Chapter 12). Its advantage is the policy interest. However, to apply this approach across Europe, some standardisation of data-collection methods is required and this may be very resource intensive. The benefits of standardisation are that comparisons can be made across countries on a like-for-like basis.

The basis of most health-care-cost consequences models is the use of an appropriate epidemiological sub-model, a model for the health-care needs and then one for estimating the costs. The details of the model required vary with the disease and the treatment needs at different stages of development of the disease. The reliability of the results of the model depends on the quality and representativeness of the data used for estimation. Health-care-cost data across Europe are of varied quality. Several possible developments were considered by the expert group, including examining the cost consequences of the hepatitis B virus (HBV) and HCV.

While health consequences were seen as a starting point and would link with other modelling work, there are a number of additional consequences of particular policy concern. The expert group considered some of the other consequences that may need modelling, especially crime costs.

Economic-market models

Economic-market models examine the interaction between the behaviour of those supplying drugs and those consuming drugs (see Chapter 13). The purpose is to generate estimates of the determinants of dynamic patterns of drug prevalence or incidence. They can be applied both at the population (macro) level and at the individual (micro) level.

Population-level models can be applied using some of the data currently available across Europe. Such models are generally applied to time series of data across the same geographic area. However, developments using individual-level data illustrate some uses of this technique to demonstrate the potential importance of economic factors in influencing drug-using behaviour and its associated consequences.

Some studies have been undertaken across Europe to examine the economic determinants of drug use, but few have been conducted on the factors influencing the supply of drugs. Because aggregate-level data are already collected, it was thought by the experts that well-founded economic models had the potential directly to inform policies, for example by identifying where to concentrate enforcement activities. At individual level, the disadvantages of these models were the demands on data and expertise. As within other areas, progress may be made if a multi-disciplinary approach were to be adopted.

Chaos models

An additional paper was presented at the second expert meeting. This was not an identified area for review, but rather an innovative area of research which was reported by a Spanish team of researchers. The potential for chaos models is to explore the spread of drug use and simulate within the model how different factors may alter the spread of drug use and its consequences. One disadvantage of the technique is the sensitivity to initial starting conditions. Another potential problem is explaining the model to policy-makers. However, the proof of the model will come with its application, as is being pursued within the Spanish group. There is scope for testing the basic model in other areas of Europe, although this is likely to be seen as a specific research task needing national research funding rather than being suitable for more general application at European level.

Where next? Research priorities

The reviews conducted for this project yielded a long list of potential research projects which are detailed in the following chapters of this monograph. Some areas are more developed than others and new work could be undertaken relatively quickly. Rather than prioritising research needs from the review papers, however, the expert group considered the potential questions that could be addressed and matched techniques to these questions. Six questions were identified by the group. These questions were seen as being of interest in themselves, and as fitting into different areas of the schema set out in Figure 2, above:

- What are the levels of drug use and problems? This could be considered over time, area and drug type. Geographic information systems may be useful for filling in gaps in information across Europe. Applications of back calculations, compartmental models and system-dynamic models could also be appropriate.
- How can the spread of new drugs be predicted? This was a potential area for conceptual scenario analysis using a variety of methods to understand social phenomena.
- What is the current drug-using career? Changes in drug-using behaviour and the interaction with treatment were seen as areas to which compartmental modelling could be applied. With data it may be possible to model gateways into drug use.

- How can social processes be modelled? These processes were seen as vital components of various modelling techniques. This is an area where there were gaps in techniques within the expert group. There was potential for considering different quantitative and qualitative data on risk factors, social networks and processes. As well as epidemiological, sociological and economic models, chaos theory may have some potential here.
- What are the current costs of drug use and the cost-effectiveness of policies designed to reduce these costs? The need to model the cost consequences of HCV and HBV was seen as a high priority.
- How can the impact of interventions be modelled? Again a variety of approaches could be used including compartmental modelling, system-dynamic models, cost and cost-effectiveness models and scenario analysis.

These questions fit different boxes or combinations of boxes set out in Figure 2, above. A multi-disciplinary team is required to develop many of the models and to begin to answer these questions. There may also be added value in developing similar approaches across a number of European countries. Pursuing modelling in each of these areas would provide inputs into more comprehensive models of the whole supply, demand and policy system.

The questions were considered further with additional experts in the seminar held in 1998 in Lisbon. A summary of the group discussion is given below. The aim of each workshop was to identify practical and feasible ideas for new research projects. In some areas projects were already under way and more work is being developed.

Modelling the health consequences of drug use – HBV/HCV/HIV

Of immediate concern was understanding the spread of these diseases. It was also seen as important to compare interventions in order to predict the future consequences and estimate the costs. There was a role for using infection rates as an indicator. A further area of work was to improve the quality of the treatment data on HIV/hepatitis.

Breaking this down into its component parts, the first issue is to estimate or model the size of the problem in terms of prevalence, consequences and risk in the general population. On interventions it was necessary to consider the effect of reporting delay, the potential to reach new injecting drug users and the issue of whether such interventions should be targeted towards risky behaviours. There is also a considerable dynamic component and a need to model the incubation period and the effects on different birth cohorts.

The data needed include the incubation time, the risk of infection, and the interaction of hepatitis with HIV. There are also requirements for data on treatment options, the proportion offered treatment, rates of success, the interaction of interferon and opiate use, and the different stages of the symptomatic phase of the disease. There are data from HIV cohorts and from behaviour surveys. Comparisons could be made between countries, but there are doubts about the reliability of some data, for example, on the proportion sharing needles.

Some modelling work has been undertaken on, and there is potential for considerable areas of new research into, preventing the spread of HIV from drug users to the general population. Further data could be input into models of transmission to the general population, partner studies and vertical transmission from mother to child. It was also necessary to build in interventions which acted on behaviour both before and after transmission. One option was to work with local prevalence projects and look at mortality, treatment, the import and export from the locality, and the hidden population that could be reached through outreach work. Costs could be considered but only after clarifying the dynamics of the disease. More specifically there were studies that could be undertaken on hepatitis B and the impact of vaccination and better treatment as well as the potential effects for the general population. Finally, there is a need to consider the routine data collected at European level and there may be some support for this type of research from former Directorate-General V, now the Health and Consumer Protection Directorate-General.

The priorities identified by the seminar were:

- modelling the spread of HIV;
- modelling interventions.

Geographically based models

Geographically based models like GIS are capable of capturing, analysing and displaying spatially referenced data. These models have been facilitated by the availability of digital map data which provide a spatial framework in which to attach the attributable data. They can be used to study the associations between location, environment and disease. Drug use is a dynamic process and spreads through a population both through micro diffusion (the spread of drug use among individuals within groups) and macro diffusion (where drug use spreads across geographic boundaries). The major current impediment is the lack of empirical data although both GIS maps and relational databases are becoming more readily available. The other data need is for the epidemiological functions, which would be part of wider GIS models.

The main outputs from GIS models are spatial-temporal maps. The models can, with dynamic models of forecasting, provide a means of visualising trends in drug diffusion and provide information for identifying populations at risk. For example, the Community epidemiology working group in the United States has provided simple maps for 25 cities displaying data, but has not in this case linked the maps and data to epidemiological functions which could be used for prediction.

To proceed there is a need to decide at which geographic level to start the analysis: city; region; or country. It was considered best to start at city level and then to extend the work. The requirements are for as much information about the city and the socio-economic and demographic data as possible. Finding epidemiological functions for these systems needs to be developed.

The group discussed the work already being undertaken in Spain, Italy, the Netherlands and the UK. In most cases these were mapping exercises and therefore the priority questions were the feasibility of incorporating models into GIS systems. The attraction of the maps for policy-makers and their important role especially in studying the spread of drug use and problems was recognised.

There was discussion about the best way to proceed with the EMCDDA on the work. Should researchers in individual countries work separately and then link up? Or should a common format be followed? It was considered important, at least, for a common decision to be made on the choice of appropriate software.

The first need was to conduct surveys of what data were available and it was considered sensible to send questionnaires to national focal points. It was also important to check the availability of suitable maps, data on drugs, and socio-economic information. The feasibility of the modelling work was the hardest component. The group's priorities were to ensure that some projects took place, possibly in pilot areas such as Amsterdam, Glasgow and Slovakia.

The two priorities for this group were identified as:

- undertaking some pilot mapping exercises;
- developing the analytic framework.

Economic modelling of markets

The group brought a number of different perspectives on the analysis of drug markets. Available data were discussed, as was whether these data could explain effects such as large changes in prices. A number of complex hypotheses relating market factors to drug use and drug-related consequences, such as deaths, were discussed by the group. The relationship between individual behaviour and how that translates into market aggregates, as displayed in some of the data presented by Leif Lenke on the impact of price changes on drug-use patterns, was also considered. The conclusion from this discussion was that there was a need for data and analysis in the following areas:

- the framework of markets across Europe;
- how individuals behave – micro-economics;
- using the first two areas to inform macro models.

There was further discussion about understanding the markets in Europe and how they may differ from those in the US. There were also some differences within Europe. Variations across drugs and types of users and the interactions between markets may well be more fluid in Europe than in the US. There was a need to gather information and to model the supplying organisations and the role of users/dealers.

From this discussion three areas of research were identified:

- a qualitative study of the individuals involved in the organisation of drug markets across Europe;
- a review and analysis of studies of the micro-economic behaviour to set out potential models which could be developed and tested within the European market;
- an examination of the potential for building macro-level models. This would consider the implications of control and other policies on markets and be capable of examining the role of other factors such as demographics.

Costs and cost–benefit studies

It is clear that the social costs of drug use and the cost-effectiveness of different policy options are of great public and policy interest, although the results of cost–benefit studies will not always directly influence the policy process. There was considerable discussion on what areas would form the priority of any research. Data and research methods were well tried within the health area. It was agreed that evaluating criminal-justice interventions was more difficult but of vital policy interest, especially given the proportion of public expenditure in that area.

Three broad areas of work were identified and discussed further:

- social-cost studies;
- cost–benefit or cost-effectiveness studies;
- public-expenditure analysis.

Social-cost studies have considerable public impact and the results of a study of the costs of illicit drugs in Canada (Single et al., 1996) were discussed. There was considerable debate about the methodology adopted. The international guidelines developed under the aegis of the Canadian Centre on Substance Abuse provide a review and a starting point for cross-European studies (Single et al., 1995). There is still some development required in a number of areas, especially the estimates of lost productivity.

Any research in this area would be best undertaken within a common framework and it is clear that a multi-disciplinary approach is required. There is also a case for a number of sub-studies, for example on specific health-care issues. In addition, other studies, for example looking at public-expenditure implications, are important.

In terms of economic evaluation there is a need for a similar set of linked activities. Work is required to develop the application of economic techniques to illicit drug interventions, for example in identifying, measuring and valuing many of the multiple outcomes from drug programmes. The most obvious place to start is to conduct evaluations within each sector, for example comparing different treatment approaches; enforcement; prevention; public nuisance; and harm-reduction interventions. The next step is to consider between-programme evaluation, for example the balance of public spending on enforcement and treatment.

The group considered the priorities from this long list to be:

- developing guidelines for social cost for Europe and undertaking a number of applications;
- developing a framework for the cost–benefit analysis of drug interventions.

Social processes and initiation of drug use

This is a key area but is still relatively under-researched. The group discussed data on the onset of drug use and came to the following conclusions. While there can be models, they have to be interpreted with care, as drug use is not the same as addiction. It is therefore important to develop measures which reflect health and social processes. This suggests a need to analyse drug careers, not the onset of drug use. However, this requires richer data sets.

Even if models are developed, care needs to be taken that results are not over-generalised. The same model would not apply to all drugs, cultures or time periods. There was also a need to recognise the role of social, environmental and policy variables.

The group prioritised the following areas:

- modelling risk factors across countries;
- measuring risk behaviours and developing methods to analyse these data.

Time trends and incidence

The participants in the Lisbon seminar came from very varied backgrounds including treatment centres, epidemiology, criminology, geography and mathematics and statistical modelling. There was a brief discussion of the review papers from the 1997 York meetings and then the group concentrated on the potential use of the back-calculation method. There was debate about the feasibility of using the method and the parallel use of survey data. There were some questions about the reliability of information and the impact of changes over time, for example in services available. It was thought, however, that there was potential for application to the analysis of first treatment and first police arrest.

The group also considered the link with policy. Why is it necessary to know incidence and time trends of heroin use? Clearly there is a need to consider the future demand for treatment services and to predict consequences such as those of hepatitis C. There is also a need to know how to chart the evaluation of overall drug strategies and policies.

The group went on to consider the study of other drugs, perhaps through the use of survey data. However, it was thought important not to dilute effort and concentrating initially on one drug may be more practical. A further extension was to consider the model with cessation rates, which again is of prime policy concern.

The first step was to consider the time lags involved and the appropriate distribution to employ. There was potential to learn from other countries outside Europe. The group

concluded that there were sufficient data and expertise, but some resources were needed to bring these together.

The main priorities were:

- developing and testing the model for heroin;
- considering the application to other drugs.

Conclusions

A number of important policy questions could be addressed by the different modelling techniques described in this monograph. Some techniques are based on complex mathematics or theories, while others can be presented in simpler terms for wider audiences. The debates at the two expert meetings in York and the later seminar in Lisbon were lively and generated many ideas for future work. It was clear that a wide agenda for research has been set out and while there was some purpose to maintaining the mechanisms to keep the wider group in contact, there was also a need for sub-groups to take some areas forward to test the practicalities of model building. Cooperation is required both across academic disciplines and between researchers and policy-makers throughout Europe. Several projects are currently under way to test some of the ideas set out in the following chapters.

References

Godfrey, C., Waby, V., Wiessing, L., Hartnoll, R. (1998). *Report of the seminar 'Drug use research, policy and dynamic modelling'*, EMCDDA, Lisbon, 7–9 May 1998, CT.97.EP.10. York and Lisbon: Centre for Health Economics and European Monitoring Centre for Drugs and Drug Addiction.

Rydell, C. P., Everingham, S. S. (1994). *Controlling cocaine: supply versus demand programs*. Santa Monica, CA: Drug Policy Research Center, RAND.

