Submitted to Statistical Science

Estimating the prevalence of drug use using mark-recapture methods

Gordon Hay and Clive Richardson

Liverpool John Moores University and Panteion University of Social and Political Sciences Athens

Abstract. Sparked by the need to inform the response to the spread of HIV/AIDS in drug injecting populations in the 1980s and the desire to base local, national and international responses to tackling drug use in the 1990s on solid epidemiological data, the mark-recapture method has increasingly been used to estimate the prevalence of drug use. Richard Cormack provided support and advice to some of the first United Kingdom and European studies to estimate drug use prevalence in this way. The approach he outlined, using macros that he developed, has led to the mark-recapture method being used to systematically assess the use of drugs such as heroin or other opioids in the United Kingdom and across Europe. We review the development of the method when applied to estimating the size of drug-using populations, including the use of Bayesian methods. We discuss its limitations and various criticisms that have been voiced.

MSC 2010 subject classifications: Primary 62P25 applications to social sciences, 62-07 data analysis; secondary 62J12 generalized linear models.

Key words and phrases: Bayesian methods, drugs, heroin use, mark-recapture, prevalence.

1. INTRODUCTION

1.1 Background

In this paper we describe how the mark-recapture method has been used to estimate the prevalence of problem drug use. We begin by examining what is meant by "problem drug use" and go on to provide a brief review of the literature on mark-recapture estimation of problem drug use, noting a "standard approach" that was used in many early studies, including a series of studies within the United

Gordon Hay is a Reader in Social Epidemiology at the Centre for Public Health, Liverpool John Moores University 15-21 Webster Street, Liverpool, United Kingdom (e-mail: g.hay@ljmu.ac.uk). Clive Richardson is Professor of Applied Statistics in the Department of Economic and Regional Development at Panteion University of Social and Political Sciences, Leof. Andrea Siggrou 136, Kallithea, Athens, Greece (e-mail: crichard@panteion.gr).

Kingdom. We examine the key assumptions associated with the mark-recapture method, particularly how concepts within the ecological application of the method relate to a covert human population. Developing on from the key assumptions, we review some articles that critically assess whether the assumptions are met. We then go on to examine the use of Bayesian methods within mark-recapture estimation of drug use and also describe the use of open-population models in estimating drug-using populations.

The mark-recapture methodology provides a sample-based solution to the problem of estimating the size of a population when a census is infeasible, likely to be unreliable or, by the nature of the population, impossible to carry out. The obstacles to counting which almost always arise in relation to animal populations, also often appear when human populations are the target. Thus the famous Petersen-Lincoln estimator was eventually reinvented in the modern demographic literature in a 1949 paper which bears – like many other innovations – the name of W.E. Deming (Sekar and Deming, 1949), although much earlier applications have been traced within Buckland, Goudie and Borchers (2000). Sekar and Deming faced the problem of estimating the total number of births and deaths in a district where vital registration was incomplete. Their solution was to use a door-to-door survey to correct the registration data in the now familiar way for a two-source mark-recapture analysis. Their work was thus the forerunner of the enormous number of applications in epidemiology which correct for incomplete ascertainment in lists with partial coverage (Hook and Regal, 1995; International Working Group for Disease Monitoring and Forecasting, 1995a,b), as well as the area of evaluating the actual coverage achieved by censuses of (human) populations by means of post-enumeration surveys (Fienberg, 1992). It includes the idea of estimating within sub-populations, partly in order to check the essential assumption of independence between each source, but also as a means of reducing the potential impact of heterogeneity in "capture" probabilities. In this respect it anticipates the applications that concern us in the present paper.

Our focus here is on one important practical application: estimating the sizes of populations of drug users, often within the confines of a single city, by application of the mark-recapture methodology. The same methods are applicable to other "hard to reach" populations, usually with public health implications, such as sex workers (McKeganey et al., 1992). These applications differ from correcting for incomplete ascertainment in that it is unlikely that there exists any list that has some pretension to approaching comprehensive coverage of a well-defined population. They thus come closer in concept to the ecological applications in which a number of samples are drawn from the population. In our context, these "samples" will be sources of information such as lists of drug users who attended treatment services or were arrested by the police within a certain interval of time. The natures of these samples - and indeed of the population itself - raise many issues which demand consideration.

