



UNIVERSITY
of
GLASGOW

Scientific Review of the Literature on Estimating the Prevalence of Drug Misuse on the Local Level

CT.97.EP.05

Centre for Drug Misuse Research:	Gordon Hay Neil McKeganey
EMCDDA:	Lucas Wiessing Richard Hartnoll
Other Contributors:	Pierre-Yves Bello Daniela D'Ippoliti Antònia Domingo-Salvany Martin Frischer Ludwig Kraus Filip Smit

Please use the following citation:

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA).
Scientific Review of the Literature on Estimating the Prevalence of Drug
Misuse on the Local Level. Lisbon: EMCDDA, July 1999.

Contact Details

Centre for Drug Misuse Research, University of Glasgow
Glasgow, G12 8QG
United Kingdom.

European Monitoring Centre for Drugs and Drug Addiction
Rua Cruz de Santa Apolónia 23/25
1100, Lisboa
Portugal.

Further copies of this report can be obtained from the EMCDDA at the above
address.

Acknowledgements

This publication is based on an EMCDDA-commissioned project which examined methods for estimating the prevalence of problem drug use at the local level. During this and a previous project, the authors relied on the support of researchers throughout the European Union and within the EMCDDA. It would not have been possible to conduct local prevalence estimation projects without the support of various agencies and bodies which supplied data on drug use in the participating cities. The authors would like to acknowledge the following individuals who contributed to these projects:

Catherine Comiskey, Sofia Friere, Maria Moreira, Alvar Norén, Päivi Partanen, Kathy Politikou, Dan Seidler, Jaap Toet, Alfred Uhl and Ari Virtanen.

Contents

	Page No
1	Introduction 1
2	Direct Methods 2
2.1	Surveys 2
2.2	Enumeration..... 3
2.3	Case Finding 3
3	Indirect Methods 5
3.1	Multiplier Methods 5
3.2	Nomination Techniques 6
3.3	Synthetic Estimation 7
4	Capture-recapture Methods 9
4.1	Historical Development 9
4.2	Two-sample Methods 10
4.3	Three-sample Methods 13
4.4	One-sample Methods 14
4.5	Epidemiological & Drug Misuse Applications 16
5	Dynamic Models 23
5.1	Open Population Capture-recapture Methods 23
5.2	Latent State Markov Models 24
6	Synthesis 25
6.1	City Estimates 25
6.2	EMCDDA Comparative Study 41
6.3	Regional Estimates 42
6.4	Summary 45
7	Discussion and Conclusions 46
8	References 51

1 Introduction

There is a clear need for accurate and comparable information on the nature and extent of drug misuse at the local level. Epidemiology is concerned with patterns of disease occurrence in human population and although it would be wrong to simply classify drug misuse as a disease, there are a range of epidemiological methods that can be used to inform policy makers and practitioners. In this report, we review the scientific literature relating to the use of epidemiological methods for estimating the prevalence of drug use at the local level.

We begin by briefly discussing the use of surveys in estimating the prevalence of drug misuse. We then detail the literature on other direct methods of prevalence estimation, such as national registers or monitoring systems. Even when data can be obtained at the local level, these methods often offer little more than the number of people in treatment or who have been arrested for possession of drugs. We note how such data can be combined within a multi-source enumeration and discuss how case-finding studies have been useful in some areas.

We then introduce indirect methods of drug misuse prevalence estimation, beginning with multiplier methods. We describe how sampling techniques such as snowball sampling can be used to provide the necessary information for a multiplier study. We also describe how the multiplier method can be extended to synthetic estimation.

We then provide an in-depth discussion of a methodology known as the 'capture-recapture method'. We review the literature which charts the historical development of the methodology, both as an ecological tool and an epidemiological tool. We comment on some of the many papers which describe using the capture-recapture method in epidemiological or drug misuse studies, and describe a few of the key review papers in that field.

We next describe the literature on two more advanced statistical techniques; open-population capture-recapture models and latent state Markov models. The various methodologies are then compared within a discussion section.

Although a scientific review of the epidemiological tools used in estimating the prevalence of drug misuse is important, there is also a need for a synthesis of recent prevalence estimation research in Europe. We therefore summarise the results of local prevalence studies, comparing the different methods used and how various definitions affect the interpretation of the estimates.

2 Direct Methods

In this section we discuss direct methods for gathering information on the nature and extent of drug misuse at the local level. We begin with surveys then describe different which sources collate information on drug misuse and how this information can be used in describing drug misusing populations. We then describe how this enumeration of drug misusers can be systematically applied within a case finding study.

2.1 Surveys

There are many instances where surveying a population will be the easiest method of finding out information, such as the percentage of people which chose to vote for one political party as opposed to another. However, even in this simple example, there may be difficulties in reaching a representative sample. For example, if the survey methodology involved telephoning people in their houses during the day, then the results of the survey may be biased towards the opinion of those that are not in full time employment. If it could be assumed that sample is selected from all aspects of the general population, the statistical variability in the population may mean that the results of such a survey should be considered in conjunction with a measure of the possible error in the percentage. This may not be problematic if there are large differences in the voting habits of the population, however if there is little to distinguish the popularity of two or more political parties, then the results of such a survey may indeed be inaccurate.

In a similar fashion, there can exist difficulties in obtaining a representative sample of the population to distinguish drug users from non-drug users. This may be because users of heavier drugs, such as heroin, or drug injectors may be less likely to be reached in the typical manner of selecting samples for surveys, for example telephone subscribers, people on local government lists or people passing by in the street.

Additionally, there exists the problem that, unlike asking someone which political party they may favour, people will often be unwilling to disclose details of what is not only an illegal activity in many areas, but also often socially disapproved of. These problems combine with the fact that even in areas where levels of drug misuse may be high, only a small proportion of the population will be using such drugs and therefore a survey which is aimed at the general population could be inefficient at reaching those that use or inject drugs like heroin.

Although not always solely concerned with drug misuse, nation-wide general populations surveys have been undertaken in many European Union countries and the EMCDDA Annual Report (EMCDDA, 1998*a*) summarises the results from these surveys. The information gathered from such surveys is often more useful in estimating the current prevalence or the lifetime prevalence of less

problematic drugs such as cannabis, ecstasy, amphetamines or cocaine.

School surveys can also be carried out using a similar methodology to general population surveys. Ignoring truancy, it can often be easier to obtain a representative sample and apply a similar questionnaire in various settings. Almost all European Union countries undertake national schools surveys and an international study (ESPAD) was carried out in 25 European countries in 1995.

There are other specific populations that can be examined to quantify levels of drug misuse, for example prisoners or arrestees. Again the problem of people not disclosing information about their drug use may occur, although it can sometimes be assumed that the levels of drug misuse in these populations will be higher therefore some of the statistical problems of survey sampling may not be so pertinent. One alternative to relying on self reports of drug use is to test hair or urine samples; something which is becoming increasingly popular when examining drug misuse within prisons.

2.2 Enumeration

In many countries of the European Union, there are registers of drug users, principally those in treatment or entering treatment; for example the SEIT reporting system in Spain, the IVV/LADIS reporting system in The Netherlands, the RELIS/LINDDA reporting system in Luxembourg and the Regional Drug Misuse Databases in most of the United Kingdom. A demographical description of those included in such recording systems, along with an indication of the main drugs used is to be found in the Annual Report on the State of the Drugs Problem in the European Union (EMCDDA, 1998a). The police also collate information in many countries of the European Union, however just as any increase or decrease in the number of people entering treatment may be partly due to the increased availability of treatment or its effectiveness, any increase or decrease in the number of drug-related arrests may only reflect operational decisions of the Police. Even if these registers were highly correlated with the prevalence of drug use, by their definition, they can only be taken to be describing the numbers of drug users in treatment or the level of detected drug-related crime. Clearly the spectrum of drug using behaviour is not restricted to those who are in treatment or those that have been arrested for drug related offences.

2.3 Case Finding

These problems with registers can be partly resolved by undertaking a case-finding exercise. This terminology is more readily applied to national studies, such as Swedish case finding studies (Olson *et al.*, 1993; Olson, 1997), however in local studies it is sometimes expressed as multi-source enumeration. As Simon (1997) notes there are several methodological difficulties in undertaking a case-finding study. The coverage of the enumeration should be examined, particularly with respect of the

representativeness of those drug users identified within the study. The method used to identify drug users should enable the inclusion of all drug users which fit the study description (sensitivity) and the exclusion of those which do not (specificity). If however if a systematic approach is taken, particularly in relation to case definitions, then the results of the exercise are useful.

A local case-finding study, i.e. a multi-source enumeration, still needs to be systematic, although problems relating to geographical coverage may not be as pronounced. Hartnoll *et al.* (1985a) describe how such a study can be undertaken, coining the phrase 'Drug Indicators Method'. Despite being over a decade old, the basic epidemiological principles outlined in that report are as relevant today as when drug misuse prevalence estimation was in its infancy. Hartnoll describes the various sources of data on drug users which can be of use within a multi-source enumeration. These include hospital admissions, psychiatric hospital admissions, hepatitis data, mortality data, treatment centres, emergency rooms, general practitioners, needle exchanges, police data, as well as local extracts from national registers detailed above.

The above methods of investigating the prevalence of drug misuse at either the local or national level mostly concentrate on known drug users, i.e. those that who are either in contact with treatment centre or are know to law enforcement agency. Although general population surveys can structure their samples to possibly include less prominent drug users, there remains the problem that these drug users still have to divulge information on a covert subject and there is often no justification to assume that they will. Thus direct methods of examining drug misuse may sometimes only offer information on the known population of drug use, and although such information may be useful in providing a minimum estimate of the level of drug misuse in a specific locality, there is often a need to look further and to obtain information on less visible drug users. In the next section we shall examine methods which can be used to examine the hidden populations of drug users and therefore provide information on the broader spectrum of drug misuse.

3 Indirect Methods

In this section we review the literature pertaining to some indirect methods of drug misuse prevalence estimation; multiplier methods, synthetic estimation, and nomination techniques.

3.1 Multiplier methods

The multiplier method combines information on the size of a known population of drug users, such as the number in treatment, with information surveyed from elsewhere. When the known population is the number of drug-related deaths, then this specific example is known as the mortality multiplier method. Instead of second source, a multiplier is used and to obtain a multiplier, data concerning the prevalence of drug-related mortality must be examined. If, for example, 2% of drug injectors die per annum (as in-depth studies of mortality such as Oppenheimer *et al.* (1994) suggest), then the number of drug related deaths can be multiplied by the inverse of this proportion (multiplied by 50) to provide an estimate of the number of injectors.

Both of the parts of this equation are subject to uncertainty. First, the collation of information on drug-related deaths (which may differ from all causes of death of drug users or injectors) is often subject to error, as Davoli (1997) and Püschel (1993) point out. Additionally HIV may have an impact on non-AIDS related mortality (Eskild *et al.*, 1993), as may treatment regimes such as methadone; (Gronbladh *et al.*, 1990). Secondly, research into mortality rates of drug users usually concentrates on drug users receiving treatment (Perucci *et al.* 1991). These uncertainties combine to make Frischer (1997) conclude that its use is only appropriate in situations where basic assumptions are met, although it is one of the simplest ways of estimating drug misuse prevalence, partly because it may only require routine data.

A novel approach is that of Newmeyer (1988), who combined methadone treatment data and data on the proportion of deaths in which methadone metabolites were found by the coroner within the multiplier method framework to estimate the number of heroin users in San Francisco.

Although other feasible data are available to which a multiplier can be applied, such as treatment databases or crime statistics, such multipliers may be little more than a guess. However an eloquent derivation of an estimate of the number of drug users in San Francisco is provided by Newmeyer and Johnson (1976) who consider data on a number of factors. These are:

- the number of burglaries reported to the Police,

- the mean value lost as reported to the Police,
- a factor which accounts for the fact that the mean value of lost goods is less than the 'new' value of the stolen goods,
- the ratio of recorded to unrecorded burglaries,
- an estimate of the mean cash return to the burglar, the proportion of burglaries that are committed by opiate addicts,
- the average daily cost of an addict's habit,
- the number of days a year such an addict is so dependent on drugs they need to commit crime to feed their habit
- the proportion of addicts who commit burglary to feed their habit.

Newmeyer combines these data sequentially to estimate that there were 9,000 addicts resident in San Francisco in 1972, however as each of the 'links' in this chain are subject to error, the fact that he obtains what appears to be a plausible estimate may be solely due to luck.

3.2 Nomination Techniques

The techniques above use existing data to provide estimates of the size of drug using populations. In some situations, the available data needs to be augmented by other data, particularly that gathered during fieldwork studies. In essence, nomination techniques are similar to multiplication methods, although the second sample which is used as a benchmark, for example those that are in treatment, is gained using an approach such as snowball sampling. Parker *et al.* (1987) describe how variations of the technique were used in an area of North West England. In that area there were 237 drug users known to services, thus Parker set about asking a sample of 60 drug users a range of questions to ascertain what proportion of the total population size this value represented. One simple question was used to ascertain the percentage of the interviewed drug users were in contact with services. This known, the total population size could be extracted as described above in relation to multiplier methods. A second approach was to ascertain how many out of each of the respondents' close acquaintances had been in contact with the agencies. To improve on these methods, identifier information can be gathered on these drug users to check that they had in fact been in contact with services.

It is assumed that the original 'core' sample is a random sample of the more general drug using population, particularly if it is they that are being used to estimate the multiplying factor. An obvious violation of this assumption would be if the core sample was recruited from a treatment sample, other problems

could arise due to the age, sex or ethnic composition of the sample. For an additional discussion of the limitations of these methods see Taylor (1997), particularly in relation to the different sampling procedures.

Korf (1997*b*) describes how snowball sampling can generate data for use in nomination techniques, particularly in relation to research undertaken in the Netherlands. Just as those drug users identified from treatment agencies may introduce a bias into the prevalence estimation, those recruited from street samples in ethnographic studies may not be representative. Thus the ‘tip of the iceberg’ as he refers to the data obtained from these sampling methods, may not represent the part under water.

Korf notes that snowball sampling itself is somewhat more than just interviewing a sample of drug users that are easy to contact, and he describes how such a study can be undertaken. After the preparation of the field work, an initial wave of respondents needs to be contacted. The statistical representativeness of the initial sample, and thus the total sampled population, will depend on many factors. It is at this stage that other drug users are nominated by the initial wave. Korf suggests that care should be taken not to simply use the first person that each drug user suggests, rather a random selection from all possible referrals should be used. The statistical representativeness of the snowball samples can be checked, particularly in relation to their reported drug use. Other factors such as age, gender or ethnic origin can be used to examine the social structure of the snowball samples. Frank and Snijders (1994) provide a comprehensive review of this topic, illustrating the methodology using data from Dutch studies and Biernacki and Waldorf (1981) also discuss the methodology at greater length, and the statistical representativeness of snowball samples is examined by Watters and Biernacki (1989).

3.3 Synthetic estimation

To address the problems due to dependency between the known population and the sampled population, the multiplier method can be extended by including other types of data within an multivariate indicator model. This method has been used to estimate drug misuse prevalence at national levels (Mariani *et al.*, 1994). A similar but far simpler method, known as synthetic estimation also works by filling in the gaps between areas which have accurate prevalence estimates, both using ancillary variables such as treatment numbers or crime rates.

While these methods may both be more relevant to prevalence estimation at the national level as they provide these estimates by piecing together local estimates, they may be of use in obtaining local-level prevalence estimates in areas where other techniques may not be applicable.