Single, E., Easton, B., Collins, D., Harwood, H., Lapsley, H., Maynard, A. (1995). *International guidelines for estimating the costs of substance misuse*. Ottawa: Canadian Centre on Substance Abuse.

Single, E., Robson, L., Xie, X., Rhem, J. (1996). *The costs of substance misuse in Canada*. Ottawa: Canadian Centre on Substance Abuse.

MODELLING DRUG USE WITH AVAILABLE DATA

INTRODUCTION

The models discussed in Part II attempt to integrate different indicators to provide meaningful interpretations of trends in problem drug use. The models vary in their assumptions and uses and are relatively data-demanding, but also may provide information that remains fairly close to direct observations, thereby increasing reliability.

In Chapter 3, Frischer and Heatlie explore the potential of linking geographic information systems to a drug-forecasting model. They envisage integrating different sources of data to produce predictions of geographic spread that can be mapped. They argue that increased mobility and increased availability of data in Europe would make such a system useful at European level. Their model is based on a limited number of parameters, but it is able to describe adequately the spread of drug use over larger geographic areas (micro-diffusion), similar to infectious-disease models.

Carla Rossi concentrates on the different sources of available data that might give information about the 'hidden' drug phenomenon in Chapter 4. She stresses the importance of a systematic approach to understanding and describing the data-production processes that result in observable records in agencies. Rossi argues that models describing observable macro-data in a simple manner are more useful for policy purposes than complex models that attempt to mimic the micro-phenomenon of the spread of drug use among individuals. Rossi discusses indicators available in different countries and suggests that appropriate mathematical and statistical functions should be used to describe the indicators taking account of their interdependencies.

In Chapter 5, Mollica and Mariani follow a similar line of thinking as Rossi by proposing the use of a multiple-indicator model for monitoring prevalence and risk of drug addiction. By combining time series of drug-related data from different sources the authors suggest that the relationships between the observed phenomena can be studied. In the proposed model, different indicators can in turn be seen as dependent or independent variables, depending on the aim of the analysis, while interactions and inter-correlations can be taken into account.

CHAPTER 3

MODELLING DRUG MISUSE IN EUROPE USING GEOGRAPHIC INFORMATION SYSTEMS

Martin Frischer and Heath Heatlie

Current methods for estimating incidence and prevalence of drug misuse tend to be retrospective and are not capable of forecasting spatio-temporal trends. Mapping of drug misuse is restricted to displays of incidence and prevalence rates. In contrast, various disciplines (e.g. ecology and environmental health) are increasingly using geographic information systems (GIS) to study associations between location, environment and behaviour. Advances in computing and graphical technology enable spatially referenced data to be linked to relational databases and epidemiological functions. GIS thus provide a powerful tool for analysing the spread of phenomena over time and space and GIS models have been used to predict disease spread from infected to susceptible populations. The development of a European drug-misuse GIS would be a logical progression in view of improvements in drug-misuse surveillance in Europe. It would provide: spatio-temporal maps of drug incidence and prevalence; dynamic methods of forecasting and visualising trends in drug diffusion; and information for identifying populations at risk and prioritising resource allocation.

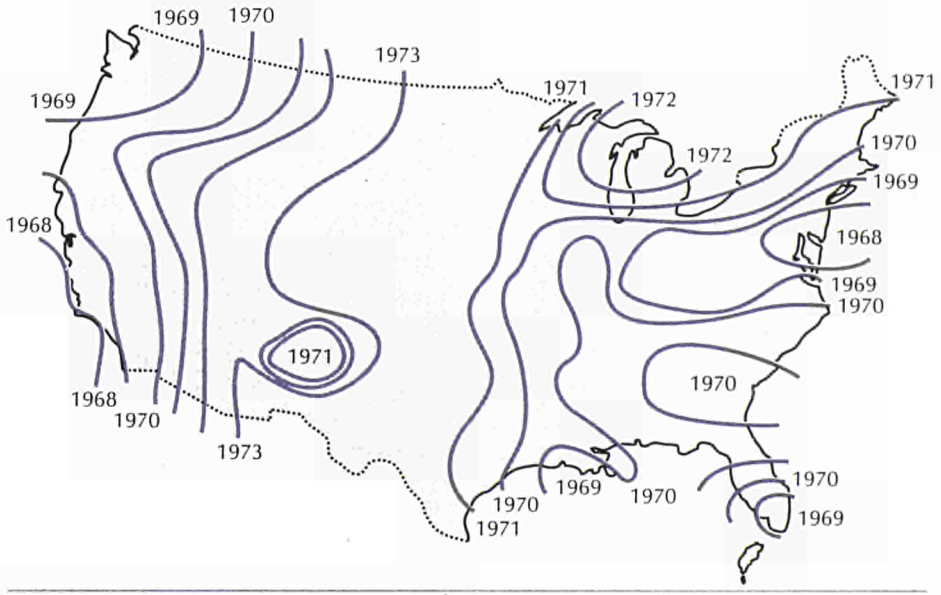
Increasing personal freedom and scope for travel within Europe have led to speculation that opportunities will arise for new forms of drug trafficking and drug use (Dorn and White, 1994). This has already occurred in the European part of the former Soviet Union (Lee, 1992) as a consequence of the social and economic changes which have taken place since the mid-1980s. However, there are few data from this region with which to paint a reliable picture of current prevalence of drug misuse. In Europe there is increasing recognition that reliable epidemiological information is important for prevention and treatment activities. In recent years there have been renewed efforts to ascertain the level of drug misuse in several countries using a range of methodologies (e.g. IFT, 1994; Sandwijk et al., 1995; ISDD, 1997). In general, methods such as population surveys and capture–recapture analyses provide period prevalence figures and sometimes incidence, but lack any intrinsic mechanism for forecasting trends.

In order to predict drug-use trends it is necessary to understand the factors that influence an individual's decision to use drugs. Concepts like personal freedom are not easily quantified, but, on the other hand, drug use has been linked to a range of intrapersonal, interpersonal and environmental factors (Ramsey and Percy, 1996). To the extent that these factors vary, consequent changes in drug use can be predicted.

Drug use is also dynamic, spreading throughout populations and across regions. Pioneering work by Hunt and Chambers (1976) in the United States focused on two processes (see Figure 1). The first process, called micro-diffusion, refers to the spread of drug use among individuals within groups. It depends on known drug users' propensity to 'transmit' drug use to new users in a similar manner to infectious diseases. The second process, macro-diffusion, refers to the transmission of drug use across geographical boundaries. This process is thought to occur in the same manner as other new phenomena (e.g. the geographical diffusion of television stations in the US in the 1960s). Hunt and Chambers were able to create a map showing the spread of heroin epidemics in the US during the 1970s by analysing data from drug-treatment programmes. Heroin use appears to have begun on the north-east coast along the chain of cities from Boston to Washington, and in southern California. Large inland and Gulf-coast cities were also early centres of epidemic use. Heroin moved to the interior, from these continental margins, spreading sequentially from cities in regions of high-population density to those of lower density.

Although Hunt and Chambers' analysis was an attempt to demonstrate that drug misuse spreads in a lawful way, their model was based on only two parameters: a simple infectious-diseases function (micro-diffusion); and empirical observations of geographical spread (macro-diffusion). The model did not link the geographical framework to other information, such as the population's socio-economic status or geographical features

Figure 1: Isochrony of peak heroin use in the United States, 1968–73



Source: adapted from Hunt and Chambers (1976)

such as transport routes. In the interval since the publication of their work, a new multidisciplinary methodology has emerged which could now incorporate such information by exploiting relational databases and spatio-temporal GIS.

What is a geographic information system?

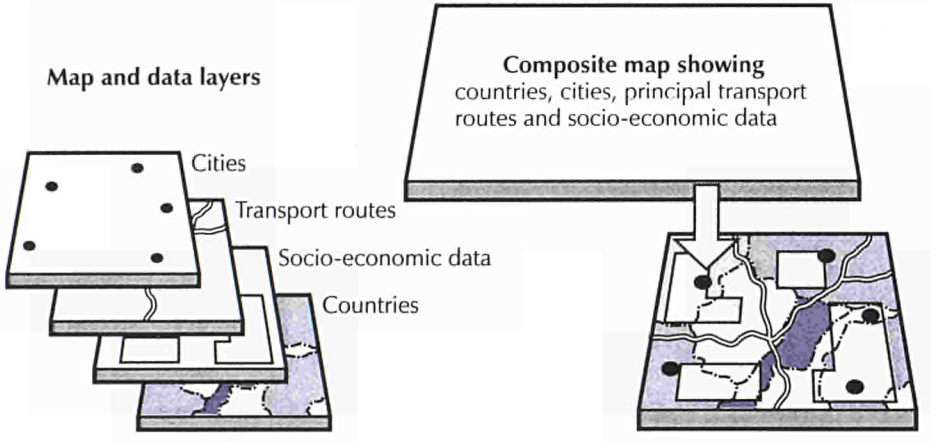
Although there are many types of geographic information system, they are all capable of capturing, analysing and displaying spatially referenced data. GIS development has been facilitated by the availability of digital map data which provide a spatial framework on which to attach the attribute data (see Figure 2). The main characteristic of a GIS is the ability to link spatially referenced data to a relational database that contains relevant information (e.g. population data, transport routes or other socio-economic data). These features, together with more powerful spatial-analysis tools, distinguish GIS from early mapping programmes that simply displayed information.

GIS capabilities

Data capture and storage

Increasingly, maps and associated attribute data (e.g. census statistics) are commercially available, but where no data exist for a project, they can be entered into a GIS from external sources by either scanning maps or tracing over a map's features using a digitising tablet. Attribute data can be entered from either existing studies or public-domain data sets.

Figure 2: Schematic representation of a typical geographic information system



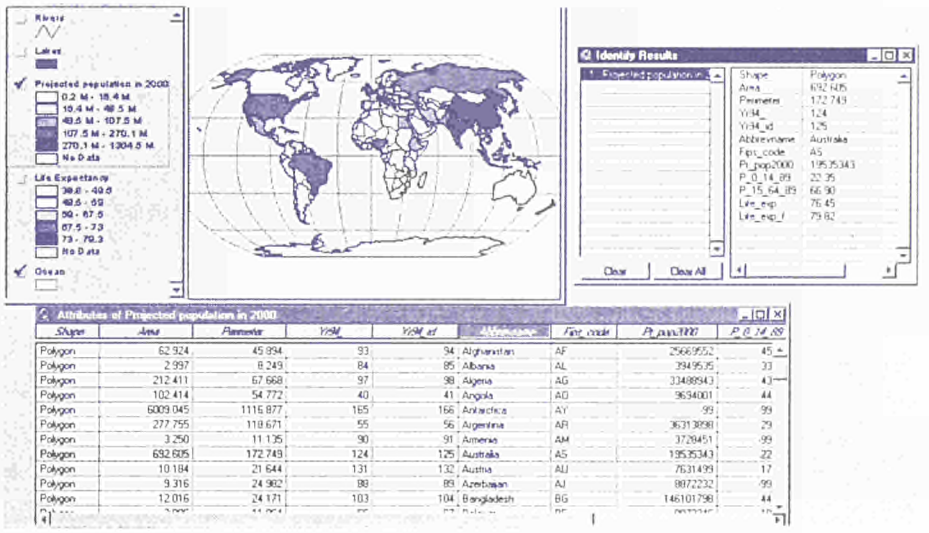
A GIS stores both spatial map data and associated attribute data. Attribute data are stored in a relational database-management system contained within the GIS and accessed by a spreadsheet or query-driven user interface. The GIS can accomplish everything that a traditional database system can by querying, selecting and manipulating data.

Map data are encoded into a set of numbers so that the geometry of the map is available to enable spatial queries, but also so that the maps can be stored digitally in one or more files. Maps are encoded by using both the coordinate information and encoded topology, so that the relationships between points, lines and areas, such as the adjacency of regions or the connectivity of lines, are known in advance.

Data selection

Data records in a GIS can be retrieved in one of two ways. First, the relational database can be searched and selections made on the basis of a feature's attributes and their values (see Figure 3). For example, European cities with household surveys within the last two years could be selected and then displayed. Second, a GIS also allows spatial retrieval, for example, the map could be searched for all European towns with populations of 50–100,000 and within 100 km of a major city. Each query could be linked to several data layers, for example, socio-economic status and transport links.

Figure 3: Geographic information systems can be queried via database or map (spatial representation)



Spatial analysis and GIS

There are generally three phases in applying GIS: visualisation; exploratory data analysis; and model building (Gatrell and Bailey, 1995).

Visualisation

Visualisation is an important tool for showing the change in disease patterns over time. Animation, embedded within a GIS, is highly effective in depicting the spread or retreat of disease over space and time. The spread of the AIDS epidemic in the United States as it moved from and within major cities was illustrated using animated maps (Openshaw et al., 1987). Figure 1, above, shows a similar process for drug misuse. Visualisation can also be used in novel ways to explore the results of traditional statistical analysis. Displaying the locations of outlier and influential values on maps and showing variation in values over space can add a great deal to epidemiologic research. Although such tools are still being developed, they would benefit greatly from a closer and more seamless link between statistical packages and GIS (Gatrell and Bailey, 1995).

Exploratory data analysis

The quantity and diversity of spatial data in a GIS can be overwhelming and exploratory methods can help the analyst make sense of data to address 'what if' questions. Advances in computing and graphics technology have made this one of the most active areas in GIS/spatial-analysis research. Among the most important exploratory methods for epidemiology and public health are methods for identifying space–time clusters or 'hot spots' of disease. For example, Marshall (1991) produced maps displaying the proximity of significant disease clusters to environmental hazards such as nuclear facilities.

Spatial-diffusion models analyse and predict the spread of phenomena over space and time and have been widely used in understanding the spatial diffusion of disease spread (Thomas, 1990). Such models are quite similar to spatial-interaction models except that they have an explicit temporal dimension. The models can predict how diseases spread, spatially and temporally, from infected to susceptible people in an area by incorporating time and space, along with basic epidemiological concepts (Haggett, 1994) and aid in understanding the emergence of infectious disease.

Spatial auto-correlation methods and space–time correlograms can be used to explore spatial and temporal patterns. These methods provide a general sense of the speed and geographic pattern of diffusion. Although the methods have not typically been incorporated into GIS, there is great potential for doing so, especially with recent advances in computer animation.