1.2 The Influence of Richard Cormack

We are indebted to Richard Cormack for providing support and advice during the early studies in Scotland, which directly followed the approach to applying the method outlined in Cormack (1985, 1992, 1989). In particular, he generously shared his suite of macros in GLIM4 (Francis, Green and Payne, 1992) to repli-

cate the approach outlined in that paper. Another important paper that informed the earlier applications in the drug use field was Fienberg (1972), along with the more general information about log-linear models in Bishop, Fienberg and Holland (1975). The support that Richard Cormack gave was perhaps a double-edged sword as it enabled researchers to be able to quickly generate mark-recapture estimates once a contingency table, denoting presence and absence from different data sources, had been constructed and translated into the format used by GLIM4, or also SPSS, which was the main statistical package used by non-specialists to carry out mark-recapture analyses. The downside was that the mark-recapture method was now being used by researchers who perhaps did not have sufficient statistical background to either apply anything more than the most standard application of the method, or more worryingly to understand or check whether they were applying the method correctly in the first place. While the same can perhaps be said about any statistical method that is used by non-statisticians, this may be a more prominent issue with the specialist or niche application of the mark-recapture method to estimate drug use prevalence.

The review of drug use applications of mark-recapture methods includes several acknowledgements of Cormack's important contributions to a variety of studies at the time when this area of application was still under development.

2. PROBLEM DRUG USE

It is a basic tenet in statistical investigation that a clear and unequivocal definition of the population of interest is vital. Therefore, any study that aims to estimate the prevalence of problem drug use needs to provide a description of the case definition employed for identifying problem drug use. Measures of drug dependence, such as the International Classification of Diseases (ICD) codes or the Diagnostic and Statistical Manual of Disorders (DSM) diagnostic criteria are not commonly used in sources of data on drug use and therefore would be of little use in this type of prevalence estimation exercise. Rather, the study considers drug use measures that are readily available in data that can systematically be collated across the country. As such, the case definition of the prevalence estimates depends heavily on the case definitions used by the contributing sources. The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) promotes the use of mark-recapture to estimate "high risk drug use" defined as "recurrent drug use that is causing actual harms (negative consequences) to the person (including dependence, but also other health, psychological or social problems) or is placing the person at a high probability/risk of suffering such harms" (EMCDDA, 2013). It is obvious that this is not a watertight, operational case definition as would be expected in an epidemiological study. In fact, "it can be considered a theoretical or conceptual definition [which] implies appropriate flexibility in reporting" and "the main point of these case definitions is to filter out experimental and occasional users who have a lower risk of harms and are not the core population for the assessment of treatment need"(EMCDDA, 2013). The difficulty of defining the population carries over into the difficulty of knowing from which population each sample has been drawn and whether, in fact, they have all been drawn from the same population. In a typical application, the sources (samples) could include lists of clients entering treatment for a drug use problem and lists of people arrested for drug use offences. Do these draw

on the same population? These are not issues that arise uniquely in relation to using mark-recapture methods among drug users, but they are always present in these applications. It could be said that, in our area of application, the usual relationship between population and sample is turned on its head. We do not have a well-defined population from which we plan the sampling. Rather we have samples that happen to be available – usually only three or four – and the population is implied by the nature of these samples. For example, in Greece, estimates have been obtained annually since 2004 by exploiting the fact that the national database of demands for treatment for a drug problem is collected in three parts, not by design but as a consequence of different services' concerns over data protection. Selecting only the clients who declare that their main substance of abuse is a drug such as heroin and treating the three parts as separate sources enables mark-recapture estimation of the population of high-risk drug users in Greece as in Richardson and Antaraki (2015). Consequently we are estimating the size of a pool of self-reported opioid abusers who are liable to seek treatment.

3. A STANDARD APPROACH

Along with recommending a common case definition such as high risk drug use, the EMCDDA also recommends a standard approach to using mark-recapture methods to estimate drug-use prevalence, following that used in early studies that benefited from Cormack's advice.