Wickens (1993) describes how information on drug use in areas where it can be measured, perhaps by capture-recapture methods, can be transferred onto

areas where no such information is available. This is known as synthetic estimation. Thus if there were two areas where drug prevalence estimates were available, then the prevalence estimates can be charted against an ancillary variable such as the crime rate for each area. Once a straight line is drawn through the two points in this chart, then it is a simple matter of calculating a prevalence estimate for an area where the crime rate is known. This straight line then infers that there is a linear relationship between drug use and the ancillary variable; for example if the crime rate doubles, then the drug misuse prevalence will double. Thus, in effect, this simple example is similar to the multiplier method, although the data from two other areas are employed. Wickens also discusses some of the problems of this method. For example, if the two areas of known prevalence have similar prevalence values and crime rate values, then there is danger in extrapolating away from these two points in a chart to produce other estimates, i.e. the linear relationship between the prevalence and the other variable may not hold. Rhodes (1993) reviews the different data sources that can be used as auxiliary variables within a synthetic estimation study, and presents an application of the approach that was developed for purposes of policy research.

Not only does the multivariate indicator method extend the simpler synthetic estimation method by including more areas where drug misuse prevalence estimates are available, it improves on the basic method by employing more advanced regression techniques, such as principal component analysis. Woodward *et al.* (1984) describe a similar method, calling it factor analysis, which again examines the relationship between drug misuse prevalence and other ancillary variables. While these methods can be useful in obtaining national, or even regional, drug misuse prevalence estimates, the fact that they often rely on capture-recapture type local estimates suggests that other methods such as the capture-recapture method would be more appropriate, if possible, at the local level.

4 Capture-recapture methods

In this section we detail the historical development and the more recent scientific literature on a methodology commonly known as capture-recapture. Other names for this method have been used, such as mark-recapture, capture-mark-recapture (CMR), dual-record system (DRS), multiple-record system (MRS) and indicator-dilutor. We begin by charting the parallel development of the methodology by epidemiologists and biologists / ecologists, culminating in a series of review papers and commentaries. We then give examples of the uses of the methodology in epidemiological studies and then go on to detail how this method can be used to estimate the prevalence of drug misuse.

4.1 Historical Development

Those responsible for counting human populations, such as government officials or epidemiologists, have long aimed to completely enumerate the populations that they are interested in. Countries such as the US or the UK undertake a regular census of their population, hoping to gather details on the complete populations. It became apparent however that the objective of a complete census may have been optimistic and that there would be some level of inherent under-ascertainment. This was particularly evident in developing countries such as India where Sekar and Deming (1949) aimed to quantify this underascertainment in a small area, using the capture-recapture method. Deming went on to revolutionise the quality and dependency of Japanese consumer goods in the early 1950s, and even now, some commentators such as Laporte *et al.* (1993) paraphrase Deming by dismissing attempts at counting human populations, including disease monitoring systems, as ‘cheap and shoddy’.

While Bailey (1952) contributed to the methodological development in animal capture-recapture studies, another of the great statisticians of the 20th century, R.A. Fisher (see Fisher *et al.* (1943)), contributed to the related methodology in which the number of animal species is estimated from a random sample of an animal population. The use of both animal applications of the methodology were presented in a novel fashion in the 1970s by Carothers (1973) who estimated the number of taxis operating in the City of Edinburgh and Efron and Thisted (1976) who estimate the number of words that Shakespeare knew.

Seber (1992) summarises the major methodological developments and provides a review of the methodology in relation to estimating the size of animal populations, whereas el Khorazaty *et al.* (1977) present a review of the ‘dual system’ method first used by Sekar and Deming. Both of these reviews include the extension to the basic methodology where three or more data sources or samples are used. Pollock (1991) also provides an important review in this subject specifically for estimating animal populations.

As noted above, the development of the methodology came from different

disciplines, including population registers and animal populations. In this section we present a description of the methodology using an example from each and augment this description with an early application to estimating the size of a drug using population.

4.2 Two-sample methods

Population registers

Sekar and Deming noted the number of births from a registrar's list and compared these data with the result of a complete house-to-house canvass. By comparing these lists, they found the number of births recorded in both which they denote as C , the number of entries only recorded on the registrars list; N_1 , and the number of births found via the canvass; N_2 . They then estimated the total number of births from these three quantities by assuming that, with respect to the registrar's list, the ratio of known to unknown births in the canvass was the same as the ratio of known to unknown births in the whole population. Sekar and Deming note that there are several inherent assumptions. These include:

- There are no coverage errors with respect to the scope of area and/or time period in which events are recorded.
- The information sources are independent (i.e. the probability of an event being recorded by one source does not depend on whether it is recorded by the other source).
- There are no misclassification errors with respect to determining whether a particular event has been recorded by both information source or only one of them (i.e. a perfect matching rule exists for linking the two information sources together in terms of the number of events which are recorded by both).

If these assumptions hold, then an estimate of the total population size can be given as

$$\hat{N} = C + N_1 + N_2 + N_1 N_2 / C$$

Sekar and Deming go on to examine the effect that some of the violation of the assumptions may have on the estimate.

Animal populations

Larson *et al.* (1994) give a brief and coherent description of the two-sample method when estimating the size of animals populations. They describes how a scientist could count the number of fish in a lake by catching a sample of fish, marking them and returning them. At a later point the scientist then catches a second sample and records the number of marked and unmarked fish caught. These samples are assumed to be two independent samples of an unknown total population of size N . The number of fish caught in both samples is denoted as n_{11} , and the total numbers caught in the first and second samples are denoted as n_1 and n_2 . In essence, the capture-recapture method equates the ratio of the first sample n_1 to the total unknown population to the ratio of marked individuals n_{11} in the second sample n_2 , or

$$\frac{n_1}{N} = \frac{n_{11}}{n_2}$$

which can be reformulated as

$$\hat{N} = \frac{n_1 n_2}{n_{11}}$$

to give an estimate of the total population size.

It should be noted that the difference between Larson's formulation and that of Sekar and Deming is due to differing notation and the simpler representation of Larson can be obtained by substituting n_{11} for C and noting that the number found only in the first list $N_1 = n_1 - n_{11}$ in Larson's notation.

Larson is one of the few epidemiologists who additionally presents an unbiased estimator which should be used when samples are small:

$$\hat{N}' = \frac{(n_1 + 1)(n_2 + 1)}{(n_{11} + 1)} - 1$$

Larson also briefly describes the assumptions, which Seber extends and adapts. These are:

- The population is closed, so that the total number of animals that is being estimated is constant.
- All animals have the same probability of being caught in the first sample.
- Marking does not effect the catchability of an animal
- The second sample is a simple random sample
- Animals do not lose their marks in the time between the two samples

- All marks are reported on recovery in the second sample.

Drug Users

These simple equations can also be translated into an example that is more easily understood by those wanting to estimate the size of a covert population such as drug users. Hartnoll *et al.* (1985b) collected and examined data concerning opiate users who had attended a drug clinic and those that had been admitted to a hospital for infectious diseases. By comparing these sources of data, they found that approximately 20%, or a fifth, of the hospital sample had also attended the drug clinic. Thus the total number of opiate users could then be estimated to be five times the number who attended the drug clinic. Newmeyer (1988) compares mortality and treatment data to obtain a similar two-sample estimate for the number of opiate heroin users in San Francisco.

These simple examples mask some of the problems of the methodology. If those who were attending in the clinic were more likely to have been admitted to the hospital then the resultant figure would be an underestimate. Thus if there is some kind of relationship between those two data sources then the estimate will be biased. Unfortunately it is often unclear if such relationships, or interactions, are present and therefore the validity of these estimates are often questionable.

Most of the other assumptions outlined by Sekar and Deming and by Seber also translate into this epidemiological application, for example:

- The population is closed; there is no movement into or out of the population in the period that is being studied;
- Each individual has the same probability of being present in each source – that is the population is homogeneous;
- Being present in one source does not effect the probability of being in another source;
- Those who are present in more than one source - the overlap cases - are identified as such.

The first assumption translates into an assumption that the sources are samples from the same population and the fourth assumes that the identification of overlaps is not subject to error. This is not always the case as the matching of similar records across sources can often be subject to error. The second and third assumptions are often harder to describe as they relate to similar problems; drug users, for whatever reason, are not all equally as likely to be present in a source and those present in one source may be more or less likely to be in another.

Bishop *et al.* (1975) detail the two-sample methodology, a ‘basic’ approach and the ‘incomplete table’ approach, the latter as described above. The ‘basic’ approach, however arrives at the simple equation described by Larson by means of maximum likelihood estimation and using a multinomial distribution. The ‘incomplete table’ approach is sufficiently transparent for those wanting to undertake a prevalence estimation study or to understand the technique, whereas the more complex approach needs only to be understood to appreciate some of the more recent theoretical developments. Bishop *et al.* also present an estimate for the asymptotic variance, which can be used to construct a 95% confidence interval, and use an example from Sekar and Deming to describe the calculation of the estimate and an associated confidence interval. Hook and Regal (1995a) also discuss how the two-sample estimates can be obtained and present the formulation of Sekar and Deming, along with both the estimators described by Larson.

4.3 Three-sample methods

One of the main problems with the two-sample capture-recapture method is that, not only is the assumption that the two samples are independent often violated, but it is usually not possible to tell if the assumption is invalid. The capture-recapture methodology can compensate for this problem by employing three or more sources. The extra information present in the third sample can be used to examine whether or not dependencies are present between data sources, and if they are, the estimate of the total population size can be adjusted accordingly. Different relationships between data sources can be described using this analysis, for example if it were thought that those drug injectors attending treatment agencies were more likely to have been tested for HIV, then such a relationship can be included. The decision to include dependencies can be taken by examining how similar the observed overlap pattern is to what would be expected if such dependencies were actually present. Thus different models can be fitted to the observed data and a preferred model would be one that closely fits the observed data.

Bishop *et al.* demonstrate that the three-sample approach in the above example is an extension of the traditional two-sample method. Hook and Regal also detail the three-sample method and, as in the two-sample approach, the more theoretical approach of Bishop *et al.* may only be suited for those wishing to delve deeper into the methodology. The Hook and Regal review, along with the reviews of the International Working Group for Disease Monitoring and Forecasting (IWGDMF, 1995a; IWGDMF, 1995b) comprehensively describe the methodology, however the specific problems relating to drug misuse prevalence estimation are seldom mentioned. In particular the problems relating to case definitions are not so pertinent in other applications. When applying this method to drug using populations, it needs to be recognised that the case definitions of contributing sources differ, particularly with respect to the severity of drug use. For example, a police source may include drug users who do not, as yet, have problems with their drug use which require medical assistance. These less problematic drug users would therefore be less likely to appear in treatment centres. A review of capture-recapture methods with an

application in criminology is present by Smit *et al.* (1994).

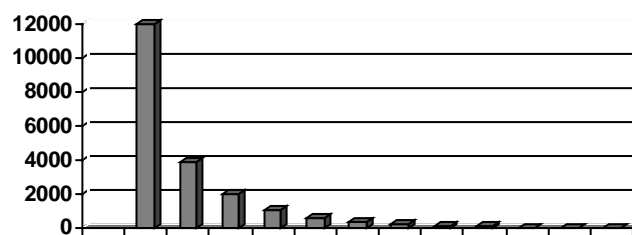
Although concise and comprehensive, the description and discussion of the methodology in review papers, such as Hook and Regal's, are aimed at the level of expertise of those that wish to estimate the prevalence of drug use. There is, of course, a far wider ranging collection of methodological papers which contribute to specific aspects of the development of the capture-recapture methodology, however such literature is not relevant for our application.

4.4 One sample methods

Although there are instances in the scientific literature where the capture-recapture method, using either two sources or more than two sources, has been used to estimate the size of populations such as drug users, there is a related method which only requires one source of data which may be applicable to drug misuse prevalence estimation. This is known as the truncated Poisson model.

Hser (1993a) applied a truncated Poisson model to data from the California Drug Abuse Data System. Using data from 1989, she extracted the number of individuals who had only been recorded once within the system, recorded twice, recorded three times and so on. Figure 1 charts these data.

Figure 1 Distribution of the number of times clients were recorded on the California Drug Abuse Data System. *Source Hser (1993a)*



Hser proposes that a 0 at 1 'drug use incident' is rare (Hser, 1993a). The distribution partially shown in Figure 1 has a Poisson form. Thus fitting a Poisson distribution to this data will offer an estimate of the number of people who are recorded in the database zero times, that is the hidden population (Blumenthal *et al.*, 1978). Wickens (1993) discusses Hser's example further. As in the traditional capture-recapture methodology, the accuracy of the estimate depends on the degree to which the observed data follow the Poisson distribution. An intrinsic assumption that must be met relates to the possible heterogeneity of the population that is being studied. If a substantial proportion of the population has little chance of contacting the treatment services that contributed to the database then the true population size will be underestimated.

The χ^2 distribution can be used to check whether the expected values obtained from fitting the Poisson distribution, and in Hser's example the extremely large χ^2 value (12,774 on 5 degrees of freedom) suggests a heterogeneity in the population with some individuals being more likely to enter treatment than others. Another critical assessment of the truncated Poisson model, in relation to estimating the prevalence of drug misuse, is provided by Simeone *et al.* (1993).

Smit *et al.* (1997) use adaptations of the truncated Poisson model as proposed by Chao (1989) and Zelterman (1988) to model treatment data from Rotterdam in 1994. Again, these estimators can be applied to data generated by counts of individuals identified from within a single data source once, twice and so on. However instead of using the complete count data, the numbers seen three or more times are only used in calculating the known population size and do not contribute to the estimator. As it can perhaps be assumed that people seen once or twice may be more similar to those never seen, employing only these counts in the estimator is intuitive. Additionally, this emphasis on the lower counts may partially resolve the problem of heterogeneity as evident in Hser's example. These estimators also have the added bonus that they are known to perform rather well even when few data are available.

Zelterman's estimator of the unknown population size, $est(n)$, is given by

$$est(n) = S / [1 - \exp(-2f_2 / f_1)]$$

and Chao's estimator is given by

$$est(n) = S + f_1^2 / (2f_2)$$

where,

f_1 = the number of persons falling in the first frequency class

f_2 = the number of persons falling in the second frequency class

S = the sum of all frequencies

These estimators also depend on the validity of the assumptions of the more general capture-recapture method, however they are not based on the independence assumption and the perfect matching assumption is less pertinent as it should be easier to match cases within a single data source. In relation to the homogeneity assumption, both estimators are known to be fairly

robust in the sense that both will *underestimate* the true population size in the presence of heterogeneity (Chao, 1989; Wilson and Collins, 1992). So, if heterogeneity is suspected, then one may reason that the estimate is a lower bound for the true population size.

Another main assumption is that an individual does not show a behavioural response to being identified at least once. This is analogous to the assumption in the three or more capture-recapture sample that the probability that an individual is identified in a given source is not influenced by their presence in another. In the case of Smit *et al.*, this would imply that those who had obtained treatment (with methadone) once would not alter from those that hadn't received it, at least in the time period that is being studied. Clearly, such an assumption may not be valid, and as the authors note, only a cynic would say that methadone maintenance is so ineffective that this assumption is not at risk of being violated anyway.

4.5 Epidemiological & Drug Misuse Applications

Epidemiological Applications

There have been many applications of the capture-recapture method in epidemiology, in particular the pioneering work of Wittes (1974). Early applications include estimating the underascertainment in birth defect registers (Hook *et al.*, 1980; Hook, 1982), and the use of the method soon spread to other diseases including cancer (Robles *et al.* 1988; Schouten *et al.*, 1994; Brenner *et al.*, 1995), measles, pertussis and tetanus (Davis *et al.*, 1993; Sutter and Cochi, 1992; Sutter *et al.* 1990; Mcgilchrist *et al.*, 1996), and more recently to HIV (Abeni *et al.*, 1994).