Model building

Modelling includes procedures for testing hypotheses about the causes of disease and the nature of disease transmission. In general, modelling involves the integration of GIS

with statistical and epidemiological methods. GIS can generate data for epidemiological models, display the results of statistical analysis, and model processes that occur over space. The first two points are evident in recent, regression-based analyses of disease risk, such as the study of Lyme disease (Glass et al., 1995). In this study, a GIS was used not only to integrate diverse data sets and calculate new variables, such as slope and distance from forest, but also to map geographic variation in disease risk, as predicted from a logistic-regression model.

Other GIS models are more explicitly spatial, expressing relationships or flows between people and places. Spatial-interaction and spatial-diffusion models are of particular relevance to the study of emerging diseases. Spatial-interaction models analyse and predict the movement of people, information and goods from place to place (Ding and Fotheringham, 1992). The flows of people between rural areas, villages, cities and countries are all forms of spatial interaction that are central to disease transmission. It is possible to identify areas most at risk for disease transmission by accurately modelling these flows and thus target intervention efforts. Spatial-interaction models reflect two principles: that interaction decreases with distance; and that it increases with population size or 'attractiveness'. Given actual flow data, values can be estimated that show the effects of distance and population size (or other 'attractiveness' factors) on interaction. The models can then be used to predict spatial-interaction patterns elsewhere. Although spatial-interaction models and GIS developed separately, some GIS now have spatial-interaction-modelling capabilities.

Applications of GIS in epidemiology

GIS have already been used to study associations between location, environment and disease. Recent applications include the surveillance and monitoring of vector-borne diseases (Glass et al., 1995; Beck et al., 1995; Richards, 1993); water-borne diseases (Clarke et al., 1991); environmental health hazards (Braddock, 1994); exposure to electromagnetic fields (Wartenberg et al., 1993); predicting child pedestrian injuries (Barnes and Peck, 1994); and prescribing patterns in primary and secondary care (Pryce et al., 1996).

In a recent study in Baltimore County, Maryland, GIS and epidemiologic methods were combined to identify and locate environmental risk factors associated with Lyme disease (Glass et al., 1995). Ecological data such as watershed, land use, soil type, geology and forest distribution were collected at the residences of Lyme disease patients and compared with data collected at a randomly selected set of addresses. A risk model was generated combining both GIS and logistic-regression analysis to locate areas where Lyme disease is most likely to occur.

GIS are being used to identify locations of high prevalence and monitor intervention and control programmes in areas of Guatemala for onchocerciasis (Richards, 1993) and in Africa for trypanosomiasis (WHO, 1989). Spatial and ecological data are combined with epidemiological data to enable analysis of variables that play important roles in disease transmission. This integration of data is essential for health-policy planning, decision-making and ongoing surveillance efforts. For example, as part of the Guinea worm eradication effort, the United Nation's Children's Emergency Fund placed pumps

in villages most infected with the disease to ensure access to a safe water supply (Kitron et al., 1994). GIS enabled researchers to locate high-prevalence areas and populations at risk, identify areas in need of resources and make decisions on resource allocation (Roger et al., 1993). Epidemiological data showed a marked reduction in prevalence in villages where pumps were introduced.

GIS and drug misuse

GIS have already been shown to be useful in epidemiological research and for facilitating interventions. While mapping of drug prevalence is currently rudimentary, the pioneering work of Hunt and Chambers provides a basic framework for developing GIS in relation to drug misuse. However, developing a practical GIS system will require careful consideration of several issues.

Scope

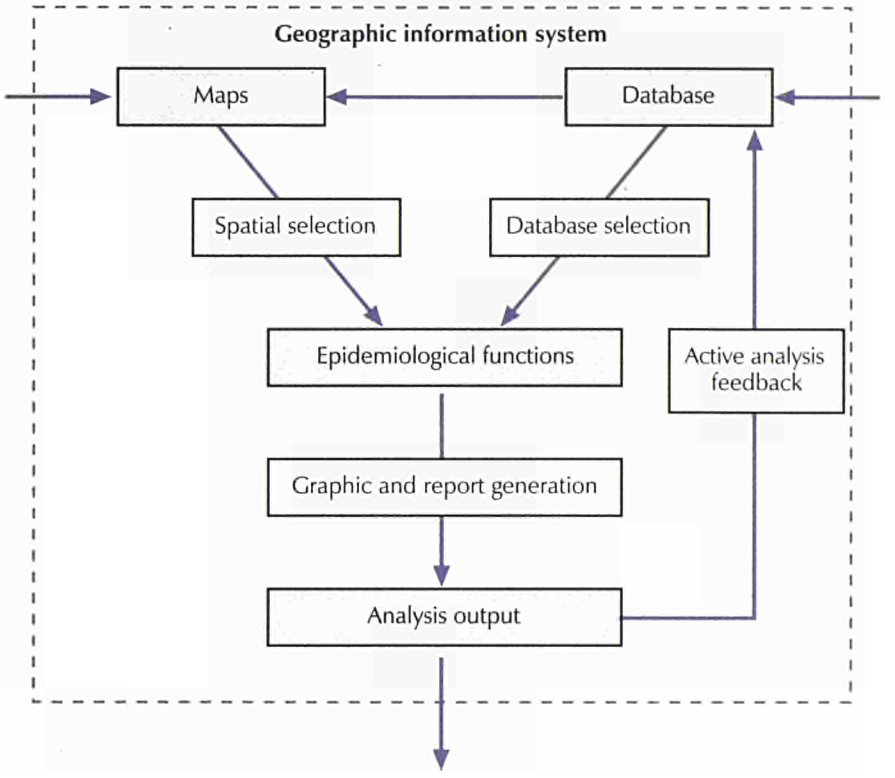
The main aim of a GIS drug-misuse system is to create a dynamic model for forecasting and displaying spatio-temporal trends. As mentioned at the start of this chapter, drug misuse is dependent on a wide range of factors, only some of which are likely to be successfully modelled. Previous research indicates that standard variables such as socio-economic status are generally not very good predictors. This is probably because diffusion of drug use is a dynamic process. Such processes have not been modelled, although there has been some interesting qualitative work on peer groups and social networks (Neagius et al., 1994). Available data on diffusion indicate that drug epidemics tend not to be indigenous, but are dependent on a variety of economic, social and political factors. The extent of diffusion also depends on individuals' predisposition to experiment with psychoactive substances. In Western Europe, increasing levels of post-war drug use have been attributed to greater individualism and consumerism. Psychological factors may also play an important role with the weakening of family relationships and lengthening periods of adolescence (Rutter and Smith, 1995).

One of the main impediments to applying GIS to drug misuse has been the lack of empirical data to include in a GIS database. However, during the 1990s there have been numerous studies and considerable improvement in European surveillance data (see EMCDDA, 1997). Furthermore, as mentioned above, GIS maps and relational databases can now be created or purchased. These developments have created a climate whereby a GIS for drug misuse can now be realistically considered. Figure 4 highlights the conceptual framework underpinning the proposed GIS for drug misuse.

Steps in creating a drug-misuse GIS

Although drug use can diffuse at various levels (micro, macro, international), the common link is a function that specifies how drug use spreads among individuals in a population. Two pilot computer programs have been developed based on deprivation status and town size. Both programs were developed for the Strathclyde region of Scotland.

Figure 4: Components of a geographic information system



The first program models drug diffusion in relation to deprivation within the city of Glasgow. It assumes that drug use starts in the 'worst' areas within Glasgow and spreads to 'better' areas. The program forecasts incidence and prevalence, in relation to six parameters:

- population size of the various areas (graded 1 to 8 in terms of affluence);
- estimated peak prevalence (can be variable over affluence);
- speed of spread across types of area;
- duration of addicts' drug-using career;
- length of epidemic cycle;
- year epidemic trends were first observed.

The second program assumes that drug use spreads from denser to less dense areas (i.e. urban to rural) within the Strathclyde conurbation (of which Glasgow is the major city). All other parameters are the same. Using the observation that epidemic trends in drug use were first noticed in Glasgow in 1980, it has been possible – with hindsight –

to generate incidence and prevalence estimates which correspond to capture–recapture estimates of injecting drug use in Glasgow in 1990.

These programs represent a preliminary attempt to model drug diffusion. Although they do not constitute a GIS, they do include rudimentary relational and geographical components. In order to create a GIS for drug-misuse forecasting, the following would have to take place to develop relational, epidemiological and mapping capabilities:

- obtaining digital data to construct the spatial framework for the study area (digital maps);
- selecting appropriate attribute data to help explain the diffusion of drug misuse (census data, national surveys, transport flows, major import routes);
- developing new models to predict the rate of diffusion (e.g. allow modelling for the effect of policy/intervention (analogy with vaccination campaign for infectious disease);
- creating an output interface to allow dynamic user interaction and temporal displays of predicted diffusion rates given a 'what if' scenario.

Conclusions

In many areas of Europe, patterns of drug use are changing. The mechanisms of diffusion are diverse: introduction of new practices by new users; tourism and migration; cross-border contact; drug transportation; and increasing opportunities for economic and international contact. GIS offer a dynamic and flexible approach to visualising the diffusion of drug use in Europe.

What are the model outputs?

The creation of a drug-misuse GIS would produce:

- spatio-temporal maps of drug incidence and prevalence;
- dynamic methods of forecasting and visualising trends in drug diffusion;
- information for identifying populations at risk and areas requiring resources for prevention and treatment.

To what extent are existing applications of these models applicable in Europe?

GIS have been used to address diverse topics in the fields of ecology, infectious diseases and environmental health.

What are the data demands of the model?

The GIS requires:

- digital maps of Europe which can be focused at city, region, country and international level;

- relational databases of socio-demographic data, population statistics, known drug-use statistics (e.g. surveys, prevalence-estimation studies);
- epidemiological functions for estimating drug incidence and prevalence.

Why have the models not been applied more extensively in Europe?

Although drug-use trends are increasingly being mapped, GIS development has been slow for a number of reasons including lack of knowledge regarding GIS capabilities, inadequate relational data and rudimentary epidemiological knowledge of drug diffusion. GIS development will require a multi-disciplinary team.

Should there be a research project to develop and apply these models in Europe?

The development of a drug-misuse GIS would create a powerful visualisation and forecasting tool with easily understandable outcomes. A drug-misuse GIS could also be a focal point for integrating and developing epidemiological understanding of trends in drug use within and across Europe.

What geographic areas do the models consider?

A flexible GIS system would be capable of modelling drug use at city, regional, country and international levels.

References

Barnes, S., Peck, A. (1994). 'Mapping the future of health care: GIS applications in health care analysis'. *Geographic Information Systems* 4: 31–3.

Beck, L. R., Rodrigues, M. H., Dister, S.W., Rodrigues, A. D., Rejmankova, E., Ulloa, A. et al. (1995). 'Remote sensing as a landscape epidemiological tool to identify villages at high risk for malaria transmission'. *The American Journal of Tropical Medicine and Hygiene* 51: 71–80.

Braddock, M., Lapidus, G., Cromley, E., Cromley, R., Burke, G., Branco, L. (1994). 'Using a geographic information system to understand child pedestrian injury'. *American Journal of Public Health* 84: 1158–61.

Clarke, K. C., Oslee, J. R., Sherry, J. M., Meert, J. P., Larsson, R. W. (1991). 'The use of remote sensing and geographic information systems in UNICEF's dracunculiasis (Guinea worm) eradication effort'. *Preventive Veterinary Medicine* 11: 229–35.

Ding, Y., Fotheringham, A. S. (1992). 'The integration of spatial analysis and GIS'. *Computers in Environmental and Urban Systems* 3–19.

Dorn, N., White, S. (1994). 'European free movement, drug users and access to health services'. *International Journal of Drug Policy* 5: 35–44.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (1997). *1996 Annual report on the state of the drugs problem in Europe*. Lisbon: EMCDDA.

Gatrell, A., Bailey, T. (1995). 'Can GIS be made to sing and dance to an epidemiological tune?'. Paper presented at the International Symposium on Computer Mapping and Environmental Health, Tampa, FL, February.

Glass, G. E., Schwartz, B. S., Morgan, J., Johnson, D. T., Noy, P. M., Israel, E. (1995). 'Environmental risk factors for Lyme disease identified with geographic information systems'. *American Journal of Public Health* 85: 944–8.

Haggett, P. (1994). 'Geographical aspects of the emergence of infectious diseases'. *Geografiska Annaler B* 76: 91–104.

Hunt, L. G., Chambers, C. D. (1976). *The heroin epidemics: a study of heroin use in the United States, 1965–1975*. New York: Spectrum Books.

Institut für Therapieforschung (IFT) (1994). *Report on the methods of estimating the extent of the drug problems in Germany*. Munich: IFT.

Institute for the Study of Drug Dependence (ISDD) (1997). *Drug misuse in Britain 1996*. London: ISDD.

Kitron, U., Pener, H., Costin, C., Orshan, L., Greenberg, Z., Shalom, U. (1994). 'Geographic information system in malaria surveillance: mosquito breeding and imported cases in Israel, 1992'. *The American Journal of Tropical Medicine and Hygiene* 50: 550–6.

Lee, R. W. (1992). 'Dynamics of the Soviet illicit drug market'. *Crime, Law and Social Change* 17 (3): 177–233.

Marshall, R. (1991). 'A review of methods for the statistical analysis of spatial patterns of disease'. *Journal of the Royal Statistics Society* 154: 421–41.

Neagius, A., Friedman, S. R., Curtis, R. et al. (1994). 'The relevance of drug injectors' social and risk networks for understanding and preventing HIV infection'. *Social Science and Medicine* 38 (1): 67–78.

Openshaw, S., Charlton, M., Wymer, C., Craft, A. (1987). 'A mark 1 geographical analysis machine for the automated analysis of point data sets'. *International Journal of Geographical Information Systems* 1: 335–58.

Pryce, A., Heatlie, H. F., Chapman, S. R. (1996). 'Buccaling under the pressure: influences of secondary care establishments on the prescribing of glyceryl trinitrate buccal tablets in primary care'. *British Medical Journal* 313: 1621–4.

Ramsey, M., Percy, A. (1996). *Drug misuse declared: results of the 1994 British Crime Survey*. London: Home Office.