A standard set of Poisson log-linear models are fitted to the overlap data recording the frequencies of patterns of presence/absence in the various data sources (see example in Table 1). Where there are three sources of data on drug users (the typical minimum number of sources for a mark-recapture study), these models are: the independence model (with no relationships or interactions between data sources), then three different models that each contain just one two-way interaction, three further models that have two such interactions and a "saturated" model that has three two-way interactions. As there is one missing cell (the hidden population we are seeking to estimate) it is not possible to estimate the three-way interaction. Comparing the value of the deviance to the χ^2 distribution could suggest the "best" model and information criteria such as the AIC value are also useful in assessing different models and estimates. An example using four sources is given in Section 4.1, whereas Table 1 gives an example of Greek data from 2012 used in a three-source analysis. The model that includes the S2.S3 interaction (in addition to all three main effects) is a good fit in each age group. The confidence intervals in Table 1 were obtained using the profile likelihood method (Cormack, 1992; Regal and Hook, 1984).

Although models that include covariates such as age group and gender are available, the typical approach in many of the earlier studies to deal with heterogeneity was to carry out separate analyses on stratified overlap tables, not least because separate estimates by age (and by gender) are requested by agencies or organisations such as the EMCDDA. In other circumstances, separate estimates are obtained naturally for different geographical areas. If these must be combined into a total, a question arises of how to construct a confidence interval around the point estimate. One solution is to use bootstrapping methods, such as those used by Gemmell et al. (2004) when summing estimates for ten areas of Greater Manchester.

TABLE 1

2^3 contingency tables (overlap tables) constructed from Greek data on presence/absence of drug users in three treatment sources. "?" denotes the unobserved "hidden" population. The total population is the total observed plus the estimated hidden population obtained from the best-fitting model (S1+S2+S3+S2.S3) in each age group.

Source			Age Group (years)		
S1	S2	S3	15–24	25–34	35–64
Yes	Yes	Yes	1	8	7
Yes	Yes	No	12	118	51
Yes	No	Yes	1	28	18
Yes	No	No	139	1149	1239
No	Yes	Yes	10	74	20
No	Yes	No	143	694	308
No	No	Yes	41	212	124
No	No	No	?	?	?
		Total observed	347	2283	1767
Estimated total population			2273	9595	9136
95% confidence interval			1478–3893	8142–11060	7546–11288

A natural objection to the standard approach described above is that it takes no account of model uncertainty. Model averaging can be applied, as in Hook and Regal (1997) who suggest using weights proprotional to $e^{-AIC/2}$. However, constructing model-averaged confidence intervals is problematic in the frequentist framework (Claeskens and Hjort, 2008). On the other hand, model averaging emerges naturally from a Bayesian analysis of the data (discussed later), which can be carried out very easily using the R program 'conting' (Overstall and King, 2014).

There is debate as to whether fitting the saturated model is appropriate as, by default, it always fits the data exactly. Because its fit cannot be assessed, it is perhaps a leap of faith as to whether relying on the saturated model gives a reliable estimate. In many of the early three-sample studies it can be noted that the saturated model gives a higher estimate of the hidden population than models with fewer interaction terms and this issue also perhaps makes some people try to avoid fitting the saturated model.

There are, of course, assumptions underlying the application of mark-recapture methods to estimating the size of animal or human populations. For our example it is assumed that the population of drug users does not change during the study. Specifically, it is assumed that drug users do not begin to use drugs or stop using drugs and that drug users do not move into or out of the area that is being studied within that time period. Clearly, these assumptions can hold only approximately at best. The shorter the period, the more likely they are to be to be valid, but the data – especially the overlaps on which mark-recapture depends – may be sparse. The standard period of one year may be a reasonable choice but should be examined. Some published studies have included sources with shorter time periods, often in order to increase the number of sources. For example, the same source of information in each trimester of the year was counted as four mark-recapture sources in Domingo-Salvany et al. (1995). The same device has also been used over two-year study periods in order to obtain first estimates in places where data sources are limited (Choi and Comiskey, 2011; Kraus et al., 2011).

It is also assumed that drug users who are in more than one data source are correctly identified as such. In the drug use example, this assumption usually

relates to how accurately overlap cases can be identified when comparing initials and dates of birth, particularly from data sources where accuracy of recording may be less than that seen in, for example, clinical settings.