One area in which the capture-recapture methodology is now well established is the monitoring of insulin dependent diabetes mellitus (IDDM). Karvonen *et al.* (1993) present a review of the recent epidemiological data on the incidence of IDDM. They examine nearly 70 registries from more than 40 countries, noting that it is now rare that papers on IDDM incidence are published without a formal check on ascertainment. One famous study, which has been used to demonstrate the methodology in the review paper of the International Working Group for Disease Monitoring and Forecasting is that of Bruno *et al.* (1992), who uses the methodology to study the prevalence of both insulin dependent and non-insulin dependent diabetes mellitus in an area of northern Italy. Four lists are used and the modelling process on the combined data source (both those who were insulin dependent and those that were not) encountered problems in dealing with heterogeneity. To remove the cause of this heterogeneity, the population could then be stratified into three groups; those that are prescribed insulin, those that are prescribed hypoglycemic drugs and those that are prescribed dietary control. When the population was stratified as such, the heterogeneity was removed and models could be fitted to the data

which produced feasible results. Indeed the results of stratifying the population was to show that while 37% of the total who were prescribed dietary control were missing from the four sources, only two percent were missing from the group that were prescribed insulin. This gave rise to the interesting question; where did the 'hidden' population who were classified as being prescribed insulin get their insulin from?

All of the above studies are cited in reviews of the epidemiological uses of the capture-recapture methodology, however another group of studies is often grouped together and discussed with these disease studies, and there appears to be little discussion about these studies in which hidden populations, such as drug users, the homeless, or prostitutes are examined. Although there are many similarities between using three or more sources of data on diabetes to estimate the number that are not seen in any source, and estimating the number of drug users that have not been in contact with agencies such as treatment agencies or the Police, there are often important differences.

First, in the case of disease studies, it is usually the completeness of different registers that is being examined, after an acceptance that although an objective of such registers is to completely ascertain the prevalence of the disease and that process is subject to error. A statistical technique can be used to measure this error, and thus correct for it. This differs from drug use capture-recapture studies in that they are more similar to animal population studies where there a distinct population is being estimated. Therefore there is a related difference in that in disease capture-recapture studies, the 'hidden' population is usually far smaller than the known. For example in Robles' examination of the Ontario Cancer registry, the estimated percentage completeness of the registry ranged from 91% to 98% for a range of different cancers, and in Bruno's examination of the incidence of IDDM in Turin, the main existing data source had included 91% of the total cases.

Another important difference between disease studies and drug use studies is that disease studies pertains to case definition. Although International Classification of Disease codes exist for drug dependence and drug-related overdose, not every drug user in the population of interest may be subject to such a classification. This problem is also pertinent for those wishing to estimate the prevalence of homelessness or prostitution.

Although studies which examine injuries (Chang *et al.*, 1997; LaPorte *et al.*, 1995; Roberts and Scragg, 1994), phenomena such as children dependent on medical support (Palfrey *et al.*, 1994), abortion mortality (Cates *et al.*, 1978), those who have who have underwent a splenectomy (Sarangi *et al.*, 1997), or other medical conditions such as Rett syndrome (Kozinetz *et al.*, 1993), haemophilia (Hewitt and Milner, 1970) or stroke (Taub *et al.*, 1996) are slightly different to the disease registers detailed above, there are other applications which are more relevant to those estimating the prevalence of drug misuse.

Rossmo and Routledge (1990) used a version of the capture-recapture method to estimate 2 criminal population; 'fleeing' fugitives and street prostitutes, using data from Vancouver. Although they do not appear to explicitly mention it, their analyses are similar to the truncated Poisson model as described above. In addition to fitting a standard Poisson distribution to their data, they also adapt the distribution to allow two homogeneous subpopulations and also to have a continuous range of probabilities that different individuals are identified. Thus they attempted to solve the problems relating to heterogeneity that Greene (1984) found in his previous study. They describe the assumptions of the methodology in the context of their two applications, noting that not only are the criminals assumed to not change their behaviour (perhaps by a prostitute learning to evade arrest), but the police are assumed to behave similarly, for example by not applying more pressure on known criminals. Although most of the paper is concerned with the effect of fitting different statistical models to the data, they also go on to use auxiliary information about the population, such as the average length of time each prostitute was on the street. They claim that this additional information provides tighter confidence intervals and gives valuable insight into the nature of the population. They conclude that their population estimation techniques are of considerable value to criminologists, even with the inherent heterogeneity of criminal populations.

Bloor *et al.* (1991) adapted this methodology to estimate the prevalence of drug-injecting and non-injecting prostitution in Glasgow. They again employ a Poisson distribution for the probability that a prostitute is detected (this time within a study in which prostitutes were contacted during fieldwork in the 'red light' district), however they separately model new contacts and repeat contacts. Thus they are able to estimate the numbers of women on the streets each night, and also the cumulative total of prostitutes who worked in that area over a six month period. Smit also adapts truncated Poisson models to estimate the size of criminal populations (Smit *et al.*, 1996; Smit *et al.*, 1994), while Watts *et al.*, (1995) discuss the applicability of using capture-recapture methods to estimate the size of a prostitute population in a Zimbabwean city.

The prevalence of homelessness has been investigated in several studies. Laska and Meisner (1993) adapt the capture-recapture method into the plant-capture method to estimate the number of street-dwelling homeless individuals on a given date in 1990. The plant-capture method is intuitively simple; plant some 'known' individuals into the population of interest (in this case by getting people to dress and pose as if they were homeless in areas where homeless people were expected to be found) and then undertake a census of that population. Those who were planted in the population could later be contacted to find out if they had been included in the census, and thus the probability that they were included could be calculated. Just as the two-sample recapture method can be thought of as calculating the proportion of the population that appears in one of the sources (then extrapolate this to obtain an estimate of the size of the larger population), the proportion of 'plants' identified can be used to estimate the total population size.

The multi-sample capture-recapture method was used by Fisher *et al.* (1994) to estimate the number of homeless and homeless mentally ill people in an area of London. They employed six samples, and also stratified the population by sex, presence of mental health problems and age, giving eight population subgroups. Instead of estimating the prevalence for these groups separately, they included the stratifications as factors within the analyses. Thus their model was extremely large, for example the final model had a scaled deviance of 338.7 on 481 degrees of freedom. Using such a large model meant they had to be careful not to ‘overfit’ the model and inflating the population estimate. Shaw *et al.*, (1996) also proposed that the capture-recapture method can be of use in estimating the size of homeless populations.

The examples detailed above, along with others cited in epidemiological review papers, are accumulating as evidence as to the benefits of using capture-recapture methods in epidemiology and to estimate the size of hidden populations. Indeed LaPorte (1994) claims that ‘the use of capture-recapture techniques could bring about a paradigm shift in how counting is done in all the disciplines that assess human populations.’ Others are not as enthusiastic, indeed Cormack in an address to the International Biometric Society criticises LaPorte’s statement that ‘much of what we know about the size, distribution, and characteristics of wildlife is based on this and other approaches to counting with incomplete enumeration’. Cormack states that when nature has allowed a genuine count of population size, capture-recapture estimates have nearly always been found to be a gross underestimate of the true value (Cormack, 1968). As Hay (1997) points out the academic literature concerning capture-recapture methods in epidemiology may be biased in that it is far easier to publish a successful study than one that failed to produce a valid estimate. Even within ‘successful’ studies Hook and Regal (1995*b*) note that ‘one cannot establish that any estimate, no matter how plausible is in fact unbiased’.

Having accepted the potential problems with the capture-recapture methodology, it is still seen to be the most appropriate one in estimating the size of drug using populations, as evidenced by the EMCDDA-funded program of research.

Drug Use Applications

We can now briefly review some of the more recent applications of the capture-recapture and related methodologies to estimate the prevalence of drug misuse at the local level. We leave the discussion of European applications of the methods to a following section in which we provide a synthesis of results prevalence estimates.

The first apparent application of the capture-recapture method to estimate the size of a drug using population appears to be that of Greenwood (1971). Newmeyer and Johnson (1976), who estimate both the incidence and prevalence of heroin use in San Francisco, consider using that method in their study, however decide that the method fails to take into account of the fact that the size of one of their main data sources was increasing throughout the time

period of interest. They therefore construct a procedure in which the probability that a drug user dies in a given quarter is combined with data from a treatment agency to give an estimate of the total number of drug users in the city in 1972. Newmeyer repeats this analysis in 1987, directly referring to it as a capture-recapture method (Newmeyer, 1988). Doscher and Woodward (1983) estimate the size of a drug using population by applying log-linear models to capture-recapture. The two samples that are used are admissions at different periods to a drug treatment program. As these data were stratified by use into two groups, heavy and moderate, a contingency table with eight cells could be produced. However in this instance, two of the cells were missing (those heavy users who were hidden and those moderate users who were hidden). The fact that there were two hidden cells meant that certain constraints were placed on the analysis, however the authors found that a model proposed by Bonett *et al.* (1986) fitted the data well. In addition, they found that stratifying the data by level of use may not have been required, as demonstrated by the absence of that interaction in the model.

Hartnoll's and Newmeyer's two-sample capture-recapture studies have already been described within this report, therefore when listing the applications to drug misuse prevalence estimation in chronological order, the work of Frischer appears to be next (Frischer, 1992). This paper appears to be the first one published which applied the three-sample capture-recapture method, as it is commonly applied now in that it uses three distinct data sources sampled over the same time period to estimate the prevalence of injecting drug use in Glasgow in 1989. Bishop *et al.* (1975) did, however, note the methodology's applicability in this field. Mastro *et al.* (1994) also use the capture-recapture method to estimate the number of HIV infected drug injectors, this time in Bangkok, however in that study, only the two-sample method is used.

While Frischer was using capture-recapture methods on data from Glasgow in 1989 and 1990 (Frischer *et al.*, 1993), Domingo-Salvany *et al.* (1995) were applying the methods in Barcelona. Brecht and Wickens (1993) apply the capture-recapture method to successive samples from one source, however this source is simulated data. Wickens (1993) uses real data from California to demonstrate the method. This five-sample approach, however, appears more to be a description of the different approaches to model fitting that can be taken, for example fitting interaction between pairs of successive samples or triplets of successive samples. Hser (1993a) also takes this approach, this time with six successive samples from treatment data, each sample corresponding to a year, and also using three samples consisting of four-month intervals within one specific year.

Larson *et al.* (1994) use the capture-recapture method to estimate the number of heroin users in the Australian Capital Territory, using data from 1988 and 1989. After detailing some of the issues that need to be considered when undertaking a capture-recapture study, in particular the geographical coverage of data source, they go on to describe 11 treatment sources which they use in conjunction with data on heroin arrests to provide an estimate. Clearly each agency had a different remit, and although Larson does not specifically discuss

it, the sources would have a different geographical coverage.

This perhaps is not too much of a problem in Larson's study, as these treatment agencies combine their data within a drug indicators project, and it is these combined data that Larson initially uses within a two-sample capture-recapture analyses with the arrest data. The two successive years are analysed separately, and Larson also makes the distinction between heroin or methadone users that define their use as a problem. After noting the limitations of the two-sample method, when Larson applies the multi-sample method, but with 11 treatment sources, there is the immediate problem of how to reduce the sources. She does this by modelling different combinations of seven samples (after merging the six smallest treatment samples), merging similar data sources systematically until settling on a three-sample analysis with a methadone clinic, a combined sample of the other treatment agencies and the heroin arrests. Other similar analyses are presented and within her discussion the issue of heterogeneity is discussed.

Larson then details what lessons can be learned from this study, particularly in relation to estimating the prevalence of drug use in other areas of Australia (Larson and Bammer, 1996). She places the capture-recapture estimates in context with a household survey and multipliers derived from nomination techniques. The capture-recapture estimates, even the two-sample estimates, are far in excess of the data from studies. It is interesting to note that the estimate derived from the household survey is only slightly greater than the minimum enumeration from the capture-recapture study for the same year. Larson also discusses other capture-recapture estimates that have been obtained for Sydney; (Kehoe *et al.*, 1992; Duque-Portugal *et al.*, 1994), noting that these studies use data from the one source but at different time periods.

To close this section on the drug misuse applications of capture-recapture and related methods, it is worth noting that as the method becomes more and more established, there will more instances where its use will not be reported in the scientific literature. This is perhaps the case in some parts of the European Union where the capture-recapture method has been repeatedly applied in locations and the results of such studies are not disseminated further.

5 Dynamic Models

In this section we describe two further methods for estimating the prevalence of drug misuse at the local level. The following methods have not, as far as we are aware, been applied within Europe, therefore we turn to America for examples.

5.1 Open-population capture-recapture methods

One of the main criticism of the traditional multiple sample capture-recapture method as described above is that it assumes that the population is closed, such that there is no movement into or out of the population during the time period that is being studied. Clearly, this assumption may not be valid, and therefore models developed by biometricians to estimate the size of animal populations may potentially be of use to those wanting to estimate the prevalence of drug misuse.

The assumption that the population is closed can be violated in several ways. In contrast to animal studies in which samples are taken at given points in time (and where it is assumed that the time taken to construct each sample is negligible), the lists of drug users from various samples are collated over a specific time interval, such as a year. During such a time interval, people can begin to use drugs or to cease using drugs. Drug users may also die during the study period, or otherwise be removed from the population of interest, for example by being imprisoned. Another violation of the closed population assumption is more pertinent in areas such as The Netherlands, where 'drug tourists' can inflate some of the samples used within an analysis. Bello and Chêne (1997) took steps to make this assumption more valid by restricting their analyses to only include those that had were resident in the Toulouse conurbation. Larson *et al.* (1994), states that in practise that violation of the closed population assumption would not seriously affect the results, however this is done within the context of the two-sample method and thus may not directly translate into the three-sample case. Larson and Bammer (1996) go on to state that when mobility of drug users is high, and they cite the Korf *et al.* (1994) study in The Netherlands as an example, then the resultant estimates could be biased such that they are overestimates.

Although review papers such Woodward *et al.* (1985) and Simeone *et al.*, (1993) discuss the use of open-population capture recapture models, in particular the Jolly-Seber model (Jolly, 1965; Seber, 1982), in estimating the prevalence of drug misusing population, it appears only to be Hser (1993a) who has applied the models to actual data. As only one data source is required, sampled from at successive time periods, treatment data was suitable for her analysis.

Within the Jolly-Seber open population model, death and immigration of drug users is included, however similar assumptions as those in the closed

population models are still in place, such as those relating to heterogeneity. One additional ‘cost’ of the open population model is that estimates for the parameters describing death or immigration are needed, however they can be obtained by analysing the data with a computer program such as “Jolly” (Brownie *et al.*, 1989). Using data which refers to Los Angeles County in 1989, Hser gives separate estimates for drug injectors, heroin users, cocaine users and amphetamine users. She also contrasts the open population estimates with the closed population estimates, and those obtained using techniques which are detailed later in this report. In comparison with the other estimates, Hser notes that the Jolly-Seber estimate underestimated prevalence by nearly 50%. Simeone *et al.* (1993) are also critical of the use of open-population models, and also closed population capture-recapture and truncated Poisson models, however their criticisms refer more to the limitations of the treatment data.

5.2 Latent state Markov models

Finally in this section, we consider one of the few instances that latent state Markov models have been used to estimate the prevalence of drug misuse. Wickens (1993) describes how a population of drug users can be classified into ‘states’ which can reflect the level of drug use and whether or not they are in treatment. This is similar to the representation of Simeone *et al.* (1993), who describe a drug using population as consisting of two ‘ponds’; drug users and those in treatment. While these methods do attempt to introduce some of the system dynamics of drug use and then use this information in prevalence estimation, the example of Simeone which has these two artificial states is over-simplistic, and more realistic attempts run into the problem of not having enough information to estimate the necessary parameters.