Richards, F. O. (1993). 'Use of geographic information systems in control programs for onchocerciasis in Guatemala'. *Bulletin of Pan American Health Organisations* 27: 52–5.

Roger, D. J., Williams, B. G. (1993). 'Monitoring trypanosomiasis in space and time'. *Parasitology* 106 (Suppl.): 277–92.

Rutter, M., Smith, D. (1995). *Psychosocial disorders in young people*. London: Academia Europaea.

Sandwijk, J., Cohen, P., Musterd, S., Langemeijer, M. (1995). *Licit and illicit drug use in Amsterdam II: report on a household survey in 1994 on the prevalence of drug use among the population of 12 years and over*. Amsterdam: University of Amsterdam.

Thomas, R. (1990). *Geomedical systems: intervention and control*. New York: Routledge.

Wartenberg, D., Greenberg, M., Lathrop, R. (1993). 'Identification and characterisation of populations living near high-voltage transmission lines: a pilot study'. *Environmental Health Perspectives* 101: 626–32.

World Health Organization (WHO) (1989). 'Dracunculiasis: global surveillance summary'. *World Health Organization Bulletin* 68: 797–8.

CHAPTER 4

USING AVAILABLE DATA WITH DYNAMIC MODELS TO ESTIMATE THE EXTENT OF PROBLEM DRUG USE

Carla Rossi

Problem drug use is a hidden phenomenon that can only be studied by indirect indicators linked to its observable consequences, such as crime, health problems, infectious diseases and death. This chapter begins with a description of possible indicators available in different countries. One of the major problems in monitoring problem drug use and in evaluating control strategies is determining the extent of the problem, that is, the size of the drug-using population. To this end, various mathematical models have been used. Specifically, these models are based on the production processes of the observable data, which produce the observable indicators. Such models allow the data collected by social agencies to be used for estimating the size of the hidden population of interest and for making inferences. Several examples are reported to show how to implement the various methods, and potentialities and problems are discussed. The models taken into account are essentially dynamic models, such as Poisson models, multiple capture–recapture models and infectious-disease models. The sources of available data are also discussed and various existing data from different sources and countries are used in the examples of application.

Models, data and indicators for hidden processes

In developing policies and evaluating interventions for controlling the use and abuse of illicit drugs, it is necessary clearly to define the objectives to be reached, the means of achieving those objectives and, most importantly, the measurements (i.e. data) to be used for monitoring interventions. To analyse a complex phenomenon like drug abuse, a systematic approach must be adopted, including the use of statistical methodologies for both describing and interpreting the current situation and proper mathematical and statistical models for making inferences and predicting trends. However, to perform these types of analyses on the collective phenomenon of the illicit drug market, it is necessary to have reliable data on the various critical events that constitute the visible part of the phenomenon and to use suitable mathematical models of the production processes of such observational data (Rossi, 1999a).

In fact, observational data, which are already available in various countries, should play an important role in developing and monitoring national and international drug policies

and in developing suitable mathematical and statistical models to support decision-making. Mathematical models describing currently available macro-data are much more useful than complex mathematical models for mirroring micro-phenomena, which require different kinds of experimental processes of data collection and would be difficult to implement. These complex models would also be very costly and time consuming, as shown for ad hoc population surveys for measuring the extent of illegal substance use and abuse among the general population or particular subgroups.

In any case, one of the primary concerns is to determine the nature and the extent of the problem so that appropriate objectives and strategies can be developed; thus the first step is to estimate the size of the population of drug users and, more specifically, of 'problem' drug users. However, since the use of psychoactive drugs is restricted or considered illicit in most countries, drug consumers and dealers are forced to conceal their habits. As a consequence, the extent of illicit drug use and abuse and the size of the drug market can only be measured by indirect methods based on mathematical models that mirror the behaviours of various measures or indicators, such as those developed to estimate 'hidden' phenomena (NIDA, 1990). Table 1 lists a set of indirect indicators developed within the *World drug report* project of the United Nations International Drug Control Programme (UNDCP). These indicators refer to some observable aspects of drug abuse, known as 'drug-related phenomena'. Indeed, it is quite difficult to identify epidemiological aspects related solely to the use of illegal drugs, particularly if such substances are not considered to represent serious health hazards (e.g. cannabis, whose use does not seem to cause any major detectable problems).

All of the indicators listed in Table 1 are relative measures obtained by dividing the absolute frequencies by the sizes of the corresponding reference populations. Statistics on drug seizures can be simply expressed in per capita units with respect to the general population. Although these statistics are of little use in analyses of a single country where absolute figures can also be used, they are extremely useful for cross-sectional analyses.

Some of the indicators consist of directly observed data recorded by social agencies, for instance data on drug seizures, which are recorded by the police, or data on persons detained, which are recorded by justice departments. Other indicators are based on different records kept by social agencies or on survey data. Since a variety of data sources are used, the quality and accuracy of the indicators are not comparable. In particular, data from surveys conducted in different countries with the aim of estimating the prevalence of drug use cannot generally be compared because of major differences in the basic methodology.

Many indicators are indirect measures of several phenomena simultaneously, in particular, most indicators of drug-related phenomena are linked to both health consequences and law enforcement and are indirect measures of the extent of the illicit drug market and, consequently, of illicit drug use as well. Thus these indicators cannot be used independently, and appropriate mathematical and statistical models are required. For instance, the absolute frequency of persons detained for drug-related offences, denoted by 'nd', can be supposed to be a function of law-enforcement policies,

Table 1: Possible indicators of drug-related phenomena

Indicator	Measuring
Drug-related deaths	Risky use of drugs/extent of drug market
Hospital admissions	Drug-related morbidity/extent of drug market
Clients of drug-treatment services	Rehabilitation policy/extent of drug market
AIDS and other infections due to drug use	Risky use of drugs/extent of drug market
Injecting drug use	Risky use of drugs
Frequency or daily abuse	Risky use of drugs/extent of drug market
Life-time prevalence of use	Extent of use
Number of seizures/quantities seized	Law-enforcement policy/extent of drug market
Warnings and similar actions	Repression policy/extent of drug market
Persons detained for drug-related offences	Law-enforcement policy/extent of drug market
Drug abusers detained for other crimes	Drug-related criminality/extent of drug market
Expenditure on education, community activities, etc.	Prevention policy

denoted by 'le', and of the extent of the illicit drug market, denoted by 'dm', as in the following equation:

$$nd=f(le,dm)$$

In particular, the equation below can be considered as quite reasonable

$$nd=f\{(le)(dm)\}=k(le)(dm)$$

assuming that nd is equal to zero if either the extent of the market or the law-enforcement is zero and that nd is reasonably proportional to both quantities. Other indicators can be estimated with similar equations. Since most of these indicators depend on dm, there is a strong mutual correlation among them.

To obtain adequate results in the processes of modelling and of making inferences, it is thus particularly important to analyse the different means used for collecting the necessary data, that is, to discuss the different sources of data and to set up suitable dynamic models of the data-production processes. In fact, the quality of data may strongly depend on how they are collected and on the specific objectives of the data collection: these aspects should be taken into account when proposing a mathematical model.

The following section, 'Data sources', describes the main sources of experimental and observational data on drug-related phenomena. This is followed by a brief description of the various drug-related phenomena and by a critical analysis of several studies conducted in different countries with the aim of investigating the peculiarities of the

various approaches and the different assumptions that are needed to develop sound inferential procedures. The final section of this chapter describes certain mathematical models, both classical and original, and evaluates their application to existing data sets from different countries, thus allowing for a better understanding of the advantages and limitations of these models.

Data sources

There are two main sources of data on drug-related phenomena: records from social agencies such as the police, hospitals or prisons, which produce observational data; and anonymous surveys for studying the different phenomena (e.g. prevalence among high-school seniors), which produce experimental data. Both sources have their advantages and limitations. Data from social agencies are usually recorded for administrative purposes rather than for monitoring the phenomena. As a result, the format of the records is not ideal for making quantitative socio-economic analyses or for easily choosing mathematical models and making inferences. Furthermore, such data can only be released in an aggregated form via statistical tables, and the aggregation often only serves some bureaucratic purpose (e.g. reporting data to the national parliament or to the United Nations). Consequently, the most useful information (i.e. correlations among different phenomena) is usually lost. Surveys can be used to gather more complete and complex information on the phenomena via an appropriate questionnaire. However, if the questionnaire is too complex and/or the questions on illicit drug use are too direct, some respondents may have difficulties in understanding the questions or they may try to conceal their illegal habits, even when anonymity is guaranteed. Other respondents may simply decide not to answer certain questions, thus creating the problem of missing data, which cannot be addressed using straightforward statistical methods. Moreover, surveys are costly and time-consuming.

Examples of both sources of data (i.e. social agencies and surveys) are provided below in order to explain how these sources can best be used and to suggest how the data-gathering process can be improved and modelled to make inferences on the various hidden processes of interest.

Surveys (experimental data)

Surveys constitute a means of systematically gathering information on a given population. Those conducted on all members of the population are referred to as 'census surveys', whereas those targeting specific sub-populations or groups are known as 'sample surveys'. Determining which of these two types of surveys is most appropriate is generally based on the specific objectives to be reached and the resources that are available: for studying the phenomenon of drug use and abuse, sample surveys are in most cases the only reasonable choice.

When designing a sample survey, particular care must be taken in choosing the sampling technique. Only random sampling allows results to be generalised to the entire population, keeping the estimation error under control. Unfortunately, random sampling

is not feasible for the hidden populations considered here. Thus, to minimise errors, the only alternative is to make suitable prior assumptions and to properly model, mathematically, the data-generating process.

Records from social agencies (observational data)

Many different agencies collect data related to illegal drug use and abuse. Data collected in connection with the process of law enforcement can be distinguished from those collected in connection with the process of health care. Unfortunately, the collection of data by social agencies is strictly related to the given country's regulations, laws and policies on drug-related phenomena, and available information only refers to the visible part of the hidden processes. In particular, for drug use and abuse, some inference can be made based on data collected on the health consequences of drug abuse, whereas for aspects linked to the illegal market or crime, data on different aspects of law enforcement can be used. In any case, for this source of data, as for data from surveys, it is crucial to make suitable assumptions and to set up proper models for the data-production processes.

Furthermore, more reliable results can be obtained by pooling data from different sources (e.g. different social agencies and sample surveys) in order to overcome problems related to the existing correlations of the indicators. Finally, to gain an understanding of the overall issue, both observable and hidden drug-related phenomena must be analysed.

Drug-related phenomena

The drug problem involves not only those phenomena linked to addiction in the strict sense – such as morbidity due to infectious and chronic diseases, drug-related deaths (caused by overdose, infectious diseases, endocarditis and cirrhosis), and the medical assistance offered to drug abusers (assistance from the national health service and periods spent in live-in rehabilitation centres) – but also many other aspects that have a considerable impact on individuals who have no connection whatsoever with drug addiction (e.g. victims of petty crimes linked to drug addiction). The potential effects of these specific aspects are considered below.

Health consequences of illegal drug abuse

The health consequences of drug abuse include all aspects of the phenomenon that have an impact on health services: illnesses linked to drug addiction (e.g. hepatitis, endocarditis, septicaemia and AIDS); hospital admissions due to intoxication or other causes strictly related to drug use; initiatives undertaken by health services to rehabilitate drug abusers (e.g. out-patient and inpatient facilities and mobile units); and juvenile mortality linked to drug addiction (overdose, AIDS, violent deaths and deaths caused by infections and other drug-related syndromes). It must be stressed that these processes, which can be modelled to make inferences about the hidden phenomenon of drug use and abuse that cause them, greatly depend on the specific drug used and on existing policies. In fact, the type of drug most commonly injected varies by country: heroin and opiates in the European Union; cocaine in the United States; and central nervous

system stimulants in Sweden. This must be taken into account when proposing models. The different drug laws, policies and therapeutic treatments for drug addicts should also be taken into account. For example, in countries where hepatitis B prevention programmes targeting intravenous drug users (IDUs) were implemented in the early 1980s, the spread of HIV and AIDS seems to be limited among IDUs.

Law enforcement

Addressing the issue of law enforcement entails analysing the following aspects: police operations for controlling money laundering or the consumption of and trafficking in substances defined by legislation as narcotic drugs or psychoactive substances; people detained for drug-related crimes (i.e. for breaking drug laws or for being drug users who have committed serious or petty crimes connected with the drug market); and relative administrative or penal proceedings.

The illegal drug market

The illegal drug market can be quantified by determining or estimating the number of people working in this economic sector. Data for making inferences about this phenomenon using appropriate models can be obtained from records of drug seizures, incarcerations and proceedings for drug-related offences.

Organised and petty crime

'Petty crime' includes those crimes linked to the drug-using lifestyle of abusers (e.g. petty larceny, bag-snatching and pick-pocketing); 'organised crime' refers to the illegal drug market (e.g. production of and trafficking in drugs, and crimes committed by criminal organisations against persons belonging to rival 'cartels'). Based on the data concerning these drug-related phenomena and on the corresponding mathematical models, it is possible to perform the following types of analyses:

- analysis of an individual country for monitoring national drug policies and conducting cost-effectiveness analyses;
- panel studies for analysing and forecasting trends; and
- cross-sectional analyses for comparing regional or national policies or monitoring international policies.

Again, it must be stressed that the statistics available from social agencies and, at times, those from surveys, are strictly related to the given country's specific drug-control regulations, which greatly determine the specific categories used to disaggregate data.

Study examples

The examples provided below include descriptions of various methodologies and of the problems encountered when modelling processes of data production and collection to estimate the size of hidden populations and monitor drug-related phenomena. These

examples highlight the fact that, in order for a model to be useful, it must include empirical data on specific laws and policies. The data used for the examples were taken from the Core Countries database (see UNDCP, 1977).

A panel study: the military conscript survey in Sweden

This study is a census survey conducted among military conscripts using a specific questionnaire. One of the questions included in the questionnaire is whether or not the individual has used illegal substances. Given that this is a direct question that requires a direct answer, it is necessary to consider the degree of reliability of the information collected (i.e. the observed frequency of use based on the number of self-reported affirmative responses) and the potential to generalise this information (i.e. to estimate the prevalence among the population based on this observed frequency).