Although we do not require sources to be statistically independent (except of course in a two-source analysis, which is not under discussion here) because their dependence is handled by interaction terms in the log-linear model, there is the possibility that the services and agencies that provide the data operate in relation to each other. Suppose that presence in one source precluded presence in another. This might be so if one source was prison or if the sources were mutually exclusive forms of treatment. Then clearly the overlap between the two should be a structural zero in the table and a more complicated analysis is required (or more simply, the two forms of treatment could be combined into one source). Another possible relationship is presence in one source implying presence in another as, for example, arrested drug users being sent automatically to a treatment service, in which case the police source would simply be a subset of the treatment source. It is also assumed that the contributing data sources should be representative. In order to meet this assumption, data sources must have equal coverage of the area they serve and also be representative of gender, age group, ethnic group, type and severity of drug use etc. That is not to say that, for example, a treatment service should have equal numbers of female and male clients, rather the probability that a female drug user in the community appears in a treatment data source is similar to that of a male drug user (or that drug users in a rural part of a county are as likely to access treatment as drug users living in a town or city). This is an issue of homogeneity or heterogeneity which is often handled by breaking the data down by gender, age and other groupings.

4. REVIEW OF THE DRUG USE APPLICATIONS OF MARK-RECAPTURE

In this section we review some of the earlier applications of the mark-recapture method in estimating the prevalence of problem drug use. Our focus is on studies carried out within the United Kingdom (beginning with Scotland where Richard Cormack is based and is most influential) which largely followed the "standard approach" described above. We consider the one-off studies carried out in Scotland which led to a series of regular national and local estimates being produced every three years, and go on to describe how prevalence estimation in England also developed from various local studies into a more systematic application of the method (and related methods) on an annual basis. We first review some of the European work in this area.

4.1 European Studies

At the European level, the motivation for many of the earlier drug prevalence estimation studies was to provide information for organisations such as the then newly established EMCDDA of the European Union or the Pompidou Group of the wider Council of Europe. The EMCDDA jointly with the Pompidou Group published a scientific monograph which included four chapters covering the application of mark-recapture methods to problem drug use prevalence estimates (Domingo-Salvany, 1997; Richardson, 1997; Bello and Chêne, 1997; Frischer, 1997). Both of the present authors have been associated with the EM-

CDDA since it was established in 1995, particularly through their national centres responsible for providing information on the issues of drug use in their countries. These National Focal Points structure much of their reporting to the EMCDDA around five Key Indicators, one of which is the size of the population of problem drug users or High-risk Drug Users. Mark-recapture is one of the methods promoted in the guidelines for estimation (EMCDDA, 1999). Thus many EU countries, plus Norway and Turkey, are regularly producing estimates using this methodology.

Important early applications in Europe include studies in Barcelona, Spain (Domingo-Salvany et al., 1995, 1998) and in Dublin, Ireland, where an estimate initially attracted mixed responses (Comiskey, 2001; Comiskey and Barry, 2001) but later seemed to be accepted (Kelly, Carvalho and Teljuer, 2003).

4.2 United Kingdom studies

In the 1980s the threat to public health of the spread of HIV/AIDS became apparent. While initial attention focused on men who have sex with men and people who inject drugs, it was recognised that HIV could spread via heterosexual sex and could impact on groups other than the specific sub-groups it had initially been seen in. In 1985, when a test for HIV became available, blood samples from patients attending an accident and emergency department who were known to be injecting were tested and over 60% were found to be positive (Robertson and Richardson, 2007). This highlighted a public health threat, but without information on the number of drug injectors in Edinburgh, little was known about the total number of people infected with the virus.

As part of the team working within the Glasgow site of a World Health Organization multi-city study on HIV and drug injecting (Des Jarlais, 1994), Martin Frisher became one of the first researchers in Europe to use the mark-recapture method in a drug-using population, specifically to estimate the number of drug injectors in Glasgow (Frischer et al., 1991; Frischer, 1992a; Frischer et al., 1993). Although the title of one of the papers suggested that mark-recapture was a new method in this context, which perhaps overlooked Hartnoll et al. (1985) and Doscher and Woodward (1983), it does expand on the epidemiological grounding of the method. That study can now be considered as ground-breaking, particularly because it was the first to use four data sources and it demonstrated that a successful study could use existing administrative data sources or data that were being collected for other purposes (such as estimating HIV prevalence) and, unlike Hartnoll previously, did not require a lot of additional data collection. It should be noted that Frischer et al. (1993) includes Richard Cormack as an author, as does a later paper by another group that provided estimates for Edinburgh (Davies, Cormack and Richardson, 1999).