Brecht and Wickens (1993) compare this methodology with the multi-sample capture-recapture method using a simulated population of drug users. With this constructed population, they find that the latent state Markov model is the most accurate estimator, however it should be remembered that their ‘hidden population’ had been artificially created.

6 Synthesis

In this section we provide a synthesis of some of the more recent prevalence estimation studies carried out within the European Union. Some of the studies detailed were included in an comparative study (EMCDDA, 1997) and the series of Annual Reports from the EMCDDA (EMCDDA, 1998a) tabulates prevalence estimates from throughout Europe.

Despite the comparative study, where to some extent the same definitions were applied in six cities, there appears to little commonality between the studies described below therefore comparisons between prevalence estimates should only be made with caution. For example, if one study examines opiate use and another looks at injecting or opiate addiction, then little can be said about the relative levels of drug use in those two settings.

Care must also be taken when comparing estimated prevalence rates when the baseline population differs. While there are moves to standardise age specific rates at 15-24, 25-34 and 35-64 years of age, the demographic characteristics of an area's drug misusing population may make direct comparisons difficult.

6.1 City Estimates

Austria

Vienna

Vienna is the capital city of Austria and has a population of over 1.5 million. As part of an EMCDDA comparative study, the prevalence of opiate use was estimated by Seidler and Uhl (1997). Four sources of data were available; police data, emergency ambulance data, hospital admissions and drug related deaths. As the hospital and ambulance data were found to be highly correlated, these two sources were merged together to provide the three sources for a capture-recapture analysis. To account for the fact that those in the drug-related death sources would have a reduced probability of being included in the other sources, the time period for collation of those data was the following year. It was therefore suggested that the prevalence estimate refers to 1993, and the estimate would refer to opiate use.

In total there were an estimated 6,747 (95% CI = 4,332 – 11,668) opiate users in Vienna in 1993, which corresponds to a population prevalence of 7.1 per thousand aged 15 to 54 (95% CI = 4.6 – 12.4). This estimate was contrasted with the various two sample estimates that could be derived from the four samples which contributed data. This estimated was contrasted with a previous unpublished estimate which assumed an annual mortality rate of 3% to estimate that there were 4,500 opiate users.

Vienna, 1993

6,474 opiate users (95% CI = 4,332 - 11,668)

7.1 per thousand (95% CI = 6.6 - 12.4) aged 15 to 54

Finland

Helsinki

While alcohol has been the traditional problem substance in Finland, the prevalence of drug use was stable in the 1980s and at a much lower level compared to other European countries. Although research in the 1990s began to examine recreational drug use, there had been few attempts to quantify the levels of drug addiction at the local level until Kaukonen *et al.* (1997) applied the capture-recapture methodology to three sources of data from Helsinki in the year 1995.

In contrast to other areas in Europe, the problematic use of amphetamines was seen to be of concern in Helsinki. Although this project was included within an EMCDDA-funded comparative study of opiate use, the inclusion of problematic amphetamine use along with opiate use demonstrated the ability of the capture-recapture methodology to adapt to local variations in the nature of drug misuse.

The three sources which were used within the Helsinki analyses were a hospital patient discharge register, a police data source which referred to opiate or amphetamine related offences and information on people who had been arrested for driving under the influence of drugs. These three sources combined to place the minimum number of opiate / amphetamine users in Helsinki in 1995 at 591. Although Helsinki is one of the smaller European capitals, with a population of just over half a million aged 15 to 54 in the Greater Helsinki area, the number of known users is comparatively small. The contingency table which described the overlaps therefore contained cells with low numbers, particularly when the data were stratified by age, sex or type of drug used. Partly because of this, models could be found for most age / sex categories, although some of the confidence intervals were wide. The authors did however feel that these models did not always adequately address the problem of heterogeneity. Given this caveat, the number of amphetamine users was estimated to be 2,356 (95% CI = 1,587 – 3,783) and the number of opiate users was estimated at 775 (95% CI = 487 – 1,393). These figures correspond to 4.3 (95% CI = 2.8 - 6.8) and 1.4 (95% CI = 0.8 – 2.5) per thousand population aged 15 to 54 respectively. These estimates were extrapolated to provide estimates of the total number of opiate or amphetamine users in Finland, and ongoing research aims to repeat the analyses for 1997.

Helsinki, 1995

2,357 amphetamine users (95% CI = 1,587 to 3,783)

4.3 per thousand (95% CI = 2.8 to 6.8) aged 15 to 54

775 opiate users (95% CI = 487 to 1,393)

1.4 per thousand (95% CI = 0.8 to 2.5) aged 15 to 54

France

Toulouse

The city of Toulouse is the capital of the Midi-Pyrénées region and of the Haute Garonne departement. Toulouse has a population of approximately 400,000 in the city, and 700,000 in the conurbation. A varied range of agencies were able to contribute data to a capture-recapture analyses (Bello, 1997; Bello and Chêne, 1997), and these included a psychiatric hospital, non-psychiatric hospitals, a low threshold agency, and legal sources including the police. These sources were asked to collect data on each drug user seen in the study period. Individuals who were identified as drug users were included, however individuals who were not declaring the use of any opiates or those who reported living outside of the Toulouse conurbation were excluded. Heroin was the most commonly used drug, and many of those included in the analyses were injectors.

To enable a three sample analyses, the data from the initial 10 sources were merged together by examining the odds ratio between pairs of sources. This process resulted in three groups of sources which could be seen as natural grouping of sources; medical, low threshold and legal. It could be argued that this process of merging sources together may have decreased some of the heterogeneity which would otherwise be inherent in the analyses, and as such there were few problems in model fitting. The total estimated number of opiate users in Toulouse in 1995 was therefore estimated to be 2,178 (95% CI = 1,780 – 2,734), which corresponded 5.4 per thousand (95% CI = 4.3 – 6.8) of the population aged 15 to 54. Stratified estimates were also easily obtained, and the pooled estimate varied little from the global estimates.

Toulouse, 1995

2,178 opiate users (95% CI = 1,780 to 2,734)

5.4 per thousand (95% CI = 4.3 to 6.8) aged 15 to 54

Germany

Berlin

As part of a larger, nation-wide study, the prevalence of 'hard' drug use in

West Berlin has been estimated using the capture-recapture methodology and also extrapolations from the results of a survey of the city's physicians. Noting that a capture-recapture study was previously undertaken in 1981 (Scarabis and Patzak, 1981), Kirschner (1997) compares the results obtained after using various methods in Berlin in 1992. The assumptions required for either approach are extensively discussed, in particular the possible consequences for estimation that any violations of these assumptions could have. Using the capture-recapture methodology, an estimate of 6,335 drug injectors is obtained, with a range of 5,780 – 7,123. The data used to obtain this estimate came from hospitals, welfare institutions, drug treatment centre, prisons and AIDS and drug-counselling services. The extrapolations from the physicians survey suggest a lower estimate of 4,744 with a wide range of 2,939 to 6,548. One reason why the latter estimate is lower could be because drug injectors in prison would not be included in the physician's survey. Although the capture-recapture study is not described in detail, an estimated range of 6,500 to 8,000 is quoted in Kirschner and Kunert (1997).

Bremen

The number of drug users in contact with services was found using the case finding method in Bremen in 1996 to 1997 (Zenker and Greiser, 1998). Medical, social and justice data were used to give a prevalence estimate of 4,347, which corresponds to a per thousand rate of 11.9 aged 15 to 54.

<p>Berlin, 1995 - 1996</p> <p>6,335 drug injectors (range = 5,780 to 7,123)</p> <p>3.1 per thousand (range = 2.8 to 3.5) aged 15 to 55</p> <p>Bremen, 1996 - 1997</p> <p>4,347 drug users in contact with services</p> <p>11.9 per thousand aged 15 to 54</p>

Ireland

Dublin

Dublin (population 1,058,714) is the capital city of Ireland, and approximately 30% of the population of Ireland live within its boundaries. Previous research identified high levels of risk behaviours within the city's injecting population, however previous research projects were unable to quantify the prevalence of problem drug use in the area, however Comiskey (1991) and Comiskey *et al.* (1992) in a 2 year survey of drug users estimated that a total of 375 people enter the drug using population each year with 198 of these being in the Dublin region.

By examining the number of drug users seeking treatment, O Higgins (1995)

found that the had almost doubled from 1990 to 1994. That review also found that the most commonly used drug was heroin. Statistics from the police confirmed that opiate use represented a significant problem. Also within Dublin, information was being collated on the number of people in receipt of methadone prescriptions and the number of people being discharged from hospital with a diagnosis of drug misuse. Thus a sufficient number of sources were available to undertake a capture-recapture analysis on the methadone data, the hospital data and the police data (Comiskey, 1997).

While the methadone data clearly refer to opiate use or problematic opiate use, and information on problematic opiate use can be extracted from the hospital data, there is less clarity on what population the information from the police data would refer to. The data were obtained from another study which focussed on the link between drug use and crime, therefore problems which some of the individuals from that source may be more social than medical.

When the data were merged together to establish the overlap pattern the minimum number of opiate users within Dublin in 1996 was found to be 6,449. For consistency with other studies, the analyses concentrated on the 6,264 opiate users who were known to be aged between 15 and 54. It should be noted that while 3,787 of these individuals were identified from the police source, and 3,169 were being prescribed methadone, only 885 were contained in both of these larger data sources. Thus a two-sample estimate would suggest that the total opiate using population would be in excess of 13,000.

When the overlap pattern was modelled within the capture-recapture analyses, there were difficulties in finding a model which adequately fitted the data. Indeed, before stratifying the data, the best model had a deviance of 15.73 on 1 degree of freedom. There was a similar lack of fit when males were considered separately from females. When the data were additionally stratified by age, there was more success in the modelling process. Although estimates were obtained for some specific age / sex groups, the estimate for the male 15 to 24 age group appears to be the most reliable one, in terms of the associated deviance and the size of the 95% confidence interval. For this age group there was an estimated 5,404 opiate users (95% confidence interval 4,980 – 5,891) which translates into a prevalence of 56 per thousand population of that age / sex group. Combining information from stratified analyses led the researcher to conclude that there were 13,460 (95% CI = 10,665 – 14,804) opiate users or 21.1 (95% CI = 16.8 to 23.3) per thousand population aged 15 to 54.

Although there were many instances in which the modelling process failed to provide a reliable estimate in terms of deviance, confidence interval or background information on the city's drug problem, the net result of the capture-recapture estimate was to provide a minimum estimate and to give information on the likely size of the total opiate using population.

Perhaps in Dublin, where the two largest data sources may be considering opiate use as 'problematic' from different angles, there problems of obtaining

a model which adequately fits the data were more acute. Indeed when the two more medical sources were used within a two sample capture-recapture estimate, then the total population was slightly under 6,000. This estimate not only contrasts with the known population when then Police data are included, but with either of the two sample estimates that are obtained when including the police data, or which the 'best' estimate combined from the various stratified estimates, or for that matter the estimate derived for the younger male age group. Further research in exploring the relationship between the definitions used by the contributing data sources and the estimates, particularly with respect to model fitting.

<p style="text-align: center;">Dublin, 1996</p> <p style="text-align: center;">13,460 opiate users (95% CI = 10,665 to 14,804)</p> <p style="text-align: center;">21.1 per thousand (95% CI = 16.8 to 23.3) aged 15 to 54</p>
--

Italy

Rome

Rome, with a population of 2.8 million, is both the capital of Italy and the main city in the Lazio Region. Information on drug addiction in Rome is available from a Surveillance System implemented by the Lazio Region. The system collects individual data on drug users attending public treatment centres and non-governmental organisations. Comparisons of the data from this system, from 1992 through to 1994, show that drug users attending treatment centres are predominantly males (more than 80%), 29 years old on average, more than 90% are heroin dependent and the injecting use is the primary route of administration for more than 75% of clients. Thus the prevalence of opiate use was examined by D'Ippoliti (1997).

Two other sources of data, hospital discharge data and data from a mobile emergency unit, were used in conjunction with the surveillance system within a three-sample capture-recapture analysis. As these two additional sources were medically orientated, and as the size of the surveillance source was far greater than even the other sources put together, the analyses in Rome could perhaps be seen as correcting for the underascertainment in the surveillance system. Models were found that fitted the overlap data well, particularly when the data were stratified by age and sex. In total, there were an estimated 14,278 (95% CI = 12,741 – 16,167) opiate users in Rome in 1996. This figure corresponds to 8.6 (95% CI = 7.6 – 9.7) per thousand population aged 15 to 54. The authors conclude that due to the nature of the sources, this estimate would refer to the prevalence of problematic opiate use. Comparisons are made with previous estimates derived from mortality multipliers which suggest that opiate use in Rome has been stable over the preceding years, and the Rome estimate is consistent with an estimate derived for the Lazio region

<p style="text-align: center;">Rome, 1996</p>
--

14,278 opiate users (95% CI = 12,741 to 16,167) 8.6 per thousand (95% CI = 7.6 to 9.7) aged 15 to 54

Luxembourg

Luxembourg City

Although a specific local prevalence estimation project has not been undertaken in Luxembourg, the country's participation in an EMCDDA-funded project to estimate the prevalence of drug misuse at the national level (EMCDDA, 1998a) has enabled an estimate for Luxembourg City (population 47,000 aged 15 to 54) to be extracted. The estimate of 760 'high risk consumers' refers primarily to opiate misusers or injecting drug users. This figure translates in a population prevalence of 16.2 per thousand, although as in other instances where estimates are extrapolated from other research, care must be taken in comparing this estimate to other cities.

Luxembourg City, 1997 760 high risk drug consumers 16.2 per thousand aged 15 to 54

The Netherlands

Alkmaar

As one of the smaller cities of the Netherlands (population 120,000) the drug using population was more readily examined using case-finding, nomination and snowball sampling techniques. Korf (1997b) explains in detail the methodology, however the comparatively low prevalence estimate of 98 opiates users in 1991 which was presented in Korf *et al.* (1991) may be attributable to the prevalence estimation technique.

Amsterdam

Although Amsterdam is the largest city in the Netherlands and drug use and its consequences have extensively been studied within the city, the prevalence of drug misuse has yet to be presented within the scientific literature. The Municipal Health Service in Amsterdam has however provided prevalence estimates using the two-sample capture-recapture method on treatment data and people requiring methadone when detained by the police (van Brussel *et al.*, 1997). Thus in 1996, the number of opiate addicts was estimated to be 5,769. When the population aged 15 to 54 is considered, these values represent 12.6. This population size estimate includes 2,205 foreigners and it is recognised in Amsterdam that the patterns of movement into and out of Amsterdam by drug using foreigners presents difficulties in obtaining concise

prevalence estimates.

Rotterdam

While there are regional and national differences in the nature and extent of drug misuse across Europe, it is perhaps in The Netherlands where the public perception and the policy response to drug misuse is the most unique. In particular, the legal situation as regards to possession of drugs for personal use practically invalidates the use of police data with a prevalence estimation exercise, and as previously mentioned, drug ‘tourists’ can present additional problems in prevalence estimation.

In the city of Rotterdam, this problem was addressed by using truncated Poisson models to provide an estimate of the number of opiate users from data on those in the city in receipt of methadone prescriptions (Smit, Toet and van der Heijden, 1997). Two established estimators as proposed by Zelterman (1988) and Chao (1989) were used, along with three estimators based on the truncated Negative Binomial regression as developed by one of the authors; van der Heijden *et al.* (in progress). A description of these methodologies is to be found elsewhere in this review, however the results of this study are presented within this synthesis.