It is clear that simple reporting of use would be misleading, as there were many missing answers and the proportion of missing data increased sharply following the implementation of stricter drug regulations in Sweden in 1980: in 1980 there was a 4 % rate of missing data, increasing to 17 % in 1994. Specifically, this increase may reflect the passing of a law in 1992 on coercive care of adult drug abusers, as well as the criminalisation of drug use in 1988 and the introduction in 1993 of a maximum sentence of six months' imprisonment for drug use. Despite the fact that the conscripts were guaranteed anonymity, these laws may have influenced their decision of whether or not to answer the question.

To increase the reliability of this kind of survey, randomised response models (RRM) can be used, even when questionnaires include questions on the use of illegal substances. Though RRM entails additional costs and time, the results obtained (observed prevalences) are more reliable and can generally be used as unbiased estimates of the population prevalences. When using RRM, an appropriate (i.e. simple) stochastic model is needed to mirror the data-production mechanism. Selected results from an experiment based both on direct answers and on answers using RRM are reported in the country profile for Italy in the UNDCP *World drug report* (1997) and in Stovali (1993).

The RRM method is aimed at reassuring the respondent that affirmative answers to questions on illicit behaviour cannot be imputed to him or her. Specifically, the question is answered indirectly by the respondent, using the following procedure:

- When designing the questionnaire, the 'problematic' question is paired with a 'neutral' question, for instance:

  1. 'Have you used illicit drugs in the past month?'
  1'. 'Were you born in the first quarter of the year?'

- The respondent is requested to use a random mechanism with two outputs to choose which question to answer. The probabilities of the outputs of the random mechanism are known by the interviewer. For instance, the respondent is asked to toss dice, without the interviewer seeing the outcome, and to answer the first question if the result is 1 or 2 and the second question otherwise.

- The respondent simply answers 'yes' or 'no' and the specific question answered is known only to him or her.

If the unknown prevalence of the illicit behaviour is denoted by p, the probability of the first output in the random mechanism (in this case 1/3) denoted by q, the probability of an affirmative answer to the second question (in this case 1/4) by m, and the relative frequency of the affirmative answers by f, then p can be estimated using the following equation:

$$p=\{f-m(1-q)\}/q \text{ if the result is greater than zero}$$

$$p=0 \text{ otherwise}$$

The resulting estimate has been shown to be correct and consistent. In a study performed in Perugia, Italy, a large number of university students were interviewed to study the performance of the procedure using 'problematic' questions on drug use and racist behaviour. The results (Stovali, 1993) were quite satisfactory and show that the estimated prevalences obtained using direct questions were consistently lower than those obtained when using RRM.

- One question was: 'Has anyone offered you illicit drugs in the past year?' The observed prevalence based on the direct method was 10 %, whereas it was 15.8 % using RRM.
- Another question was: 'Have you used cannabis or derivatives in the past year?' The observed prevalence was 5.3 % with the direct method and 24.6 % with RRM.
- A further interesting question was: 'Have you used amphetamines or similar drugs in the past year?' The observed prevalences were 1.6 % and 5.9 %, respectively.

The procedure was also applied by the city of Verona, Italy, in 1981 and again at the end of the 1980s to estimate the prevalence of illicit drug use, of drug-offering near schools, and of other types of behaviour linked to the illicit drug market and to drug use among high-school students.

The three questions on drug use were as follows:

- 'Have you used any kind of drug in the past year?' (one-year prevalence);
- 'Have you used any hard drug (i.e. injected) in the past year?';
- 'Has anyone offered you illicit drugs in the past year?'

Panel analyses based on records from social agencies

Two similar panel analyses performed in Italy and Sweden are described and compared below.

As an example of an analysis based on panel data from social agencies, time trends of incarceration in Italy are examined, including an assessment of the impact on the Italian population and the possible influence of different laws on the frequency of detention of IDUs (Table 2). In brief, a 1990 law on drug abuse seems to have resulted in an

Table 2: Mean number of inmates, detention rate of IDUs and percentage of inmates represented by IDUs (IDU %) by year, Italy, 1986–95

Year	Mean no. of inmates	Detention rate (× 100 000)	IDU %
1986	33 609	59	18
1987	31 773	55	16
1988	31 382	55	24
1989	30 680	53	25
1990	25 931	45	28
1991	32 971	58	32
1992	45 538	80	32
1993	50 748	90	30
1994	52 410	92	29
1995	49 027	86	29

Source: Italian Ministry of Justice (1997)

increased detention rate of IDUs and an increased percentage of prisoners represented by IDUs (IDU %). A 1993 referendum that depenalised the personal possession of drugs also seems to have influenced the IDU %, apparently resulting in a decreasing trend. To test the hypothesis that the IDU % has significantly changed since 1993, a comparison can be made, using a standard hypothesis test for the means of non-paired data, between the average IDU % for 1991–92 and that for 1994–95. The comparison produces a u statistic ($u=10$) that is highly significant for both unilateral and bilateral tests compared to the standard normal variate. The same result can be obtained comparing the means before and after 1990. This, of course, does not mean that the 1990 law or the 1993 referendum had a causal effect on the trends of the phenomena analysed, but only that there is a significant statistical association. The study of the possible causality does not just involve statistical methods, which can only provide some clues and indications for investigating other means and information and a proper dynamic model of the entire process.

The second example of a panel analysis refers to incarceration in Sweden (Table 3). The IDU % shows an initial increase followed by a trend of stabilisation. Thus it does not seem to be influenced by the enactment of a new law on the coercive care of adult drug abusers in 1982, the criminalisation of drug use in 1988 or the introduction of a maximum sentence of six months' imprisonment for drug use in 1993. Only the difference between the percentages before and after 1982 is statistically significant.

A cross-sectional analysis: the impact of drug-related deaths in different countries

Deaths caused by drugs are often mistakenly considered to be one of the main indicators of the impact, whether positive or negative, of initiatives for controlling drug abuse.

Table 3: Mean number of inmates and IDUs and percentage of inmates represented by IDUs (IDU %) by year, Sweden, 1980–95

Year	Mean no. of inmates	Mean no. of IDU inmates	IDU %
1980	3 795	962	25
1981	4 116	1 168	28
1982	4 112	1 291	31
1983	4 280	1 487	35
1984	3 708	1 158	31
1985	3 617	1 336	37
1986	3 414	1 381	40
1987	3 792	1 509	40
1988	3 933	1 654	42
1989	3 944	1 653	42
1990	4 036	1 597	40
1991	3 803	1 586	42
1992	4 005	11 672	42
1993	4 290	1 742	41
1994	4 581	1 813	40
1995	4 408	1 773	40

Source: Swedish Council for Information on Alcohol and other Drugs (1996)

Since the data provided in Table 4 are not uniform, at least with regard to the year under study, comparisons can only be used to show how the methodology works and to provide examples of potential problems. It should also be noted that the definition of 'drug-related death', also referred to as 'substance-abuse-related mortality', differs by country and that it depends on the type of drug most commonly correlated with death. Some countries only consider overdoses as drug-related deaths, whereas others include drug-related accidents and deaths due to long-term abuse. Data-collection techniques may also vary and thus provide different results. When data were available from different national sources, mainly social agencies, even in the same country, discrepancies were detected in the reported numbers of drug-related deaths, as demonstrated, for example, in Italy (Conti et al., 1999).

The main problem with cross-sectional analyses is thus related to the non-comparability of country statistics, due to the use of different definitions and data-collection systems. These considerations should be taken into account when developing a model of health consequences of drug use and abuse or a model of drug-related mortality, especially when death statistics are used for the purposes of estimation, evaluation and monitoring.

To overcome this problem and to obtain an overall picture of the situation in various countries, it is possible to use ordinal indicators (grading) instead of numerical measures. This entails arranging the different countries according to indicators and using, for

Table 4: Number of drug-related deaths and per capita equivalent in various countries, 1991 and 1992

Country	Year	No. of drug-related deaths	Per capita equivalent (residents in thousands)
Australia	1992	820	0.046
Austria	1991	116	0.015
France	1991	411	0.007
Germany	1992	2 099	0.026
Italy	1992	1 217	0.021
Spain	1991	579	0.015
Sweden	1994	203	0.023
Switzerland	1991	405	0.060
UK	1992	1 421	0.024
USA	1992	5 601	0.022

Note: EU Member States are given in bold type
Sources: UNDCP (1994, 1997)

comparative purposes, a grading system, which is less affected by shifts and errors in the basic data (this procedure is commonly adopted for analysing the quality of life in various countries or cities).

A cross-sectional analysis: the impact of the HIV/AIDS epidemic among IDUs

The European Centre for the Epidemiological Monitoring of AIDS, Paris, collects quarterly individual data on AIDS cases notified in European countries. Table 5 shows the 1995 incidences per million IDUs.

The rankings can be used to perform cross-sectional analyses of the major health consequences of drug abuse, namely drug-related deaths and AIDS, by totalling the ranking figures in each country (Table 6).

In Switzerland, the ranking for drug-related deaths is equal to that of AIDS, and if considering the total of the two rankings, there are similarities between Austria and the United Kingdom and between France and Sweden. However, when considering the individual rankings in each country, there are considerable differences among the countries listed. These variations are due to the fact that the deaths related to injecting drug use, either directly or indirectly (i.e. those related to AIDS among IDUs and those strictly related to drugs), have different distributions in the different countries. It must be noted that since the two death processes involve the same population of IDUs engaging in risk behaviour, they represent competing risks. Thus the two figures are negatively correlated within each country and require a suitable mathematical model to be interpreted and used for making inferences. In any case, as explained below,

Table 5: AIDS incidence among IDUs and ranking in selected European countries, 1995

Country	AIDS incidence (× 1 000 000 IDUs)	Ranking
Austria	5	4.5
Denmark	6	6.0
France	22	7.0
Germany	3	2.5
Netherlands	5	4.5
Italy	61	10.0
Portugal	41	9.0
Spain	115	11.0
Sweden	3	2.5
Switzerland	32	8.0
UK	2	1.0

Source: European Centre for the Epidemiological Monitoring of AIDS (1996)

Table 6: Overall rankings of two major health consequences of drug abuse for selected European countries

Country	Ranking of drug-related deaths	Ranking of AIDS	Total
Austria	2.5	4.5	7.0
Denmark	–	6.0	–
France	1.0	7.0	8.0
Germany	7.0	2.5	9.5
Netherlands	–	4.5	–
Italy	4.0	10.0	14.0
Portugal	–	9.0	–
Spain	2.5	11.0	13.5
Sweden	5.0	2.5	7.5
Switzerland	8.0	8.0	16.0
UK	6.0	1.0	7.0

Note: – indicates data not available

particular care must be taken when using AIDS and death processes to make inferences about the hidden population of IDUs. In the presence of the AIDS epidemic among IDUs, some of the popular estimates obtained by simple multiplication methods on the basis of death statistics can be quite inconsistent. Thus, the two processes must be carefully modelled (proper point processes) in order to obtain useful and comparable estimates.

Multiple sources: improving inferences on drug-related phenomena

As mentioned above, one of the crucial aspects of monitoring drug-related phenomena is estimating the extent of drug use and abuse, which can only be done using indirect methods. Provided below is a description of how to use some of these methods, which depend on the type of data available from social-agency records and, in some cases, from surveys.

Estimates obtained by stratified calibration samples

This methodology can be used when a calibration population is available and the prevalence of the hidden attribute of interest (e.g. illicit drug use) can be observed or estimated among this population. For example, for a calibration sample stratified by variables such as gender, age and employment status, among others, with single strata denoted by S1, S2, ... Sk, the corresponding observed prevalences of the hidden attribute denoted by p1, p2, ... pk, and the sub-populations corresponding to the strata denoted by N1, N2, ... Nk individuals, the size of the hidden population can be estimated using the following equation:

$$Nh = N1p1 + N2p2 + ... + Nkpk$$

This method provides good estimates only if the strata and the sub-populations are homogeneous; in other words, the hypothesis that the data observed in the various sub-samples are randomly chosen (simple random-sampling model) must be assumed to be true. This is the crucial hypothesis of this method and it should always be verified before the method is applied.

Estimates obtained by probabilistic modelling

As pointed out, it is crucial to develop suitable basic hypotheses on the data-production process and on the structure and properties of the available sample data. It is also important properly to integrate data and information from different sources in order to obtain better estimates of the size of the hidden populations.

In the following examples, data from different sources are used to make inferences about the prevalence of 'heavy' (HU) and 'non-heavy' (NHU) drug use in different countries. 'Heavy use' is defined as daily or injecting use.

Estimating prevalence using data modelled by truncated Poisson processes

Table 7 presents data on the total number of drug users notified to the different social agencies in Sweden during 1992, disaggregated with respect to the number of notifications per person during the period and with respect to the type of drug use (i.e. heavy use and non-heavy use).

Table 7: Number of drug users notified to different social agencies in Sweden, 1992

Number of notifications	HU		NHU		Total
	Number	%	Number	%	
1	3 675	40	8 088	77.86	11 763
2	2 207	24	1 709	16.45	3 916
3	1 390	15	414	3.99	1 804
4	788	9	135	1.31	923
5	492	5	34	0.33	526
6	250	3	4	0.04	254
>6	301	3	4	0.04	315
TOTAL	9 103	100	10 388	100.00	19 491

Note: data disaggregated with respect to the number of times a single person was notified during the period and with respect to type of use (HU or NHU)
Source: Olsson et al. (1992)

A simple mathematical model that could be used for heavy and non-heavy users is the truncated Poisson model, approximating the binomial model, which allows for the calculation of the probabilities related to the number of notifications per person per year (see Figure 1 below), in other words, the probability that a single drug user, whether engaging in heavy use or not, is notified k times during a specific year (k=1,2, ...). On the basis of these probabilities, it is possible to predict the expected number of persons notified k times in a specific year and, by comparing these data with the actual observed data, it is possible to estimate the proportion of the population that is hidden (i.e. the size of the population which corresponds to zero notifications during that year).