Information was collated on 2,029 individuals who were included in the methadone prescription database. These data covered the whole of the City of Rotterdam (population 345,675 aged 15 to 54) and referred to the calendar of 1994. Using Zelterman’s estimators suggests that there are 3,727 opiate users in Rotterdam, whereas Chao’s estimator provides the slightly smaller estimate of 3,565. The confidence intervals attached to both estimates, 3,497 – 3,990 for Zelterman’s estimator and 3,348 – 3,818 for Chao’s, do not contradict either as the true opiate using population size. Both estimators were applied to data stratified by age and by gender, although the resultant pooled estimates are not substantially different from the total estimates. Both estimates suggest that the prevalence of problematic opiate use is just over 10 per thousand of the population, 95% CI = 10.1 to 11.5 for the estimator of Zelterman.

The authors go on to discuss the various estimators, in particular the estimators proposed by Van der Heijden which employed covariates such as age, sex and nationality. This discussion focuses on the issue of heterogeneity and conclude that both Zelterman’s and Chao estimators, which address this issue, would be the preferred estimate. The note that these estimates do not substantially contradict an estimate of Wiessing *et al.* (1995) who used the multiplier method to indicate a population size in the range 2,400 to 3,500.

Utrecht

The prevalence of opiate use in Utrecht has also been estimated using a range of methods. ten Den *et al.* (1995) used the two-sample capture-recapture

methodology on police data and methadone data to show that there were approximately 950 opiate users in that city in 1993. Nomination techniques and network analyses were used to complement the capture-recapture analysis.

<p>Alkmaar, 1991 98 opiate users 1.8 per thousand aged 15 to 54</p> <p>Amsterdam, 1996 3,564 opiate addicts (5,769 including foreigners) 7.8 per thousand (12.7) aged 15 to 54</p> <p>Rotterdam, 1995 3,727 opiate users (95% CI 3,497 to 3,990) 10.8 per thousand (95% CI 10.1 to 11.5) aged 15 to 54</p> <p>Utrecht, 1993 950 opiate users 6.3 per thousand age 15 to 54</p>
--

Portugal

Setúbal

The city of Setúbal, with a population of 85,292, is the third largest in Portugal and previous research has shown that Setúbal county has a significant drug problem and associated high levels of AIDS and hepatitis. Information from the Ministry of Justice suggests that this area has a significant problem of drug related crime. The prevalence of opiate use in 1996 was estimated by Freire and Moreira (1997).

Despite these indicators of the area's drug problem, there are difficulties in collating information on drug users. Although four sources of information on drug misuse were identified, only two sources could be accessed. These were drug addicts who were in treatment in a specialised public drug treatment centre and drug addicts who were in treatment at general health centres. Clearly both sources would refer to drug users, and in this case opiate users, who had a medical problem relating to their use of these drugs.

In an attempt to explore any relationship between these two sources which could have made a two-sample estimate invalid, the data from the health centres was split into two semesters for the years 1996. While this can only partially address this problem, particularly as the expected interaction between

the two semesters was shown to be needed in the analysis, an estimate of the prevalence of opiate use in Setúbal city in 1996 was obtained, as were stratified estimates by age and sex.

The estimate of 894 (95% CI = 620 – 1,423) corresponds to 18.2 per thousand of the population aged 15 to 54 (95% CI = 12.6 –29.0), and thus confirms that drug misuse is a significant problem in that area. Again the case definitions of the contributing sources would suggest that the estimate refers to the more problematic end of the opiate using spectrum.

<p style="text-align: center;">Setúbal, 1996</p> <p style="text-align: center;">894 opiate users (95% CI = 620 to 1,423)</p> <p style="text-align: center;">18.2 per thousand (95% CI = 12.6 to 29.0) aged 15 to 54</p>
--

Spain

Barcelona

Two studies, by the same research team, have estimated the prevalence of opiate use in Barcelona in 1989 and in 1993. In the original paper (Domingo-Salvany *et al.*, 1993) three sources of data were used; hospital emergency rooms, treatment admissions and heroin overdose deaths. Two different strategies were adopted; emergency room data only, divided into three trimesters and all three sources over the one time period. All data refer to opiate users in the age range 15 to 44.

With their first strategy, information on 2,075 addicts were analysed using various two-sample and three-sample analyses, with the four sample analysis being used to test for dependence between trimesters. The population estimates obtained by comparing pairs of trimesters ranged from 2,466 to 3,516, however the authors noted that they could not assume independence, particularly between contiguous trimesters. When the combinations of three samples were employed, none achieved an adequate fit. The four sample analysis suggested that all two-way interactions were present, therefore the authors applied a formula which assumed dependence between three samples to each combination. These estimates, which ranged from 4,679 to 6,298, were higher than the two-sample estimates.

The second strategy, using three distinct samples, also began by examining the various two-sample combinations and stratified the data by age and sex. The two-sample estimates were higher than the two-sample estimates using the trimesters of the emergency room data; ranging from 6,744 to 8,171. When the three-sample analysis was undertaken, the independence model provided a satisfactory fit and thus estimated the prevalence of opiate use at 6,831. This corresponded to a population rate of 9.2 per thousand; 13.2 per thousand male and 4.8 per thousand females with the majority being under 29. The discussion section of the paper describes some of the issues relating to applying the

methodology, in particular the difficulty of delimiting appropriately the population under study.

The second study, using data from 1993 (Domingo-Salvany *et al.*, 1998), concentrates more on the methodological issues, specifically identifying population definition, source heterogeneity and assessment of an adequate model as the main problems they encountered.

In this study, which extended the geographical remit, a fourth source of data, from a prison, was available. Emergency room and treatment data were also available for the whole of the Barcelona metropolitan area, however mortality data were only available for Barcelona city. This could then possibly enable a direct comparison to the previous study. In total information was gathered on 3,207 individuals in the metropolitan area, with a further 83 individuals appearing in the mortality data for the city. Again pairs of samples were examined initially, before the three-sample analysis was explored, stratified by age and sex. The Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were used to assist model selection and weighted estimates were also obtained.

Various estimates were presented, using different data sources, geographical areas and methods for model selection. Two estimates are highlighted; 12,894 (95% CI = 10,594 – 16,132) individuals in the metropolitan area and 9,176 (95% CI = 7,188 - 12,222) in the city. These correspond to 11.2 (95% CI = 7.2 – 11.0) and 13.0 (95% CI = 10.1 – 17.2) per thousand aged 15 to 44 respectively. Although these figures suggest an increase in prevalence in the city since 1989, the inclusion of the prison data could have widened the ascertained population base. The authors discuss the adequacy of the mortality source in the previous estimate.

Madrid

The prevalence of opiate addiction in 1992 in Madrid has been extensively studied by Paredes *et al.* (1994). Four samples are employed; deaths, drug treatment agencies, AIDS registers and prisons. The known population from each source is described in depth, examining the data by socio-demographic information and by location within the *Comunidad de Madrid*. A three-sample capture recapture analysis is performed on the death, drug treatment and AIDS data, which estimates the number of heroin addicts to be 49,260 (95% CI = 38,804 – 59,716). These figures correspond to 21.1 (95% CI = 16.6 – 25.6) per thousand population age 15 to 44, however after considering estimates derived from multiplier methods, the authors conclude that in total, there were approximately 41,000 heroin addicts in the Madrid metropolitan area. This represents 17.6 per thousand of the population aged 15 to 44 in that area.

<p style="text-align: center;">Barcelona, 1993</p> <p style="text-align: center;">9,176 opiate users (95% CI = 7,188 to 12,222)</p> <p style="text-align: center;">13.0 per thousand (95% CI = 10.1 to 17.2) aged 15 to 44</p> <p style="text-align: center;">Madrid, 1992</p> <p style="text-align: center;">41,000 heroin addicts</p> <p style="text-align: center;">17.6 per thousand aged 15 to 44</p>
--

Sweden

Malmö

The case-finding method, which has been applied nationally in Sweden, was supplemented by the capture-recapture method to estimate the prevalence of severe drug use in Malmö (population 236,684) in 1992 (Olson *et al.*, 1993). The case definition used in this study was that an individual has injected at least once in the first year, or that they used any illegal drug (including cannabis and ecstasy) either daily or almost daily. 44% were opiate users and 95% had injected within the last year, although amphetamine was the most commonly injected drug. Data were obtained from needle exchanges, treatment centre, social services and detention centres. Using this definition, there were between 1,100 and 1,300 severe drug abusers, figures which correspond to just under 1% of the population aged 15 to 54.

Stockholm

The case-finding method was again used to estimate the prevalence of drug use in Stockholm (population 727,339), the Swedish capital in 1995, 1996 and 1997 (Olson *et al.*, 1993; Finn, 1997; Finn, 1998). This time the definition was more akin to drug users in contact with services, including cannabis or ecstasy users who were in contact with social services. The prevalence estimates were relatively stable over this time period, rising from 1,656 to 1,792. These figures correspond to just over 0.4% of the population, and these estimates should be considered along with the case definition which excludes drug users not in contact with services.

<p style="text-align: center;">Malmö, 1992</p> <p style="text-align: center;">1,100 to 1,300 severe drug abusers</p> <p style="text-align: center;">8.8 to 10.4 per thousand aged 15 to 54</p> <p style="text-align: center;">Stockholm, 1997</p> <p style="text-align: center;">1,792 drug users in contact with services</p> <p style="text-align: center;">4.3 per thousand aged 15 to 54</p>
--

United Kingdom

Aberdeen

The prevalence of opiate and benzodiazepine use was estimated in the City of Aberdeen in 1997 (Centre for Drug Misuse Research, 1998). This city, in the north east of Scotland has a population of 215,903. Six sources of data were available; the police, a drug treatment agency which offered substitute prescribing, a low threshold agency, a needle exchange, general practitioners contributing to the Scottish Drug Misuse Database and people who have had their drug-related problems assessed in terms of their offending behaviour in order to assist the courts in assigning sentences. These assessments are carried out by social workers and are known as Social Enquiry Reports. Due to the nature of the services in the city, the police data and the Social Enquiry Report data were merged to give a single source, whereas joint working between general practitioners and the city's drug treatment agency led to those sources being merged.

Thus in total, there were 1,129 opiate or benzodiazepine user identified in Aberdeen. A four-sample capture-recapture analyses was undertaken on these data, and the total number of drug misusers was estimated to be 2,396. The data were stratified by age and by sex, and analyses were undertaken in these groups. The authors preferred to quote the sum of the stratified estimates therefore they conclude that there were 2,519 drug misusers; this figure comprises 1,941 males and 578 females. This corresponds to 20 per thousand of the population aged 15 to 54 in the city.

Dundee

Two papers report on a prevalence estimation project undertaken in Dundee (Hay and McKeganey, 1996; Hay, 1997). In total five sources were available in Scotland's fourth largest city (population 88,515 aged 15 to 45); police, the city's drug treatment agency, HIV test data, general practitioners' notifications to the Scottish Drug Misuse Database and data from a concurrent HIV behavioural and seroprevalence study. The original paper presents the prevalence estimate using the first four sources of data, the information on 855 individuals was used in the capture-recapture analyses to estimate that there were 2,557 (95% CI = 1,974 - 3,458) drug misusers in the city, which corresponded to a population prevalence of 28.8 (95% CI = 22.3 - 39.0) per thousand aged 15 to 54. The definition that this study gave for drug misuse was the use of opiates or benzodiazepines; the latter class of drugs being commonly used in the absence of opiates.

In contrast to other studies, the time period which the data refer to was long, ranging from 1990 to 1995. This raises questions about the validity of the closed population assumption. Despite this long period, there were not sufficient data to stratify the analysis by age or sex.

The second paper contrasts the results obtained by looking at different combinations of the five available sources, including all three-sample analyses and all four-sample analyses. This paper again suggests that the number of drug misusers in the city was approximately 2,500, however it presented examples using three and sometimes four sources, where the best model offered substantially lower estimates. The author concluded that although the capture-recapture methodology may be the best for estimating drug misuse prevalence, there may be instances where the estimates are unreliable, even when the models fit the data well.

Glasgow

The prevalence of injecting drug use was estimated in 1989 and in 1990 by Frischer (Frischer, 1992; Frischer *et al.* 1993). In the first year, three sources were employed within a capture-recapture analyses; treatment agencies, an HIV reporting scheme and the police. The second year additionally used data from the city's needle and syringe exchanges.

In the first analyses, data were gathered on 1,738 individuals, and the formulae described by Bishop *et al.* (1975) were used to estimate the total size of the injector population at 13,050. This assumed a relationship between the HIV sample and the treatment sample. Following Drucker and Vermund (1989), Frischer applied a downward correction factor to his estimate, in this case by taking the lower limit of the symmetric 95% confidence interval. He therefore concluded that there were 9,424 injectors in Glasgow, with an associated confidence interval 6,964 to 11,884. The analyses were repeated for age and sex stratified groups, however in some of these groups the symmetric confidence intervals encompassed negative values. The 9,424 injectors corresponded to a population prevalence of 15 per thousand (95% CI 11.0 – 18.9 with 22 per thousand male and 8 per thousand female in the 15 to 55 age group.

Most of the methodological problems with Frischer's first analyses were rectified in the second study where the fourth source enabled more choice in model selection. As in the first analyses, Frischer used data on all non-cannabis offences that had occurred in Glasgow. Although over 50% were for opiate offences, 20% were for amphetamines or cocaine and 10% were for hallucinogens such as LSD or MDMA. Although it is claimed that drug users in Glasgow inject a wide range of drugs, it is unclear what proportion of this sample would actually be drug injectors.

In total, information was gathered on 2,886 individuals, and the four-sample capture-recapture analyses suggested that the best model included interactions between the HIV and treatment sources, the HIV and needle exchange source and between the needle exchange and treatment source. There also appeared to be a three way interaction between these three sources, thus there were effectively two independent sources of information; the police list and a combined list from the other sources. Thus the analyses could be reduced to a

two-sample analyses, although which ever method is used, the total number of injectors is estimated at 8,494. The 95% confidence interval was this time calculated following Cormack (1992), giving an upper limit of 9,721 and a lower limit of 7,491. This figure corresponds to 13.5 per thousand of the population aged 15 to 55 (95% CI = 11.9 – 15.5). Again the analyses were repeated on data stratified by age and sex. This paper combined injectors prevalence information with information on the prevalence of HIV in the injector population, suggesting that there were 93 HIV-infected current injectors in Glasgow.

Liverpool

The estimation of the prevalence of drug misuse in Liverpool in 1991 was combined with a spatial analysis of known addiction by Squires *et al.* (1995). In this instance, drug misuse was restricted to opiate or cocaine use and the three-sample capture-recapture method was applied to data from drug dependency units, an infectious disease unit and the police. Data on 1,427 individuals were analysed and this provided an estimate of the drug using population of 2,344 (95% CI = 1,972 – 2,716). The 95% confidence interval was obtained using the formulae of Bishop *et al.*, however the symmetric interval gives a range with higher limits than the method proposed by Cormack (1992). The population prevalence equates to 5.2 (95% CI = 2.5 – 6.0) per thousand of the total population, and in the 15 to 29 age group, the number of drug misusers was estimated to be 16.9 (95% CI = 13.9 – 19.9) per thousand population. The spatial analyses showed that drug misuse was concentrated in areas of the city and there was a correlation with material deprivation, in particular unemployment.

<p style="text-align: center;">Aberdeen, 1997</p> <p>2,519 opiate/benzodiazepine misusers (95% CI = 2,048 to 3,200) 19.8 per thousand (95% CI = 16.1 to 25.2) aged 15 to 54</p> <p style="text-align: center;">Dundee, 1990 -1994</p> <p>2,557 opiate/benzodiazepine users (95% CI = 1,974 to 3,458) 28.8 per thousand (95% CI = 22.3 to 39.0) aged 15 to 54</p> <p style="text-align: center;">Glasgow, 1990</p> <p>8,494 drug injectors (95% CI = 7,491 to 9,721) 13.5 per thousand (95% CI = 11.9 to 15.5) aged 15 to 55</p> <p style="text-align: center;">Liverpool, 1991</p> <p>2,344 opiate / cocaine users (95% CI = 1,972 to 2,716) 16.9 per thousand (95% CI = 13.9 to 19.9) aged 15 to 29</p>
--

6.2 EMCDDA Comparative Study

It is clear from the studies detailed in the preceding section that different prevalence estimation methods have been used and even when a particular method has been employed, there have been variations in the implementation, in particular the case definition. However, in a methodological pilot study, funded by the EMCDDA (EMCDDA, 1997), the prevalence of opiate use was estimated in seven cities, six of which employed the three-sample capture-recapture method. One of the main objectives of the project was to achieve comparability in applying the method throughout Europe therefore several definitions were common to each city. In particular this pilot study focussed on the use of opiates. It was, however, noted that the case definitions used by many of the contributing sources would mean that many of the resultant estimates would be of problematic opiate use.

The individual research projects have previously described, however the known and estimated total number of opiate users, the prevalence in the 15 to 54 age group, and 95% confidence intervals are presented in Table 1.

Table 1 Prevalence of opiate use in seven cities.

City	Known Users	Total Users		Prevalence (%)	
		Est.	95% CI	Est.	95% CI
Dublin	6,264	13,460	10,665-14,804	2.11	1.68-2.33
Helsinki	175	775	487-1,392	0.14	0.09-0.25
Rome	6,896	14,278	12,741-16,167	0.86	0.76-0.97
Rotterdam	2,029	3,716	3,497-3,990	1.07	1.01-1.14
Setúbal	339	894	620-1,423	1.82	1.26-2.90
Toulouse	799	2,178	1,780-2,734	0.54	0.44-0.68
Vienna	1,028	6,747	4,332-11,668	0.67	0.43-1.16

As each city stratified the analysis by age and gender, the numbers of male and female opiate users estimated in each city, along with the numbers of drug opiate users stratified by age are presented in Table 2 and Table 3 for those cities which used the three-sample capture-recapture methodology.

Table 2 Prevalence of male and female opiate use in seven cities.

City	Males		Females		Male : Female
	Estimate	%	Estimate	%	
Dublin	6,831	2.2	3,179	1.0	2.15:1
Rome	12,649	1.5	2,368	0.3	5.34:1
Setúbal	696	2.9	208	0.8	3.35:1
Helsinki	536	0.2	310	0.1	1.73:1
Toulouse	1,709	0.8	466	0.2	3.67:1
Rotterdam	1,485	0.8	544	0.2	2.73:1
Vienna	5,746	1.2	554	0.1	10.3:1

Table 3 Prevalence of opiate use in the young and old age groups.

City	Young			Old		
	Known	Estimate	%	Known	Estimate	%
Dublin	5,604	10,964	2.94	660	1,067	0.41
Rome	4,521	10,365	1.17	2,375	4,447	0.58
Setúbal	284	884	3.54	55	85	0.72
Helsinki	54	122	0.11	121	464	0.11
Toulouse	649	1,709	0.74	131	426	0.24
Vienna	613	3,393	0.71	94	3,354	0.73

There appeared to be some difference in the prevalence of opiate use in the different cities; ranging from the low levels of use in Helsinki to the higher levels found in Dublin and Setúbal. The male to female ratio was lowest in Helsinki, and higher in Rome, Setúbal and Toulouse, whereas the highest prevalence values in the young age groups were to be found in Dublin and Setúbal.

6.3 Regional Estimates

In this section we describe some of the few studies which have sought to use

the apply methods, such as the capture-recapture methodology, in non-urban settings.

Portugal

Setúbal

The Setúbal county was examined by Godinho *et al.* (1998) who aimed to provide an estimate of the number of heroin users in 1995, 1996 and 1997. Three samples were used; a police sample, infectious disease consultations at the local hospital and the region's treatment centres. 1,074 individuals were identified as using heroin within those three years, and from these data, an estimate of 4,218 users (95% CI = 2,646 to 7,385). These figures equate to just over 4% of the total population and 7% of the population aged 15 to 54. This figure contrasts with the estimate of Freire and Moreira (1997) who estimated that the rate in Setúbal City was 1.8% of the population aged 15 to 54. The discrepancy in these two estimates could partially be due to the longer time period studied at the regional level or by the addition of a police data source within the analyses.

Spain

Navarra

The prevalence of heroin use in the Autonomous Community of Navarra, a region in the north of Spain, was estimated using multi-agency case finding (Urriaga-Dominguez *et al.*, 1993). This method consisted of collecting information from public treatment centres for drug dependencies, therapeutic hospitals, hepatitis' case register, HIV+ case register, prisons, forensic and various social services. The study covered the whole area (512,512 residents) during the year of 1990. Overall, 1,231 persons were counted as heroin users representing a prevalence rate of 2.4 cases per thousand residents. They were mainly male (3 to 1), and 72% of cases were aged between 20 and 29. Most of the opiate users lived in urban areas and it was found that there were large differences among the different urban zones. Unlike the capture-recapture method, the problems of geographical heterogeneity should not be so pertinent as extrapolating from these known drug users in not undertaken within a case finding study.

UK

Cheshire

The South and East Cheshire districts, describe as a rural district, was examined by Brugha *et al.* (1998). Just over half the population of that area lived in one of 12 small to medium towns. Data were collated from four sources; community drug teams, needle exchanges, the Police and general

practitioners. These data referred to opiate, amphetamine or cocaine misusers present in that area in 1993. A log-linear model was applied to the overlap pattern between these four sources, giving a prevalence estimate of 1,094 (95% CI: 682-4,153). This large confidence interval would be due to the complexity of the model, which included two three-way interactions. A less complex model gave a higher estimate of 1,639 misusers but a narrower confidence interval of 1,136 to 2,643. The authors concluded that the prevalence per thousand of the total population was 2.5; which is about half that of the nearby city of Liverpool. They do admit however that the target population was small and heterogeneous and that there was a wide confidence interval attached to the estimate.

Grampian

As part of an on-going program of prevalence research in Scotland, the Centre for Drug Misuse Research (CDMR, 1998) has estimated the prevalence of opiate or benzodiazepine misuse in the Grampian area in the North East of Scotland. This area includes the city of Aberdeen and a range of small towns, many of which depend on the fishing industry. One such town, Fraserburgh, has attracted much media interest due to the high numbers of opiate overdoses, and the prevalence estimate not only gave estimates for four sub-areas of the region (in addition to the city of Aberdeen) but also Fraserburgh. In total seven distinct data sources were available, but the coverage of these sources varied across the region so separate three or four sample analyses were conducted in each sub-area. These data sources include drug treatment agencies, needle exchanges, the police and data on people referred to social work departments when considering sentencing for drug-related crimes such as theft or shoplifting.

The total number of opiate or benzodiazepine misusers in Grampian was estimated to be 3,626 (95% CI 2,674 to 5,918). This corresponded to 1.2% of the population aged 15 to 54. Not surprisingly, there was significant variation between the sub-areas of the region, with the city of Aberdeen having a higher prevalence at 2.0%. The suspected high prevalence of opiate or benzodiazepine misuse in Fraserburgh (which has a population aged 15 to 54 of only 13,000) was confirmed using the three-sample method to be approximately 2.5% in the relevant age group.

Lanarkshire

The Health Board Area of Lanarkshire stretches south-west from the city of Glasgow and includes neighbouring towns which are sometimes described along with Glasgow as the 'Clydeside Conurbation'. In addition, the local government areas of North Lanarkshire and South Lanarkshire include urban areas which were part of the city of Glasgow until local government re-organisation in 1995. Due to the geographical variation in coverage of data sources, and the requirement of the study funders for both Health Board and local government estimates, six different areas were examined by the Centre

for Drug Misuse Research (CDMR, 1997) to estimate the prevalence of opiate or benzodiazepine misuse in 1996/1997.

In addition to the provision of drug treatment agencies varying across the region, the patterns of substance use varied immensely across the region. Many drug treatment agencies dealt with alcohol problems, and it was clear that alcohol misuse was a more significant problem in some areas. This presented an additional problem in that in some data sources, it was sometimes unclear which drugs an individual actually used. To alleviate this problem, the matching between sources was done on all data and the contingency tables were adapted in light of overlaps where opiates or benzodiazepines were noted in one source but un-named drug were noted in others.

In total, information on the known population of 1,146 was used to estimate that there were just over 5,000 misusers in the Health Board area. Thus the prevalence aged 15 to 54 was 1.6%, with a 95% confidence interval of 1.3 to 2.0. Apart from the more rural area of the region and a small area where there were difficulties in finding a known population due to the lack of service provision, the prevalence rates for the various sub-areas were consistently between 1.3 and 1.7%.

Wales

Wales, with a total population of approximately 3 million, is the third largest of the four countries which make up the United Kingdom. In 1994, a prevalence estimation study aimed to use the capture-recapture methodology to provide estimates for the whole of Wales and the eight sub-areas (known as counties) which the country is split into (Bloor *et al.*, 1997). Two case definitions were employed; injecting drug use and serious drug use. The latter definition additionally included heroin smokers however the seriousness of the individuals drug use was defined by the contributing sources.

There was only one data source which covered the whole country; the Welsh Drug Misuse Database. This data sources records new attendances at drug treatment agencies, general practitioners and prisons. Enough data were available to identify injectors. Treatment agency data were available in some areas, as were needle exchange data, probation or police data and HIV test data. Thus with eight counties, two case definitions and different combinations of data sources, a multitude of different models were applied to the collated data. In total, the prevalence of serious drug use in Wales was estimated at 8,357 (95% CI 5,307-11,407). This corresponded to 5.34 per thousand of the population aged 15 to 55. In the county which includes the capital, Cardiff, the prevalence of injecting drug use was 1,058; a rate per thousand aged 15 to 55 of 4.59. This prevalence rate compares to 8.71 for serious drug use. Although it was possible in some instances to stratify the data geographically, it should be noted that at the county level, the area will include both urban and non-urban locations therefore geographical heterogeneity may have posed a

problem for the study.

6.4 Summary

In this Section we have presented a brief summary of some of the more recent prevalence estimation studies throughout the European Union. We have focussed more on studies which have employed epidemiological methods such as case-finding or the capture-recapture methodology partly because these studies have been more visible in that many of them have been published in the scientific literature. There are, of course, many other valid instances where prevalence information at the local level has been used to inform policy or practice; ranging from specific local needs assessments to information on local prevalence gathered from social surveys. While such information on the prevalence of drug misuse at the local level may be useful, we have detailed in this section the more prominent literature, much of which has been subject to peer review. It should, however, be remembered that there is no pre-determined correct method of estimating the prevalence of drug misuse at the local level. The capture-recapture method has been applied in many areas and a related publication (EMCDDA, 1998*b*) details some of the methodological issues that need to be considered when applying the method. However, alterations to the basic methodology have been made in many of the local studies showing that local patterns of drug use have to be considered when applying any methodology.

7 Discussion and Conclusions

In this report we have reviewed the scientific literature on using statistical methods to estimate the prevalence of drug use at the local level. One method shows particular applicability to this task, and we have described the literature concerning this methodology, known as capture-recapture, at greater length. A parallel publication (EMCDDA, 1998*b*) presents methodological guidelines to assist those that wish to use the method.

Reuter (1984) notes that ‘numbers without purpose are numbers without quality’. We therefore begin this discussion about the various prevalence estimation techniques by examining the policy relevance.

While Stimson and Judd (1997) examine the relationship between science, policy and drugs strategy, Cohen (1997) discusses, in particular, the relationship between prevalence estimation and policy interests. Stimson details some of the questions that policy makers may pose to those able to estimate the prevalence of drug use, such as those relating to the nature and extent of the problem and whether or not it is changing. Other questions are of interest to policy makers, such as the cost and availability of drugs, however these are outwith the focus of this report.

We have concentrated on drug misuse prevalence at the local level and Cohen (1997) notes ‘It is often at the local level – often the city or local municipality – that most decisions are taken, and where the relevance of good data and understanding for drug policy practice is highest.’ Cohen also notes that good local data is often a prerequisite for reliable national prevalence estimates.

Cohen discusses the need for prevalence estimation under two broad policy models; the repressive model and the harm reduction model. Within the repressive model, prevalence estimates can measure the short term effects of the policy, however any change in the indicators of prevalence, for example heroin related emergency room episodes may not actually reflect changes in the number of drug users, as previously discussed in this report. Under the other model, the acceptance that drug misuse is continuing enables more specific research into the nature of drug use, in particular in relation to harm reduction policies such as needle exchanges. McKeganey (1998) discusses the change in emphasis that occurred in Scotland over the preceding decade, and how prevalence estimates have, to some extent, informed policy decisions.

The discussion of the link between prevalence estimation and policy formulation has most often been staged in the United States, particularly as epidemiological instruments, such as the National Household Survey of Drug Abuse and the Drug Abuse Warning Network are well established. Reuter (1993) discusses the link between prevalence and policy, stating that prevalence estimation has only played a modest role in decision making at the national level, the figures are used within the rhetoric of the national drug policy. Although such instruments are not as established in Europe, the

Member states of the European Union share a concern about the levels of drug use in their countries. However, each member state has different policy reactions to the problem, therefore the policy debate in the United States may not be so pertinent in Europe, where Cohen notes that it is often at the local level that most decisions are made.

So if the policy makers are ready with the questions they wish to pose to prevalence estimators, what are the tools available to help provide answers to these questions? In the preceding sections of this report, we have detailed the methods in isolation, however there are papers, mainly American, where different prevalence estimation methods are contrasted and compared.

Hser (1993*b*) begins her summary of the available methods by re-examining the interaction between policy interests and prevalence estimation, discussing how prevalence estimation should be used for resource allocation and what degree of accuracy is needed for policy makers. She briefly defines the common themes, such as case definitions and what sources of data can be used, then details what prevalence estimation techniques are available and what their data requirements are.

Other papers, such as Hser (1993*a*), Wickens (1993) and Woodward *et al.* (1984) compare and contrast different prevalence estimation techniques, including different variations on the common theme of capture-recapture, however such papers conclude that there is no single best estimation procedure. The suitability of a method depends on the data that are available and the nature of the population whose size is being estimated.

Korf (1997*a*) discusses four types of methods which were used to estimate prevalence in the Netherlands; surveys, extrapolations from sources, case counting and nomination techniques, and capture-recapture. See also Korf *et al.* (1994). Although survey methods may only be of use under certain circumstances, Korf details the results of several studies that were undertaken in The Netherlands, however he concludes that the capture-recapture methodology appears to be the most feasible method to estimate the size of heroin using populations, particularly in larger cities.