The different estimates obtained for the parameters of the Poisson model using different pairs of k values and data from Table 7 indicate that once a person has been notified, he or she is more likely to be identified again than those who have never been notified. In other words, the intensity of the notification processes is lower for the first notifications, compared to successive ones, for several reasons:

- Once a person is known to be a drug user, it is easier to identify him or her again as a drug user.
- Certain drug users tend to engage in riskier behaviour with respect to the rest of the drug-using population and are easier targets for the notification process.
- Once individuals have been identified at least once, they may become less cautious and engage in riskier behaviour because they are aware of already being known as drug users.

These considerations are confirmed by the observation that the notification processes for non-heavy use have lower intensities than those for heavy use, since individuals practising NHU engage in less risky behaviour (i.e. they are not daily users or injectors).

Figure 1: Truncated Poisson model

The general formulae of the models are:

$$P(HU = k) = N \frac{n^k e^{-n}}{k!}$$

$$P(NHU = k) = M \frac{m^k e^{-m}}{k!}$$

where k refers to the number of notifications, as reported in the first column of Table 7, and n and m are the intensities of the notification processes per person, related respectively to HU and NHU. Both are unknown and must be estimated using the available data.

If the expected number of persons notified for a given value of k are considered, the following equation can be used:

$$E(HU = k) = N \frac{n^k e^{-n}}{k!}$$

$$P(NHU = k) = M \frac{m^k e^{-m}}{k!}$$

where N and M are the overall sizes of the hidden populations of drug users to be estimated (HU and NHU, respectively).

If the observed data are used as estimates of the expectations, suitable values can be found for the unknown quantities, and the fitting of the Poisson model to the observed data can be checked.

Thus the processes of notification are not homogeneous Poisson processes. Using some mathematics, appropriate generalisations of the Poisson process (Cox processes) can be applied to better model the notification processes for heavy and non-heavy use and to obtain estimates for the different intensities and population sizes. However, in the present case, it is possible to estimate at least the order of magnitude of the populations simply by using the mean values of the intensities, specifically, calculating the mean of the intensities obtained by different pairs of equations. This is done (excluding the first notification) using $k<7$ given that individuals notified a greater number of times ($k>6$) account for a small percentage of the cases, providing data that may be less reliable.

Using this simplified procedure, the following results are obtained:

$$n=2.6,\ N=19\ 000;\ m=0.7\ M=24\ 000$$

where n and m are the intensities (rate per person per year) of the notification processes per person for heavy and non-heavy use, respectively, and N and M are the overall sizes of the hidden populations of drug users (heavy and non-heavy use, respectively).

The estimated order of magnitude for N is totally consistent with the estimate obtained by Olsson et al. (1992) by using multiple capture–recapture methods.

Estimation methods based on multiple capture–recapture

The same individuals may appear on records from several social agencies. For instance, heavy heroin users may appear in the database of health-care services (HCS), in the database of the Prison Administrative Authority (PAA) and in the National Register of AIDS Cases (RAIDS).

If some of the registers refer to the same period of time (e.g. a specified quarter or semester), a simple multiple-classification model can be used to estimate the average size of the hidden population during the period, under the basic hypothesis that the registration processes are independent. For example, if there are simultaneous registrations in the databases of the PAA and RAIDS, it can be supposed that the basic hypothesis holds and the data can be reported using a summary table, such as that illustrated below (Figure 2).

Figure 2: Cross-tabulation of data from the PAA and RAIDS in the same period

		PAA		
		+	–	
RAIDS	+	N++	N+–	N+*
	–	N–+	N– –	N–*
	N*+	N*–	N	

Notes: N = the size of the hidden population (to be estimated)
N++ = the observed absolute frequency of the persons registered in both databases
N– = the size of the population not observed in either of the databases (unknown)
N+– = the observed absolute frequency of the persons registered only in RAIDS
N–+ = the observed absolute frequency of the persons registered only in the PAA
N–* = the size of the population not registered in RAIDS (unknown)
N+* = the observed absolute frequency of the persons registered in RAIDS
N*+ = the observed absolute frequency of the persons registered in the PAA
N*– = the size of the population not registered in the PAA (unknown)

If the hypothesis of independence between the registrations is satisfied, then the estimate of N simply results in

$$N=(N+*)(N*+)/(N++)$$

If the two sources are correlated, then estimates are obtained from above (negative correlation between the two sources) or from below (positive correlation).

The method can be easily generalised to more than two data sources and to dynamic situations and can be applied to data gathered by different social agencies in different countries. Provided below are two examples of how the method can be used.

Estimating the number of heavy heroin users in Italy

Figure 3 refers to data from the PAA and RAIDS in Italy.

Based on these data

$$N=218\,000$$

which is the estimated size of the population of heavy heroin users (i.e. the population exposed to the risk of being registered (captured) by both the PAA and RAIDS). However, this figure probably represents a low estimate, due to positive correlation between the two recording systems.

Figure 3: Cross-tabulation of observed data from the PAA and RAIDS, Italy, first half of 1996

		PAA		
		+	−	
RAIDS	+	106	1 594	1 700
	−	13 494	*	*
		13 600	*	*

Note: * indicates missing data
Sources: Italian Ministry of Justice (1997) and COA (1997)

Estimating the number of heavy drug users in the United Kingdom

The UK Home Office *Statistical bulletin* includes tables on different drug-related phenomena, in particular, data on notifications of heavy drug users and of drug-related deaths among heavy users. The two processes can reasonably be considered as independent and as affecting the same hidden population. Thus data can be classified according to Figure 2 above, resulting in Figure 4 below.

Straightforward calculations produce the following estimate:

$$N=136\,000$$

This estimate is quite consistent with those provided by official UK agencies and reported by the UNDCP (1997).

Figure 4: Cross-tabulation of observed data incidences, 1992

		Drug-related deaths		
		+	–	
Notifications	+	259	24 444	24 703
	–	1 162	*	*
	1 421	*	*	

Note: * indicates missing data
Source: Home Office (1994)

Estimating the prevalence of heavy heroin users involved in acquisitive crimes

Most heavy heroin users in Italy are involved in criminal activities to make money to buy drugs and, as in other countries, they mainly commit acquisitive crime (e.g. theft and robbery) or deal in drugs. Some of these individuals are identified, reported and prosecuted for such crimes and are thus included in crime statistics and in the database of the PAA.

According to PAA statistics for the second half of 1995, the absolute frequency (Na) of imprisonment of IDUs for acquisitive crimes was 5 788 (the same order of magnitude for 1996). Crime statistics and data from a multipurpose investigation (Indagine multiscopo), reported by the Italian National Institute of Statistics (ISTAT, 1993), provide the following additional information, which allows for calibration:

- the proportion of persons involved in acquisitive crimes that are known to the police and for whom the criminal proceedings have begun is f=5 % (this figure has remained fairly constant in recent years);
- the proportion of acquisitive crimes that are reported to the police is about q=45%.

If all of the data and information from the two sources are pooled, the order of magnitude N of the hidden population of heavy drug users (mainly heroin users) involved in acquisitive crimes in the second half of 1995 can be estimated using the simple formula (homogeneous model):

$$N=Na/(fq)$$

The result obtained by straightforward calculation is:

$$N=257\ 244$$

However, this is presumably a high estimate because of the positive correlation of the events A={an IDU committing an acquisitive crime} and B={being reported to the police and prosecuted}, due to possible previous notifications.

The order of magnitude obtained is extremely consistent with the estimated size of the hidden population of heavy users, obtained previously on the basis of AIDS and PAA data. It must also be pointed out that the present estimation does not include IDUs who do not commit acquisitive crimes but who nonetheless deal in drugs. It also excludes those not involved in any criminal activity. Indeed, the number of drug users imprisoned for dealing in drugs in the same period is quite high: 6 353. Unfortunately, the same information cannot be used on the proportion of dealers who are reported and prosecuted. Thus data on imprisonment of drug dealers cannot be used as above to improve the estimates. It must also be taken into account that an unknown, and possibly high, proportion of IDUs are involved in both criminal activities.

All of the estimation methods reported above can also be used to estimate the time trends of problem drug use when applied in different time periods. For this purpose, more suitable dynamic models can be used, as shown below.

Estimating the prevalence of heavy heroin users using a Markov model

Markov and semi-Markov models can be used to make projections for various phenomena related to drug use. In particular, they can be used to model the spread of the HIV/AIDS epidemic among IDUs and to make inferences about the prevalence of heavy injecting drug users. In the following example, one of the 'simple' dynamic models for the HIV/AIDS epidemic, developed within the EU Concerted action (PL 931723) multinational scenario analysis concerning epidemiological, social and economic aspects of HIV-AIDS on society, is used to estimate the trend of this prevalence in Italy in the period 1990–96, using data from the Italian Ministry of Justice and the PAA. The relationship between the two competitive processes of death (i.e. deaths directly related to drugs and deaths from AIDS) is also addressed (for details on the model and the estimation methods, see Rossi and Schinaia, 1998, and Pasqualucci et al., 1998).

The rationale behind this kind of simple dynamic model is:

- to use only macro parameters that can be mostly estimated using epidemiological data and studies reported in the literature in order to obtain robust forecasting and estimation results;
- to generate projections for available data (in the present case, AIDS-incidence data and mortality data, reported quarterly) also taking into account the processes of under-reporting and reporting delay that affect actual observed data;
- to base the model on only a few internal parameters that mirror different kinds of behaviour affecting the infectivity process, to be estimated by comparing projections and observed data, using suitable statistical methods (e.g. maximum likelihood and minimum χ^2).

The rationale for using such models to make inferences on the dynamics of the IDU is:

- to use external information to model the relations between injecting drug use and HIV infection, in order to estimate the prevalence of injecting use on the basis of the evaluated HIV/AIDS prevalence;
- to compare the results with those obtained by other sources or methods, in order to validate the estimates and to increase accuracy and robustness.

Some similar ideas can be found in Kaplan and Soloshatz (1993).

As an example, the case of an epidemic model which provides the overall prevalence of HIV and AIDS cases separately can be used. By using external information on the proportion of IDUs among HIV-infected individuals, denoted by P(IDU/HIV)=p(HIV) and provided by the HIV surveillance system, and the proportion of HIV-infected individuals among IDUs, denoted by P(HIV/IDU)=q(HIV) and provided by various sources (sample surveys among incarcerated IDUs or IDUs in treatment), the model can provide an estimate of the total number of IDUs, denoted by N(IDU), using the following simple calibration formula based on the total number of HIV-infected individuals, denoted by N(HIV), projected by whichever epidemic model is used:

$$N(IDU) = \frac{N(HIV)p(HIV)}{q(HIV)}$$

Similarly, AIDS prevalence can be used in place of HIV prevalence:

$$N(IDU) = \frac{N(AIDS)p(AIDS)}{q(AIDS)}$$

Unfortunately, the denominators q(HIV) (i.e. the estimated proportion of HIV-infected individuals among IDUs) and q(AIDS) are not always available from routine statistics. For example, neither of the sources available in Italy (i.e. sample surveys of IDUs in the health-care services and surveys of those in prison) distinguish between cases of HIV and AIDS in their official published reports. Thus, only the denominator q(HIV or AIDS) can be used, modifying the above formula as follows:

$$N(IDU) = \frac{N(HIV)p(HIV) + N(AIDS)p(AIDS)}{q(HIV\ or\ AIDS)}$$

In simple terms, the formula divides the current numbers of HIV-infected life-time IDUs, obtained by an epidemic model of the HIV/AIDS epidemic from AIDS cases, by the estimated current rate of infection in IDUs. It must be noted that the results of the estimation method are based only on the projections of the absolute prevalence (number of cases) of HIV and AIDS, obtained by whichever epidemic model is used (numerator of the calibration formula), and on the estimate of the proportion of individuals with HIV or AIDS among IDUs (denominator of the calibration formula), coming from official statistics, which are easily obtainable. It must be noted that the results are sensitive to uncertainties both in the numerator and the denominator, and these uncertainties become particularly important when present in the denominator. For example, if the estimated denominator is 2 % and the country's actual denominator is 4 %, the error of

the estimate of N(IDU) is 100 %. Thus the accuracy of the estimated denominator is important for obtaining good results, especially for those countries where the HIV/AIDS epidemic is small.

The calibration formula can be applied either at national level, when national estimates of the denominator are available, or at local level, when only local estimates are available (the numerator can always be estimated at local or national level). The local estimates of N(IDU) can be used as anchor points for other methods for obtaining national estimates. This possibility is particularly interesting for those countries where local prevalence estimates of problem drug use are not available or where obtaining these estimates is problematic or costly. The results obtained by the use of the Mover-Stayer model on Italian data are reported in Table 8 together with the analogous estimates obtained by the multiple capture–recapture method (Rossi, 1999b). These estimates are quite consistent with other previously reported estimates.

The population of IDUs was estimated to have increased from the beginning of the 1990s until 1995, followed by a decreasing or stabilising trend which coincided with an increase in reported drug-related deaths. This effect may be mostly due to the fact that the competing risk of death related to the AIDS epidemic has been decreasing in recent years among IDUs, with a consequent increase in the risk of dying from drug-related causes. However, it must also be considered that there is a delay in the process that generates drug-related deaths from drug addiction. In fact, data from harm-reduction programmes and hospital admissions for non-fatal overdoses indicate that for each death from overdose there are approximately ten non-fatal overdoses. Thus, to model the mortality process, non fatal-overdoses should also be considered and such events, which have an intensity ten times that of drug mortality. These cause a time lag between the onset of injecting drug use and the death process (UNDCP, 1997). Therefore, the recent increase in the number of drug-related deaths observed in Italy has probably been generated by the increase in injecting drug use in previous years. To further study

Table 8: Number of IDUs in Italy estimated using the Mover -Stayer model for the HIV/AIDS epidemic and multiple capture–recapture model, and the number of drug-related deaths by year, 1990–96

Year	IDUs (MS)	IDUs (CC)	DRD
1990	180 000	140 000	1 161
1991	211 000	190 000	1 383
1992	224 000	230 000	1 217
1993	244 000	190 000	888
1994	279 000	225 000	867
1995	310 000	235 000	1 195
1996	297 000	220 000	1 564

Notes: MS = Mover-Stayer model
CC = capture–recapture
DRD = drug-related deaths

this phenomenon, the incidence curve of injecting drug use should be estimated, since the prevalence curve is not a sensitive enough indicator.

These considerations confirm that the number of deaths caused by drugs cannot be used as one of the main indicators of the impact of policies developed over the years or across different countries. They also demonstrate the hazards of estimating the prevalence of injecting use by calibration or multiplication methods based solely on death statistics.

Other health consequences, such as hepatitis B and C, can also be used as indicators. However, the problem with these indicators, at least for Italy, is that notifications of these infections is not compulsory.

It is possible to compare the trends of the estimated prevalences of IDUs with other kinds of indicators related to the hidden population, such as the incidence and prevalence of IDUs in public or private health-care services, the incidence and prevalence of IDUs in jail, and statistics on heroin seizures. First, it is possible to calculate the mutual correlation coefficients between couples of indicators. These coefficients can be used to determine whether the two indicators considered tend to show similar behaviours (both of them increasing or both decreasing) or different ones (one increasing and the other decreasing). In particular, when the correlation coefficient is close to 1, there is a strong tendency for the two indicators to display similar behaviour, whereas when the coefficient is close to –1, different behaviour is likely. To facilitate the examination of such mutual coefficients, they can be arranged in a table known as a 'correlation matrix'. Table 9 is a correlation

Table 9: Correlation matrix related to selected variables (PRCOM), 1990–96

	PRPHC	INPHC	INCOM	PRPRI	INPRI	SEIZK	IDU (MS)	IDU (CC)	PRCOM
PRPHC	1.00	0.97	0.59	0.86	0.50	0.22	0.94	0.87	0.32
INPHC		1.00	0.72	0.88	0.55	0.22	0.95	0.91	0.59
INCOM			1.00	0.83	0.88	0.21	0.50	0.79	0.90
PRPRI				1.00	0.83	0.08	0.75	0.84	0.74
INPRI					1.00	0.19	0.29	0.65	0.91
SEIZK						1.00	0.11	0.49	0.33
IDU(MS)							1.00	0.79	0.30
IDU(CC)								1.00	0.72
PRCOM									1.00

Notes: PRPHC = prevalence of IDUs in public health-care services
INPHC = incidence of IDUs in public health-care services
INCOM = incidence of IDUs in private health-care services
PRPI = prevalence of IDUs in prison
INPRI = incidence of IDUs in prison
SEIZK = seizures (in kg)
IDU (MS) = injecting drug users using the Mover-Stayer model
IDU (CC) = injecting drug users using the multiple capture–recapture model
PRCOM = prevalence of IDUs in private health-care services

matrix that allows a straightforward comparison to be made among the estimated prevalences of injecting drug use and the other indicators of interest.

All of the values of the correlation coefficients between the statistics directly related to the extent of injecting heroin use and the prevalence estimates are significantly high, except those related to seizure statistics, which strongly depend on repression policies and the type of police interventions. These high values of the correlation constitute an indirect validation of the estimation procedures used. This confirms the agreement between some important indicators of drug-related phenomena and the prevalence estimates reported above.

Conclusions

The main purpose of this chapter was to demonstrate the use of available routine data and of dynamic models for estimating the number of problem drug users. With respect to other methods which can be applied to estimate the same hidden population, the models presented here allow for the use of data that are generally available in most Western countries.

In any case, due to the various uncertainties, related both to data and models, several models and methods should be applied whenever possible in order to obtain the best description of the hidden population of interest. In the previous sections, this joint use of different methods allowed some conclusions to be drawn, at least regarding trends and the order of magnitude of the population.

The observed agreement in the prevalence trends estimated by different methods allows some indicators to be used for monitoring different control interventions, at least from a qualitative point of view. It must again be stressed that better analysis could be performed to monitor interventions on the basis of the incidence curve of injecting drug use.

References

Centro Operativo AIDS (COA), Istituto Superiore di Sanità (1997). *Sindrome da Immunodeficienza acquisita in Italia, aggiornamento dei casi notificati al 31 dicembre 1996*. Rome: COA.

Conti, S., Farchi, G., Masocco, M., Rezza, G., Toccaceli, V., Rossi, C. (1999) 'Mortality from overdose in Italy: 1984–1993'. *Statistica applicata* 11: 515–23.

European Centre for the Epidemiological Monitoring of AIDS (1996). *HIV/AIDS surveillance in Europe 1994–1996*. Paris: European Centre for the Epidemiological Monitoring of AIDS.

Home Office (1994). 'Home Office Research and Statistics Department "Statistics of drug addicts notified to the Home Office", United Kingdom, 1993'. *Statistical Bulletin* 10.

Istituto Nazionale di Statistica (ISTAT) (1993). 'I fatti delittuosi: caratteristiche degli eventi e delle vittime'. *Indagine multiscopo sulle famiglie, anni 1987–91* 5.

Ministero di Grazia e Giustizia, Dipartimento per l'amministrazione penitenziaria, Ufficio studi, ricerche, legislazione e automazione, Divisione II – Statistica (1997). *Indagine nazionale sui soggetti tossicodipendenti e affetti da virus HIV negli istituti di pena al 31/12/96*. Rome: Italian Ministry of Justice.

Kaplan, E., Soloshatz, D. (1993). 'How many drug injectors are there in New Haven? Answers from AIDS data'. *Mathematical and Computer Modelling* 17: 109–15.

National Institute on Drug Abuse (NIDA) (1990). *Collection and interpretation of data from hidden populations*. Research Monograph No. 98. Rockville, MD: NIDA.

Olsson, O., Byqvist, S., Gomer, G. (1992). *Det tunga narkotika-missbrikets omfattning i Sverige 1992*. Report no. 28. Stockholm: Swedish Council for Information on Alcohol and Other Drugs (CAN).

Pasqualucci, C., Ravà, L., Rossi, C., Schinaia, G. (1998). 'Estimating the size of the HIV/AIDS epidemic: complementary use of empirical bayesian back calculation and the mover–stayer model for gathering the largest amount of information'. *Simulation* 71 (4): 213–27.

Rossi, C. (1999a). 'Monitoring drug control strategies: hidden phenomena, observable events, observable times'. *International Journal of Drug Policy* 10: 131–44.

Rossi, C. (1999b). 'Estimating the prevalence of injecting drug users on the basis of Markov models of the HIV/AIDS epidemic: applications to Italian data'. *Health Care Management Science* 2: 173–9.

Rossi, C., Schinaia, G. (1998). 'The mover stayer model for HIV/AIDS epidemic in action'. *Interfaces* 28 (3): 127–43.

Stovali, M. (1993). 'Un esempio di applicazione della tecnica delle risposte casualizzate'. *Induzioni* 7: 49–57.

Swedish Council for Information on Alcohol and Other Drugs (CAN) (1996). *On the alcohol and drug situation in Sweden*. Report no. 96. Stockholm: CAN.

United Nations International Drug Control Programme (UNDCP) (1997). *World drug report*. Oxford: Oxford University Press.

United Nations International Drug Control Programme (UNDCP) (1994). *Substance abuse related mortality: a worldwide review*. Vienna: UNDCP.

CHAPTER 5

MULTIPLE-INDICATOR MODELS

Roberto Mollica and Fabio Mariani

Dynamic multiple-indicator models start from the choice of indicators and data available. However, drug-related problems are related to the context in which they occur. Dynamic evolution of drug use, markets and consumption may vary among countries, and in each country among different local and geographic areas, such as rural, suburban or urban sites. For the rest of this chapter, the geographic area under study is referred to as the 'location'. The target and risk-exposed population is only a part of the total population of the considered geographic area. Law enforcement and penalties, addiction treatment and care availability, diseases related to illicit drug use and consequent social costs also vary across and within countries.

These general considerations show that working on a multiple-indicator model involves being aware of the overlapping aspects which contribute to defining the model and its outputs. In this chapter, a monitoring model is proposed which could predict changes in the main aspects of drug use and related diseases over time. To pursue this aim the use of cause–effect relationships among some indicators is hypothesised. Because of the delay between the effect to be observed and the factors which cause that effect, a dynamic model capable of providing predictions can be created. Starting from these concepts, the model building and its difficulties are discussed in turn.

As known, every 'location' will have its own approach to drug use and related problems. A common model which could be applied to different locations would not necessarily emerge from individual studies. Specific sets of indicators relevant to the study location may be developed. This would, however, complicate the data-collection requirements of such models. If comparisons are being made across large geographic areas, then there may be considerable differences between the models.

The other main characteristic of dynamic models is the collection of data over time. The chosen indicators have also to be real and timely to give up-to-date information.

How the model works

In this chapter, a simple but effective definition of indicators is proposed to allow the same collection for different populations and/or different 'locations' at different times.

To make model applicability easier, the following indicators are suggested as shared characteristics which can be collected in every context. They are classified into three

main groups: social (S); legal (L); and medical (M). It is then demonstrated how the defined indicators could be combined in a dynamic model.

Social indicators

This group involves the most meaningful parameters for social data and situations commonly observed in an addict's life. They are considered as risk factors and can also be used as predictor factors. The data for these indicators can be acquired from national data sources on the general population and can be considered as indicators of social disruption.

- *S1 disaggregate familiar nuclei*: separation and divorce can be linked to this item as indirect demonstrations of life stress.
- *S2 total unemployed people*: this reflects social general entity of welfare. Lack of welfare generally leads to crime and illicit activities.
- *S3 interruption after primary school*: this is akin to a condition that reflects an early predisposition to anti-social behaviour.

Legal indicators

This second group of indicators includes detected offences related to drug problems. Such indicators may be very sensitive to context changes like the spread of new drugs and new kinds of addiction. It may be possible to use such indicators as early-warning measures of emerging trends.

- *L1 drug-related law offences*: easy and simple evidence about the size and characteristics of the black market and major crimes concerning drug use and problems. In the case of seizures, the type of drug seized may be indicative of market/consumer orientation.
- *L2 minor offences*: generally these indicate how great is the use of illicit drugs as street addicts are involved in this kind of offence to maintain their drug needs.
- *L3 under-age arrests*: this item could address the 'at risk' population which will still be present in the future.

Medical indicators

While the social and legal indicators are probably easy to acquire from national and/or local reports, not all locations may currently have the proposed health and treatment data. However, it is hypothesised that these would match the aims of developing dynamic models.

- *M1 treated drug users*: in particular, this item is valid in the locations where treatments are offered to addicts. Some kinds of treatment define the population, as methadone treatment does for heroin addicts. For these reasons, the variable could not be considered for the analysis of some drug issues.

- *M2 drug-related deaths*: this includes not only overdoses, but also crime-related or violent deaths that can be linked to drug involvement (car crashes). Such data may be available from coroners' reports or autopsy evidence.
- *M3 emergency room*: all emergency-room admittances for suspected drug use could be detected and recorded. It may be useful to separate patient's demands for treatment (if available) or police involvement (if compulsory).
- *M4 AIDS cases related to intravenous drug use (IDU)*: This indicator could involve only the numbers at a late stage of HIV infection. While HIV prevalence and incidence estimates now have greater accuracy, the numbers are not easily routinely recorded. Moreover, new therapies for HIV could delay progression to AIDS. The number of new subjects who have undergone anti-HIV treatment could be a good series for this indicator, but these data may not be easy to collect.

The variables listed above represent the main and minimal list that allow the development of a proposal for a dynamic model. Most of these data were reported in EMCDDA (1995). This suggests that such data are of great importance in drug-addiction monitoring and they are studied and surveyed as being discriminate entities and indicators for risk conditions. Because in some cases the same variable is investigated and collected using different tools or times, the difficulties of matching information among European countries could be relevant to the discussion about developing these models.

Each item can be observed at different times and they can all be clustered at the same time (Figure 1). These clusters represent single-point-in-time observation about real situations and risk conditions at that time. Consecutive observations indicate the course of variables, and in this way some correlation can be found among them. It is interesting looking for correlations between risk factors (for example all S and L3 variables) and point effects (for example M and other L items) to create the model. In fact, if risk factors show a correlation with the next-point-in-time evidence with a certain delay, the first data points of risk factors could be used to predict indicators of problems for the following period (Figure 2). It is obvious that variables interact with each other at different strengths and that a coefficient that allows adjustments between indicators using a mathematical model is therefore required. Moreover, by developing the model in subsequent periods it would be possible to obtain greater and greater accuracy as

Figure 1: Example of clustered observations

Timed observations	S1	S2	S3	L1	L2	L3	M1	M2	M3	M4
T1	1.3	15.6	4.8	1 300	1 2650	560	30 000	1 200	5 940	4 896
T2	2.3	17.5	4.6	1 650	1 6580	467	40 000	1 060	6 860	5 691
T3	1.8	19.5	5.1	1 540	1 5640	592	45 000	950	5 620	5 979

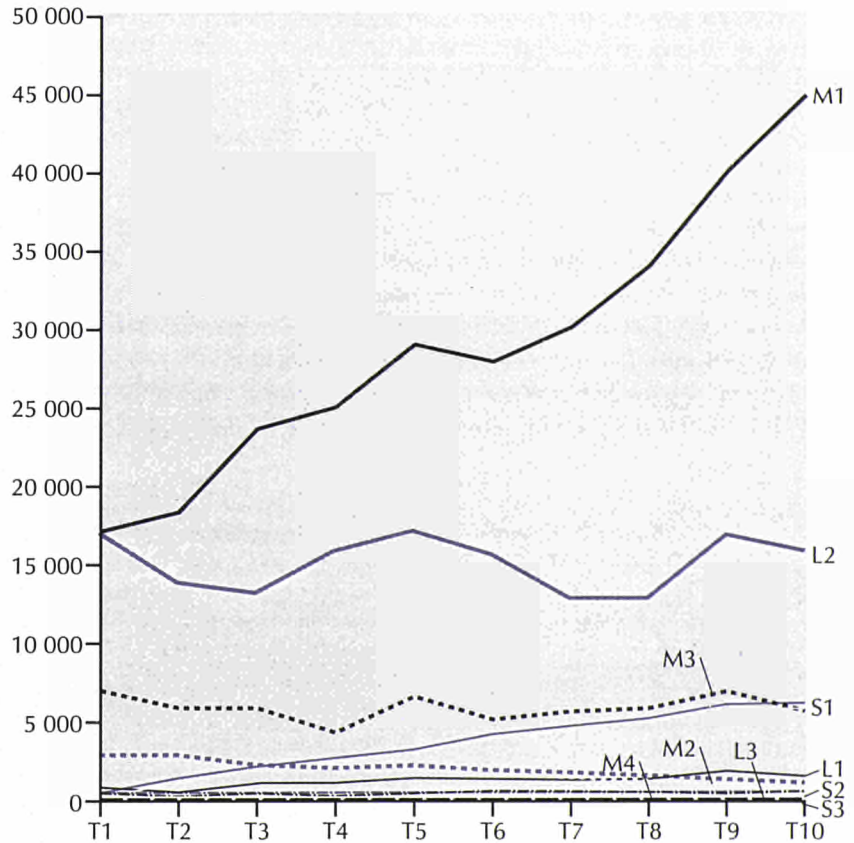
Cluster

Note: data are indicative but not real

the historical list of observations grows. For this reason, such a model can be used and developed with limited data sets and then amended over time.