It should be remembered that every technique detailed in this report is an simplification of the real-life processes that affect drug using populations. Although the different techniques rely to a greater or lesser extent to a series of assumptions, these assumptions all relate to the fact that drug using populations are diverse. Any attempt to summarise their attributes into a simple model, perhaps by assuming a common mortality rate or a constant probability that they are in contact with a treatment agency, may introduce a level of bias which may render the estimate inaccurate, or even misleading. That is not to say that we should merely give up in our attempts to count drug using populations; we should proceed with caution, assessing the estimates and the related assumptions with care. As LaPorte (1994) notes, 30 to 40 years of work was needed to evaluate the capture-recapture method in estimating

animal populations. Although the epidemiological uses of that, and other statistical techniques, are becoming more and more accepted, there is still some way to go before these methods can be applied in a straightforward manner to each and every drug misuse prevalence estimation problem.

To varying extents, each of the papers detailed above present a discussion about the applicability of the capture-recapture method to estimating the size of a hidden population. Several key themes appear; the heterogeneity of the underlying population, the validity of the closed population assumption, the matching of individuals between data sources and the construction of confidence intervals and model fitting.

Heterogeneity appears to be the most prominent issue that needs to be addressed when using capture-recapture methods. Perhaps to the detriment of the studies, all of the estimates derived from capture-recapture methods assume that those drug users who are hidden from sources such as treatment agencies or the police are similar to those that have been identified, particularly with respect to the seriousness of their drug problem. While authors have sought to address this heterogeneity by stratifying the population by age, gender, or even severity of drug problem, there has been little attempt in this area to employ models specifically designed to included heterogeneity, such as those discussed by Agresti (1994), Evans *et al.* (1994), or even those by Hook and Regal (1993). Indeed the differences in stating the basic assumptions of the capture-recapture model, such as the assumptions listed by Seber (1992), and translated by those such as the International Working Group for Disease Monitoring and Forecasting (IWGDMF, 1995a) or Hook and Regal, are sometimes unclear. Not only is it assumed that each individual in the population has the same probability of being identified from each source, but having been identified within one source does not increase that individuals probability of being identified from another. That is different from dependencies between sources where the same increased, or decreased, probability of being identified in each source is common to the whole population.

It is only in American studies that serious attempts are made to include the fact that drug using populations may be transient and therefore the closed population assumption may be violated. Studies such as Hser's, however, do not appear to be readily applicable in Europe, certainly not in the previous EMCDDA-funded pilot study. It may be that, as some authors claim, the effect of any changes in population size are negligible, but that has not been backed up by any conclusive studies. However, it should be noted that the effect of movement into (and out of) the population will differ between studies which draw samples at different points in time and those which collate samples over a time interval.

It is becoming more and more common for the capture-recapture method to be used to estimate the prevalence of drug misuse at the local level. There are several reasons for this, not least the continuing emphasis that the EMCDDA

and other bodies are placing on the method. The multi-city pilot study that was initiated after the Strasbourg Scientific Seminar demonstrated that the methodology could be applied in a range of areas, stretching from Helsinki in the north to Setúbal in the south, from Dublin in the west to Vienna in the East. The Annual Report on the State of the Drugs Problem in the European Union describes how, in addition to the EMCDDA funded pilot study, local prevalence estimates have been obtained using the methodology in the majority of member states. Indeed the methodology is currently being applied to estimating prevalence at the regional or even the national level.

The scientific literature relating to this methodology, in our particular area of interest and related ones such as research into homelessness or prostitution, is growing. Several review papers were published in the last five years, providing a comprehensive review of the epidemiological applications of the method, however these fail to address some of the more specific problems relating to drug misuse prevalence estimation. There also exists a wider, more theoretical, range of statistical papers which aim to address specific methodological aspects, such as compensating for heterogeneity, however these are often difficult to understand without a thorough grounding in the theory behind log-linear models.

The methodological guidelines contained in a parallel report (EMCDDA, 1998*b*) attempt to fill in the gaps in the literature, by giving a step-by-step guide to implementing the method and discussing the potential problems that may be faced. Topics such as potential data sources, matching between data sources and model fitting are discussed, however it is usually not until a capture-recapture study has commenced that some potential problems come to light; many of which will be specific to the particular study. Care must be taken when resolving these problems that potential for producing valid estimates is not compromised.

Although capture-recapture is becoming more widely used, that specific methodology is far from being a panacea; an easily used portable tool that can be used to churn out prevalence estimates everywhere it is applied. As previously stated within this report, the instances where the method has provided useful estimates, as evidenced by the peer-review process of academic journals, far outweigh the instances where the method has failed. And there can be various reasons why the methodology would fail, either due to the lack of sufficient data or by the failure to fit an adequate model.

As many authors note, capture-recapture methods are only one group in a raft of measures which, under certain circumstances, can be used to estimate the prevalence of drug misuse at the local level. Not all of these have been applied in Europe, and when they have, this has been done in a fragmented manner. Additionally there have been few instances in Europe where more than one estimation technique have been applied in the same area over the same time period.

Some of these techniques have an exciting potential for use in the Europe, needing little more (or in some cases less) data than the three sample capture-recapture method. However each and every method relies to some extent on a series of assumptions, some of which will be readily understood, other will need case studies and practical evaluation in the context of the results that they produce. These methods include multiplier methods, synthetic estimate, factor analysis, latent state Markov models, and to some extent open population capture-recapture methods and the truncated Poisson model. Nomination methods, where fieldwork sampling techniques such as snowball sampling are used to gather information from drug users that are not in contact with agencies, can also provide valuable information which can be used to estimate the prevalence of drug use in some areas.

Although the capture-recapture methodology increasingly appears to be the most commonly used method within the European Union, to gauge the applicability of other methods, there should be research into the comparability of prevalence estimation techniques, not just the statistical methods detailed in this report, but also those that are used to estimate prevalence at the local level and population surveys. These comparisons should be multi-faceted, including the effect of case definitions, the availability of data and expertise needed to undertake a study.

To conclude, this report has described drug misuse prevalence techniques at the local level, and in particular the capture-recapture method. The peer-reviewed scientific literature appears to conclude that this method is the most promising of a range of methods, but the choice of method is often dictated by the availability of data and the use of the resultant estimates. The challenge now is to increase policy makers and prevalence researchers' knowledge of this methodology, and to describe how this methodology relates to the others that show potential.

8 References

- Abeni, D.D., Brancato, G., and Perucci, C.A. (1994) Capture-recapture to estimate the size of the population with Human Immunodeficiency Virus Type-1 Infection. *Epidemiology*, **5**:411-414.
- Agresti, A. (1994) Simple capture-recapture models permitting unequal catchability and variable sampling effort. *Biometrics*, **50**:494-500.
- Bailey, N. (1952) Improvements in the interpretation of recapture data. *Journal of Animal Ecology*, **21**:120-127.
- Bello P, -Y. (1997) Estimates of opiate use in Toulouse conurbation in 1995. In European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) *Methodological pilot study of local prevalence estimates*. Lisbon: EMCDDA.
- Bello, P.-Y., and Chêne, G. (1997) A capture-recapture study to estimate the size of the addict population in Toulouse, France In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*. EMCDDA, Lisbon
- Biernacki, P., and Waldorf, D. (1981) Snowball sampling - Problems and techniques of chain referral sampling. *Sociological Methods & Research*, **10**:141-163.
- Bishop, Y., Fienberg, S., and Holland, P. (1975) *Discrete Multivariate Analysis*. MIT Press, Cambridge MA.
- Bloor, M., Leyland, A., Barnard, M., *et al.* (1991) Estimating hidden populations: a new method of calculating the prevalence of drug-injecting and non-injecting female street prostitution. *British Journal Of Addiction*, **86**:1477-1483.
- Bloor, M., Wood, F., and Palmer, S. (1997) *Estimating the prevalence of injecting drug use and serious drug use in Wales*, University of Cardiff, Cardiff.
- Blumenthal, S., Dahiya, R., and Gross, A. (1978) Estimating the complete sample size from an incomplete Poisson sample. *Journal of the American Statistical Society*, **73**:182-187.
- Bonett, D., Woodward, J., and Bentler, P. (1986) A linear model for estimating the size of a closed population. *British Journal of Mathematical and Statistical Psychology*, **39**:28-40.

Brecht, M.-L., and Wickens, T. D. (1993) Application of multiple-capture methods for estimating drug use prevalence. *The Journal of Drug Issues*, **23**:229-250.

Brenner, H., Stegmaier, C., and Ziegler, H. (1995) Estimating completeness of cancer registration - an empirical- evaluation of the 2 source capture-recapture approach in Germany. *Journal of Epidemiology and Community Health*, **49**:426-430.

Brownie, G., Hisnes, J., and Nichols, J. (1989) Constant-parameter capture-recapture models. *Biometrics*, **42**:561-574.

Brugha, R. F., Swan, A. V., Hayhurst, G. K., *et al.* (1998) A drug misuser prevalence study in a rural English district. *European Journal of Public Health*, **8**:34-36.

Bruno, G., Bargerò, G., Vuolo, A., *et al.* (1992) A population-based prevalence survey of known diabetes-mellitus in Northern Italy based upon multiple independent sources of ascertainment. *Diabetologia*, **35**:851-856.

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

Cates, W., Smith, J., Rochat, R., *et al.* (1978) Assessment of surveillance and vital statistics data for monitoring abortion mortality, United States, 1972-1975. *American Journal of Epidemiology*, **108**:200-206.

Centre for Drug Misuse Research (CDMR) (1997) Estimating the prevalence of drug misuse in Lanarkshire: A report for the Lanarkshire Drug Action Team. Glasgow; CDMR.

Centre for Drug Misuse Research (CDMR) (1998) Estimating the prevalence of opiate and benzodiazepine misuse in Grampian: A report for Grampian Health Board. Glasgow; CDMR.

Chang, Y. F., McMahon, J. E., Hennon, D. L., *et al.* (1997) Dog bite incidence in the city of Pittsburgh: A capture-recapture approach. *American Journal of Public Health*, **87**:1703-1705.

Chao, A. (1989) Estimating population size for sparse data in capture-recapture experiments. *Biometrics*, **45**:427-438.

Cohen, P. (1997) The relationship between drug use prevalence estimation and policy interests. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*. EMCDDA, Lisbon.

Comiskey, C. (1991) Mathematical models for the transmission dynamics of the Human Immunodeficiency Virus (HIV) in Ireland. *Proceedings of the sixth European conference on mathematics in industry*. 125-128. Stuttgart: BG Teubner.

Comiskey, C, Ruskin, H and Wood, A. (1992) Mathematical models for the transmission dynamics of HIV in Ireland. Dublin, Dublin City University.

Comiskey, C. (1997) Estimating the prevalence of opiate drug use in Dublin, Ireland during 1996. In European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) *Methodological pilot study of local prevalence estimates*. Lisbon: EMCDDA.

Cormack, R. (1968) The statistics of capture-recapture methods. *Annual review of Oceanography and Marine Biology*, **6**:455-506.

Cormack, R. M. (1992) Interval estimation for mark-recapture studies of closed populations. *Biometrics*, **48**:567-576.

Cox, S., and Shipley, M. (1997) Counting the uncatchable? An epidemiological method for counting drug misusers. *Social Psychiatry and Psychiatric Epidemiology*, **32**:19-23.

Davis, S. F., Strebel, P. M., Atkinson, W. L., *et al.* (1993) Reporting efficiency during a measles outbreak in New York City, 1991. *American Journal of Public Health*, **83**:1011-1015.

Davoli, M. (1997) Establishing mortality rates from cohort data. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe, EMCDDA Scientific Monograph Series*. EMCDDA, Lisbon.

D'Ippoliti, D (1997) Prevalence estimation of drug addiction in Rome. European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) *Methodological pilot study of local prevalence estimates*. Lisbon: EMCDDA.

Domingo-Salvany, A., Hartnoll, R. L., Maguire, A., *et al.* (1995) Use of Capture-Recapture to Estimate the Prevalence of Opiate Addiction in Barcelona, Spain, 1989. *American Journal of Epidemiology*, **141**:567-574.

Domingo-Salvany A, Hartnoll, R, Maguire, A, Brugal, T, Albertin, P, Caylà, J, Casabone, J and Suelves, J. (1998) Analytical considerations in the use of capture-recapture to estimate prevalence: Case studies of the estimation of opiate use in the metropolitan area of Barcelona. *American Journal of Epidemiology*, **148**, 732-740.

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

Draper, D. (1995) Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society, Series B*, **57**:78-79.

Drucker, E. and Vermond, S. (1989) Estimating population prevalence of human immunodeficiency virus infection in urban areas with high rates of intravenous drug use: a model of the Bronx in 1988. *American Journal of Epidemiology*, **130**, 133-142.

Duque-Portugal, F., Martin, A. J., Taylor, R., *et al.* (1994) Mark-recapture estimates of injecting drug users in Sydney. *Australian Journal of Public Health*, **18**:201-204.

Efron, B., and Thisted, R. (1976) Estimating the number of unseen species: How many words did Shakespeare know? *Biometrika*, **63**:435-47.

el Khorazaty, M., Imrey, P., Koch, G., *et al.* (1977) Estimating the total number of events with data from multiple-record systems: a review of methodological strategies. *International Statistical Review*, **45**:129-157.

Eskild, A., Magnus, P., Samuelson, S., *et al.* (1993) Differences in mortality rates and causes of death between HIV positive and HIV negative intravenous drug users. *International Journal of the Addictions*, **22**:315-320.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). (1997) Methodological pilot study of local level prevalence estimates. EMCDDA, Lisbon.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). (1998a) Annual report on the state of the drugs problem in the European Union. EMCDDA, Lisbon.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (1998b). Methodological Guidelines to Estimate the Prevalence of Problem Drug Use on the Local Level. Lisbon: EMCDDA, December 1998.

Evans, M. A., Bonett, D. G., and McDonald, L. L. (1994) A general-theory for modeling capture-recapture data from a closed population. *Biometrics*, **50**:396-405.

Finn, E. (1997) Socialtjänstens kontakter med missbrukare, psykiskt störda och hemlösa, 1996. FoU-rapport, 15, Stockholm.

Finn, E. (1998) Socialtjänstens kontakter med missbrukare, psykiskt störda och hemlösa, 1997.

Fisher, N., Turner, S. W., Pugh, R., *et al.* (1994) Estimating numbers of homeless and homeless mentally ill people in north east Westminster by using capture-recapture analysis. *BMJ*, **308**:27-30.

Fisher, R., Corbet, A., and Williams, C. (1943) The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*, **12**:42-58.

Frank, O and Snijders, T. (1994) Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics*, **10**:53-67.

Freire, S and Moreira, M (1997) Estimating the number of opiate users in Setúbal. In European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) *Methodological pilot study of local prevalence estimates*. Lisbon: EMCDDA.

Frischer, M. (1992) Estimated prevalence of injecting drug-use in Glasgow. *British Journal of Addiction*, **87**:235-243.

Frischer, M., Leyland, A., Cormack, R., *et al.* (1993) Estimating the population prevalence of injection drug use and infection with human immunodeficiency virus among injection drug users in Glasgow, Scotland. *American Journal Of Epidemiology*, **138**:170-181.

Frischer, M. (1997) Estimating the prevalence of drug abuse using the mortality multiplier method: an overview. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*. EMCDDA, Lisbon.

Godinho J, Costa, H and Padre-Santo, D. (1998) Estimativa da prevalência de consumidores de heroína no Concelho de Setúbal. *Toxicodependências*, **4**, 27-32.

Greene, M. (1984) Estimating the size of a criminal population using an open population approach. *Proceedings of the American Statistical Association, Survey Research Methods Section*, pp8-13.

Greenwood, J. (1971) Estimating the number of narcotic addicts. Bureau of Narcotics and Dangerous Drugs, United States Department of Justice, Washington DC.

Gronbladh, L., Ohlund, L., and Gunne, L. (1990) Mortality in heroin addiction: impact of methadone maintenance. *Acta Psychiatrica Scandinavica*,

82:223-227.