Theoretically, such a model could be used to investigate a specific event, or drug-related problem, as the outcome of the modelling process. The application of such a model may vary on the basis of specific need. If, for example, someone wants to forecast the number of emergency-room admittances in order to reinforce staff with a toxicologist, the dependent variable will be emergency-room admittance. All other variables contribute to the model, as independent factors, in generating the outcome needed. However, the variables or indicators would have a different impact in other examples. Starting from the risk indicators, the at-risk population and the number potentially involved in drug use could be predicted, but some other events (the availability of new out-patient treatment resources, for example) could interact with this prediction. This is because the presence of any kind of treatment protects addicts from fatal accidents. A similar interaction of indicators in a model is constructed to estimate the number of

Figure 2: Example of trends in variables



drug-related deaths. The introduction of methadone treatment protects from street opiate overdoses but not deaths from car crashes or those due to the abuse of other substances. On the basis of their involvement in the model, variables can be independent factors or co-factors and, moreover, both multiple correlation and multiple collinearity may apply in the model. Another example concerns predicting minor offences. The number of minor offences recorded is related to the number of people involved in illicit drug use, especially the number of street users, to the rate of unemployment and to the presence/absence of treatment units. Moreover, the model could address the effectiveness and efficiency of new laws regarding addicts or new therapeutic approaches.

Two Italian experiences illustrate the potential. The first example is from a study in Sardinia into methadone take-home prohibited by a new law at the beginning of the 1990s. It was seen from a time series of indicators that the application of the new law resulted in a dramatic reduction in the treated population. In fact, many addicts preferred to quit treatment hoping to abstain from any drug use rather than lose the job they were able to preserve or had found during treatment. The second example is from Milan and was a study performed by one of the authors of this paper. It was seen that low-threshold methadone intervention at a public out-patient unit located in city suburbs resulted in a growth in the treated population (ten-fold in three years), a reduction in unemployment rates, low HIV incidence and improvements in welfare through monitoring indicators.

The model suggested shows flexibility and can be personalised to suit specific needs. For example (how many people will attend treatment units...) in 'locations' (...in Italy...) over time (...in next two years?).

Some weak points to these data-driven models must, however, be underlined. For example, people presenting risk factors could be already involved in drug problems which cannot be recorded at that time. They can only be recorded when there is the evidence of involvement and when variables (estimates at a point in time) demonstrate what is the known situation of the drug use. The sense of the model, however, is not to fit a drug-addict population estimate, even if a capture–recapture model could implement these data, but rather to investigate dynamic relationships.

Other issues emerge if a specific population, such as those aged 15 to 44, is considered, and if how great is the phenomenon to be observed is defined. A selected population will have limited variations over a long time and small average changes from one time period to the next. Absolute observation could be recorded and used in the model, but this would not reflect the prevalence rate.

Another bias is delay between causes and effects. Time intervals may vary in these models from months to years. In a study of a population of addicts, it took an average of ten years (with low standard deviation) between the onset of addiction and enrolment in a therapeutic programme. This means that the course of addiction follows some general rules: onset during adolescence and youth; continuous use until young adult age; and treatment attempts. The first period corresponds to the S1, S3 and L3 variables, the second to S2, L1, L2, M2 and M3, and the last one to M1 and M4. It could not be

the same for other countries where prevention strategies allow early admittance to treatment or where law-enforcement agencies impose increasingly severe penalties or, again, where HIV/AIDS has a low impact on social health (for example the north to south gradient of prevalence and incidence in Italy). As generally a small 'location' will keep the same characteristics over several years, it is possible to predict the future situation of a requested target outcome. If the number of minor offences or unemployment are found to follow primary-school interruption over a one-year period it will then be possible to hypothesise the importance of next year's minor offences or unemployment predicted by other co-factors or collinearities if the number of last year's school interruptions is known. As has been reported, the ten years' gap between illicit drug-use onset and admittance to treatment could give an accurate idea of how many people will be treated within the coming years. This would allow for planning the development of new units or, conversely, the closure of opened ones.

The dynamic model controls itself every time data are updated. Unexpected results can be used to allow the model to verify its accuracy and correct contingent trends. As observations can be scheduled differently (every month, every year or every week), it is also possible to develop a short-, medium- or long-period forecast when cause–effect relationships are linked by different time periods. The figures below explain model construction and development with hypothetical data. Each cell represents a value, each column a variable and each row a cluster. In Figure 3, a 15-period observation is simulated and each cell is labelled by variable name and registration time to make it recognisable at the next step.

In Figure 4, the delay periods which link the cause–effect relationships among the variables are hypothesised.

At this point there are both tabled data and delay times. Moving columns (variables) by the time of their delay gives a new figure (Figure 5); each row shows the situation arranged by variable time links, and new clusters are now generated. Linked clusters would match the real temporal relationships between variables.

Blue labels are data which could be estimated if variables show correlations with each other. In other words, these are the target indicators the dynamic multiple-indicator model is intended to predict. Simple bivariate correlations allow a first indication about how variables are coupled and related, and it could be the first meaningful target to reach. To understand how much a variable is influenced by others, regression models have to be applied. The evidence gained in such analysis gives the strength of the relationship between the independent factors and the dependent one. If any relationship is observed, blue-labelled cells can be estimated as there is a forecast for each of them. In this hypothetical example, most of the variables have long delay periods in reference to M1 and M4 factors making possible a long-term forecast about these two variables which are 'well controlled' by real data; M1/T25 cell hypothesis will be less precise because they depend on other derived values (all blue ones).

Figure 3: Data table

Clusters	Variables									
	S1 T1	S2 T1	S3 T1	L1 T1	L2 T1	L3 T1	M1 T1	M2 T1	M3 T1	M4 T1
	S1 T2	S2 T2	S3 T2	L1 T2	L2 T2	L3 T2	M1 T2	M2 T2	M3 T2	M4 T2
	S1 T3	S2 T3	S3 T3	L1 T3	L2 T3	L3 T3	M1 T3	M2 T3	M3 T3	M4 T3
	S1 T4	S2 T4	S3 T4	L1 T4	L2 T4	L3 T4	M1 T4	M2 T4	M3 T4	M4 T4
	S1 T5	S2 T5	S3 T5	L1 T5	L2 T5	L3 T5	M1 T5	M2 T5	M3 T5	M4 T5
	S1 T6	S2 T6	S3 T6	L1 T6	L2 T6	L3 T6	M1 T6	M2 T6	M3 T6	M4 T6
	S1 T7	S2 T7	S3 T7	L1 T7	L2 T7	L3 T7	M1 T7	M2 T7	M3 T7	M4 T7
	S1 T8	S2 T8	S3 T8	L1 T8	L2 T8	L3 T8	M1 T8	M2 T8	M3 T8	M4 T8
	S1 T9	S2 T9	S3 T9	L1 T9	L2 T9	L3 T9	M1 T9	M2 T9	M3 T9	M4 T9
	S1 T10	S2 T10	S3 T10	L1 T10	L2 T10	L3 T10	M1 T10	M2 T10	M3 T10	M4 T10
	S1 T11	S2 T11	S3 T11	L1 T11	L2 T11	L3 T11	M1 T11	M2 T11	M3 T11	M4 T11
	S1 T12	S2 T12	S3 T12	L1 T12	L2 T12	L3 T12	M1 T12	M2 T12	M3 T12	M4 T12
	S1 T13	S2 T13	S3 T13	L1 T13	L2 T13	L3 T13	M1 T13	M2 T13	M3 T13	M4 T13
	S1 T14	S2 T14	S3 T14	L1 T14	L2 T14	L3 T14	M1 T14	M2 T14	M3 T14	M4 T14
	S1 T15	S2 T15	S3 T15	L1 T15	L2 T15	L3 T15	M1 T15	M2 T15	M3 T15	M4 T15

Conclusions

The applicability of dynamic multiple-indicator models depends on the relationships between the indicators. During practical experience of such models, unexpected and unsuspected effects were found, and sometimes it is not so easy to justify these results. They could be caused by a false response or a new phenomenon. Using the same process it is possible to perform retrospective studies on selected populations when explanation of an event is needed or is not completely understood. The study of low-threshold treatment with the aim of harm reduction was an interesting example. It is generally believed that such interventions are targeted at compromised and marginalised people, but the results showed that an unsuspected population of employed, highly educated and HIV-negative individuals were contacted and recruited. The results from the indicator model showed that a still-not compromised population susceptible to harm-reduction programmes had been captured by this programme.

Figure 4: Delay periods

	S1	S2	S3	L1	L2	L3	M1	M2	M3
M4	9	9	9	7	6	7	2	7	7
M3	2	2	2	0	1	0	9	0	
M2	2	2	2	0	1	0	9		
M1	11	11	11	9	8	9			
L3	2	2	2	0	1				
L2	3	3	3	1					
L1	2	2	2						
S3	0	0							
S2	0								

As described, the main features of the model outlined in this chapter can address multiple uses. First, it has potential as a predictor tool to estimate the future development of specific problems in different fields of addiction. Alternatively, it could be used to adjust hospital resources if surveillance methods suggest that these interventions are needed. The second use is to verify, retrospectively, any expected or unexpected consequence, its relationship with the original target and to check model adequacy for future forecasting. This last use suggests an updating system calibrating the model for the next estimation and adjusting the model itself. A third use is to study the evidence of cause–effect relationships between legal or government or local logistic interventions and drug demand or social-health consequences.

Both European countries and the European Monitoring Centre for Drugs and Drug Addiction would be interested to know in advance what will be the impact of drug and drug-related problems and this type of model could give some important information in this regard. Model development and shared data-collection methods would permit comparison among countries. The variables suggested in this chapter are present in all countries. The complexity of the task could generate difficulties in model use because relationships between variables vary across 'locations', and so an analysis of historical data is needed to generate equations which consider multiple correlation, multicollinearity, regression and correction coefficients. Even if the model will correct itself over time, some minimal previous observations are needed to address the starting point of the model estimation. This means that, in spite of the easy accessibility and recovery of data, model development must be preceded by a data-collection system.

Figure 5: Data arranged by delay periods

Linked clusters	Variables									
							M1 T1			
							M1 T2			
							M1 T3			M4 T1
							M1 T4			M4 T2
							M1 T5			M4 T3
							M1 T6			M4 T4
							M1 T7			M4 T5
							M1 T8			M4 T6
					L2 T1		M1 T9			M4 T7
				L1 T1	L2 T2	L3 T1	M1 T10	M2 T1	M3 T1	M4 T8
				L1 T2	L2 T3	L3 T2	M1 T11	M2 T2	M3 T2	M4 T9
	S1 T1	S2 T1	S3 T1	L1 T3	L2 T4	L3 T3	M1 T12	M2 T3	M3 T3	M4 T10
	S1 T2	S2 T2	S3 T2	L1 T4	L2 T5	L3 T4	M1 T13	M2 T4	M3 T4	M4 T11
	S1 T3	S2 T3	S3 T3	L1 T5	L2 T6	L3 T5	M1 T14	M2 T5	M3 T5	M4 T12
	S1 T4	S2 T4	S3 T4	L1 T6	L2 T7	L3 T6	M1 T15	M2 T6	M3 T6	M4 T13
	S1 T5	S2 T5	S3 T5	L1 T7	L2 T8	L3 T7	M1 T16	M2 T7	M3 T7	M4 T14
	S1 T6	S2 T6	S3 T6	L1 T8	L2 T9	L3 T8	M1 T17	M2 T8	M3 T8	M4 T15
	S1 T7	S2 T7	S3 T7	L1 T9	L2 T10	L3 T9	M1 T18	M2 T9	M3 T9	M4 T16
	S1 T8	S2 T8	S3 T8	L1 T10	L2 T11	L3 T10	M1 T19	M2 T10	M3 T10	M4 T17
	S1 T9	S2 T9	S3 T9	L1 T11	L2 T12	L3 T11	M1 T20	M2 T11	M3 T11	M4 T18
	S1 T10	S2 T10	S3 T10	L1 T12	L2 T13	L3 T12	M1 T21	M2 T12	M3 T12	M4 T19
	S1 T11	S2 T11	S3 T11	L1 T13	L2 T14	L3 T13	M1 T22	M2 T13	M3 T13	M4 T20
	S1 T12	S2 T12	S3 T12	L1 T14	L2 T15	L3 T14	M1 T23	M2 T14	M3 T14	M4 T21
	S1 T13	S2 T13	S3 T13	L1 T15	L2 T16	L3 T15	M1 T24	M2 T15	M3 T15	M4 T22
	S1 T14	S2 T14	S3 T14	L1 T16	L2 T17	L3 T16	M1 T25	M2 T16	M3 T16	M4 T23
	S1 T15	S2 T15	S3 T15	L1 T17	L2 T18	L3 T17	M1 T26	M2 T17	M3 T17	M4 T24

Finally, developing these types of models involves a shared and simple monitoring activity with predictor characteristics. Many institutions involved in drug problems would like to know in advance how many people will use illicit drugs or will be HIV positive in the AIDS phase or will be exposed to risk of death or, again, will have legal problems as a drug consumer. Moreover, the model has built in the possibility of verifying its own forecasts and developments in model effectiveness and efficiency. The opportunity of retrospective clinical and social studies will allow the acquisition of experiences which lead to increasingly helpful suggestions to reduce drug demand.

References

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (1995). *Annual report on the state of the drugs problem in the European Union*. Lisbon: EMCDDA.