Hartnoll, R., Daviaud, E., Lewis, R., *et al.* (1985a) Drug problems: assessing local needs. A practical manual for assessing the nature and extent of problematic drug use in a community. Drug Indicators Project, London.

Hartnoll, R., Mitcheson, M., Lewis, R., *et al.* (1985b) Estimating the prevalence of opioid dependence. *Lancet*, **i**:203-205.

Hay, G. (1997) The selection from multiple data sources in epidemiological capture-recapture studies. *Statistician*, **46**:515-520.

Hay, G., and McKeganey, N. (1996) Estimating the prevalence of drug misuse in Dundee, Scotland: An application of capture-recapture methods. *Journal of Epidemiology and Community Health*, **50**:469-472.

Hewitt, D., and Milner, J. (1970) Prevalence of hemophilia in Ontario, 1966. *Canadian Medical Association Journal*, **102**:174-177.

Hook, E. B. (1982) Incidence and prevalence as measures of the frequency of birth defects. *American Journal Of Epidemiology*, **116**:743-747.

Hook, E. B., Albright, S. G., and Cross, P. K. (1980) Use of Bernoulli census and log-linear methods for estimating the prevalence of spina bifida in live births and the completeness of vital record reports in New York State. *American Journal of Epidemiology*, **112**:750-758.

Hook, E. B., and Regal, R. R. (1993) Effect of variation in probability of ascertainment by sources ("Variable Catchability") upon "capture-recapture" estimates of prevalence. *American Journal Of Epidemiology*, **137**:1148-1166.

Hook, E. B., and Regal, R. R. (1995a) Capture-recapture methods in epidemiology - methods and limitations. *Epidemiologic Reviews*, **17**:243-264.

Hook, E. B., and Regal, R. R. (1995b) Internal validity analysis: a method for adjusting capture-recapture estimates of prevalence. *American Journal Of Epidemiology*, **142**:S48-S52.

Hser, Y.-I. (1993a) Population estimation of illicit drug users in Los Angeles County. *The Journal of Drug Issues*, **23**:323-334.

Hser, Y. I. (1993b) Prevalence Estimation - Summary of common problems and practical solutions. *Journal of Drug Issues*, **23**:335-343.

International Working Group for Disease Monitoring and Forecasting. (1995a) Capture-recapture and multiple-record systems estimation I: History and

theoretical development. *American Journal Of Epidemiology*, **142**:1047-1058.

International Working Group for Disease Monitoring and Forecasting. (1995b) Capture-recapture and multiple-record systems estimation II: Applications in human diseases. *American Journal Of Epidemiology*, **142**:1059-1068.

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

Karvonen, M., Tuomilehto, J., Libman, I., *et al.* (1993) A review of the recent epidemiologic data on the worldwide incidence of type 1 (Insulin Dependent) Diabetes-Mellitus. *Diabetologia*, **36**:883-892.

Kaukonen, O, Kinnunen, A, Partanen, P, Seppälä, T, Simpura, J and Virtanen, A. (1997) Estimating the number of amphetamine and opiate users in the Greater Helsinki area. In European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) *Methodological pilot study of local prevalence estimates*. Lisbon: EMCDDA.

Kehoe, L., Hall, W., and Mant, A. (1992) Estimates of the number of injecting drug users in a defined area. *Australian Journal of Public Health* (1992) **16**:232-237.

Kirschner, W. (1997) Case-finding, capture-recapture and double counting: A Berlin case study. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*. EMCDDA, Lisbon

Kirschner, W and Kunert, M (1997) Umfang und Struktur von i.v. Drogenabhängigen in Deutschland (1995). Anonymes Monitoring in den Praxen niedergelassener Ärzte. Berlin, EFB.

Korf, D, Hes, J and van Aalderen H. (1991) Waar je mee omgaat. Aids-risico's in de Alkmaarse drugscenes. Alkmaar: Brijder Stichting.

Korf, D. J., Reijneveld, S. A., and Toet, J. (1994) Estimating the number of heroin users - a review of methods and empirical-findings from the netherlands. *International Journal Of The Addictions*, **29**:1393-1417.

Korf, D. (1997a) Comparison of different estimation methods in The Netherlands. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in the European Union*. EMCDDA, Lisbon.

Korf, D. (1997b) The tip of the iceberg: Snowball and nomination techniques, the experience of Dutch studies. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*.

EMCDDA, Lisbon.

Kozinetz, C. A., Skender, M. L., Macnaughton, N., *et al.* (1993) Epidemiology of Rett syndrome - a population-based registry. *Pediatrics*, **91**:445-450.

LaPorte, R. E. (1994) Assessing the human condition: capture-recapture techniques. *BMJ*, **308**:5-6.

LaPorte, R. E., Dearwater, S. R., Chang, Y. F., *et al.* (1995) Efficiency and accuracy of disease monitoring systems - application of capture-recapture methods to injury monitoring. *American Journal Of Epidemiology*, **142**:1069-1077.

Laporte, R. E., McCarty, D. J., Songer, T. J., *et al.* (1993) Disease Monitoring. *Lancet*, **341**:1416.

Larson, A., and Bammer, G. (1996) Why ? Who? How ? Estimating numbers of illicit drug users: Lessons from a case study from the Australian Capital Territory. *Australian and New Zealand Journal of Public Health*, **20**:493-499.

Larson, A., Stevens, A., and Wardlaw, G. (1994) Indirect estimates of "hidden" populations: capture-recapture methods to estimate the numbers of heroin users in the Australian Capital Territory. *Social Science & Medicine*, **39**:823-831.

Laska, E. M., and Meisner, M. (1993) A plant-capture method for estimating the size of a population from a single sample. *Biometrics*, **49**:209-220.

Mariani, F., Guaiana, R., and Di Fiandra, T. (1994) An epidemiological overview of the situation of illicit drug abuse in Italy. *Journal of Drug Issues*, **24**:569-595.

Mastro, T. D., Kitayaporn, D., Weniger, B. G., *et al.* (1994) Estimating the number of HIV infected injection drug users in Bangkok: A capture-recapture method. *American Journal Of Public Health* **84**:1094-1099.

Mcgilchrist, C. A., McDonnell, L. F., Jorm, L. R., *et al.* (1996) Loglinear models using capture-recapture methods to estimate the size of a measles epidemic. *Journal Of Clinical Epidemiology*, **49**:293-296.

McKeganey, N. (1998) Drug misuse in Scotland: Policy, prevalence, and public health. *Journal of Drug Issues*, **28**:91-105.

Newmeyer, J., and Johnson, G. (1976) The heroin epidemic in San Fransisco: Estimates of incidence and prevalence. *The International Journal of the*

Addictions, **11**:417-438.

Newmeyer, J. A. (1988) The prevalence of drug use in San Francisco in 1987. *Journal of Psychoactive Drugs*, **20**:185-189.

O Higgins, K. (1995) *Treated drug misuse in the Greater Dublin area. A review of five years 1990-1994*. Dublin, Health Research Board.

Olson, B. (1997) Problems of definition and other aspects of case finding: the Swedish nationwide studies. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the Prevalence of Problem Drug Use in Europe*. EMCDDA, Lisbon

Olson, O., Byqvist, S., and Gomer, G. (1993) *Det tunga narkotikamissbrukets omfattning i Sverige 1992*. Centralförbundet för alkohol- och narkotikaupplysning, Stockholm.

Oppenheimer, E., Torbutt, C., Taylor, C., *et al.* (1994) Death and survival in a cohort of heroin addicts from London clinics: a 22-year follow up study. *Addiction*, **89**:1299-1308.

Paredes, D, del Llano, del Llano, J, Ruiz, M and Harnoll, R. (1994) Estimación de la prevalencia de la adicción a la heroína en la Comunidad de Madrid durante 1992. Madrid, Comunidad de Madrid, Plan Regional sobre Drogas.

Palfrey, J. S., Haynie, M., Porter, S., *et al.* (1994) Prevalence of Medical Technology Assistance Among Children in Massachusetts in 1987 and 1990. *Public Health Reports*, **109**:226-233.

Parker, H., Newcombe, R., and Bakx, K. (1987) The New Heroin Users: prevalence and characteristics in Wirral, Merseyside. *British Journal of Addiction*, **82**:147-157.

Perucci, C., Davoli, M., Rapiti, E., *et al.* (1991) Mortality of intravenous drug users in Rome: a cohort study. *American Journal of Public Health*, **81**:1307-1310.

Pollock, K. (1991). Modeling capture, recapture and removal statistics for estimation of demographic parameters for fish and wildlife populations - past, present and future. *Journal of the American Statistical Society*, **86**:225-238.

Püschel, K. (1993) Drug related death - an update. *Forensic Science International*, **62**:121-128.

Reuter, P. (1984) The (continued) vitality of mythical numbers. *Public*

Interest, **75**:135-147.

Reuter, P. (1993) Prevalence Estimation and policy formulation. *Journal of Drug Issues*, **23**:167-184.

Rhodes, W. (1993) Synthetic estimation applied to the prevalence of drug use. *Journal of Drug Issues*, **23**:297-321.

Roberts, I., and Scragg, R. (1994) Application of capture-recapture methodology to estimate the completeness of child injury surveillance. *Journal of Paediatrics and Child Health*, **30**:513-514.

Robles, S. C., Marrett, L. D., Clarke, E. A., *et al.* (1988) An application of capture recapture methods to the estimation of completeness of cancer registration. *Journal of Clinical Epidemiology*, **41**:495-501.

Rossmo, D. K., and Routledge, R. (1990) Estimating the size of criminal populations. *Journal of quantitative criminology*, **6**:293-314.

Sarangi, J., Coleby, M., Trivella, M., *et al.* (1997) Prevention of post splenectomy sepsis: a population based approach. *Journal of Public Health Medicine*, **19**:208-212.

Scarabis, H and Patzak, M (1981) Die Berliner heroinszene. Arbeitsergebnisse aus der suchtforschung. Basel, Band 1.

Schouten, L. J., Straatman, H., Kiemeney, L. A. L. M., *et al.* (1994) The Capture - recapture method for estimation of cancer registry completeness - a useful tool. *International Journal of Epidemiology*, **23**:1111-1116.

Seber, G. (1982) *The estimation of animal abundance and related parameters*. Charles Griffin, London.

Seber, G. (1992) A review of estimating animal abundance II. *International Statistical Review*, **60**:129-166.

Seidler, D and Uhl, A. (1997) Estimating the number of opiate users in Vienna. In European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) *Methodological pilot study of local prevalence estimates*. Lisbon: EMCDDA.

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

Shaw, I., Bloor, M., Cormack, R., *et al.* (1996) Estimating the prevalence of

hard-to-reach populations: the illustration of mark-recapture methods in the study of homelessness. *Social Policy and Administration*, **30**:69-85.

Simeone, R. S., Nottingham, W. T., and Holland, L. (1993) Estimating the size of a heroin-using population - an examination of the use of treatment admissions data. *International Journal Of The Addictions*, **28**:107-128.

Simon, R. (1997) Estimating prevalence using the case-finding method: an overview. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*. EMCDDA, Lisbon.

Smit, F., Brunenberg, W., and Van der Heijden, P. (1996) Het schatten van populatiegroottes toepassingen en een voorbeeld. *Tijdschrift Sociale Gezondheidszorg*, **74**:171-6.

Smit, F., van der Heijden, P., and van Gils, G. (1994) Enkele weinig gebruikte methoden om het aantal plegers van misdrijven te schatten. *Tijdschrift voor criminologie*, **36**:96-119.

Smit, F, Toet, J and van der Heijden, P (1997) Estimating the number of opiate users in Rotterdam using statistical models for incomplete count data. In European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) *Methodological pilot study of local prevalence estimates*. Lisbon: EMCDDA.

Squires, N. F., Beeching, N. J., Schlecht, B. J. M., *et al.* (1995) An estimate of the prevalence of drug misuse in Liverpool and a spatial analysis of known addiction. *Journal Of Public Health Medicine*, **17**:103-109.

Stimson, G., and Judd, A. (1997) Estimating the scale and nature of drug problems: the relationship between science, policy and drugs strategy. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*. EMCDDA, Lisbon.

Sutter, R. W., and Cochi, S. L. (1992) Pertussis hospitalizations and mortality in the United States, 1985-1988: Evaluation of the completeness of national reporting. *Journal of the American Medical Association*, **267**:386-391.

Sutter, R. W., Cochi, S. L., Brink, E. W., *et al.* (1990) Assessment of Vital Statistics and Surveillance Data For Monitoring Tetanus Mortality, United-States, 1979-1984. *American Journal of Epidemiology*, **131**:132-142.

Taub, N. A., Lemicstojcevic, N., and Wolfe, C. D. A. (1996) Capture-recapture methods for precise measurement of the incidence and prevalence of stroke. *Journal Of Neurology Neurosurgery And Psychiatry*, **60**:696-697.

Taylor, C. (1997) Estimating the prevalence of drug use using nomination

techniques: an overview. In EMCDDA Scientific Monograph Series, Number 1, *Estimating the prevalence of problem drug use in Europe*. EMCDDA, Lisbon.

ten Den, C, Bieleman, B, de Bie, E and Snippe, J (1995) *Pijn in het hart. Onderzoek naar aard en omvang van de harddrugproblematiek in de stad Utrecht*. Groningen-Rotterdam, Stichting IntraVal.

Urriaga-Domínguez, M, Extramiana-Cameno, E, Moreno-Iribas, C, Ardanaz-Aicua, E and Martínez-Trinidad, M. (1993). Aproximación a la prevalencia de usuarios de heroína en Navarra en 1990. *Gaceta Sanitaria*, **35**, 56-62.

van Brussel *et al.* (1997) Amsterdam, Municipal Health Service

van der Heijden, P.G.M, Zeltermann, D, Engbertsen, G.B.M and van der Leun, J. (in progress) Estimating the number of illegals in the Netherlands with the truncated Poisson regression model.

Watters, J. K. and Biernacki, P. (1989) Targeted sample: options for the study of hidden populations. *Social Problems*, **36**:416-130.

Watts, C. H., Zwi, A. B., and Foster, G. (1995) How to do (or not to do)...Using capture-recapture in promoting public health. *Health Policy and Planning*, **10**:198-203.

Wickens, T. D. (1993) Quantitative methods for estimating the size of a drug-using population. *The Journal of Drug Issues*, **23**:185-216.

Wiessing, LG, Toet, J, Houweling *et al.* (1995) Prevalentie en risicofactoren van HIV-infectie onder druggebruikers in Rotterdam. Bilthoven, RIVM; GGD, Rotterdam

Wilson, R., and Collins, M. (1992) Capture-recapture estimation with a sample of size one using frequency data. *Biometrika*, **79**:543-53.

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

Woodward, J., Bonett, G., and Brecht, M. (1985) Estimating the size of a heroin abusing population using multiple-recapture census. In B. Rouse, N. Kozel, and L. Richards, eds., *Self-report methods of estimating drug use, NIDA Research Monograph*. National Institute on Drug Abuse, Rockville MD.

Woodward, J. A., Retka, R. L., and Ng, L. (1984) Construct validity of heroin

abuse estimators. *International Journal of the Addictions*, 19:93-117.

Zelterman, D. (1988) Robust estimation in truncated discrete distributions with applications to capture-recapture experiments. *Journal of Statistical Planning and Inference*, **18**:225-237.

Zenker, C. Greiser, E. (1998) Erprobungsvorhaben zur Prävalenzschätzung des regionalen illegalen Drogenmißbrauchs und seiner Folgen. Bvremen: Bremer Institut für Präventionsforschung und Sozialmedizin (BIPS),