While pioneering the application of mark-recapture methods to estimating the prevalence of problem drug use, Frisher was also required to defend the method (Frischer and Leyland, 1992) and respond to criticism (Frischer, 1992b), such as that in the commentary piece by Armstrong and Hayes (1992).

Later on in Scotland, a study set out to build upon the work of Frisher by applying the method in Dundee (Hay and McKeganey, 1996; Hay, 1997). Richard Cormack, although not credited as an author of the paper, provided support and advice to the project team, particularly by allowing free access to his GLIM4

macros (Cormack, 1985). The case definition of opiate and/or benzodiazepine use was applied in an attempt to recognise the "hard drug use" situation in that city at a time when a "heroin drought" had led to people using (including by injection) benzodiazepines. Although that case definition was employed for a specific situation, it has stuck for all subsequent drug use prevalence estimation studies in Scotland, leading sometimes to difficulty in making comparisons to opiate use estimates elsewhere in the UK or Europe. A similar study was carried out in urban and rural areas of the north east of Scotland (Hay, 2000).

These local studies developed into a series of national prevalence estimation exercises on a three-yearly basis that used a common methodological approach stretching from 2000 to 2012 where trends can be examined (Hay and Gannon, 2006; Hay et al., 2009; Information Services Division, 2014).

An important point in the historical development of mark-recapture estimation of the size of drug-using populations in the United Kingdom was Hartnoll et al. (1987), who collated information on drug users in an area of London. That study could be seen as arising from a local initiative which tried to quantitatively assess drug use in a particular area, including applying prevalence estimation methods such as multiplier methods or the two-sample mark-recapture method. While aware of the benefits of using more than two sources of data within a mark-recapture method, the published analyses in Hartnoll et al. (1985) were restricted to the two-sample mark-recapture method. Hartnoll's work could, in part, be seen as responding to the increasing public health concerns about drug use when levels of heroin use in some areas such as London were rising after staying relatively stable since the introduction of the Misuse of Drugs Act in 1971.

Other local estimates in the United Kingdom produced in the 1990s include Squires et al. (1995) who produced an estimate for Liverpool, Beynon et al. (2001) who compared and contrasted estimates across the North West of England (an area that experienced relatively high levels of opiate use in the early 1980s) and Brugha et al. (1998) who applied the method in a more rural area. Elsewhere in the United Kingdom, studies estimated the prevalence of drug use in Northern Ireland (McElrath, 2002; Hay et al., 2006) and in Wales (Bloor, Wood and Palmer, 2000).

In 2002 the UK Government commissioned three methodological pilot studies in England to estimate drug-use prevalence. One used mark-recapture methods in Greater Manchester, updating previous estimates of the prevalence of opiate use (Beynon et al., 2001). Another used mark-recapture methods to estimate the prevalence of opiate use in Liverpool, Brighton and parts of Inner London (Hickman et al., 1992) and a third project used the multivariate indicator method (MIM) to estimate the number of opiate users in England (Frisher, Heatlie and Hickman, 2007). MIM, which was considered by the EMCDDA to be appropriate for estimating the prevalence of opiate use at the national level (Kraus et al., 2003), is essentially a multiple linear regression model which places available prevalence estimates (derived using the mark-recapture method) in a regression model with available "indicator" data such as numbers of opiate users in treatment or other indicators that are thought to be correlated with opiate use.

Following on from these pilot studies, the UK Home Office commissioned a series of annual estimates of the prevalence of opiate use in England, along with estimates of the number of people who use crack cocaine or the number of people

who use either opiates or crack (Hay et al., 2009, 2010, 2011). The intention was to provide mark-recapture estimates for each of the 149 local areas of England, which is the level at which the planning and provision of drug treatment services and other responses to drug use is decided, and at which information on prevalence is particularly needed. While the decision to stratify the analyses at the local area level was more about providing local estimates, it also helped to address the heterogeneity that would be present at the regional or national level. Where the mark-recapture method failed to provide what was subjectively considered to be a valid estimate then MIM would be used to fill in the gaps.

The four data sources on problem drug use which could be used within these mark-recapture analyses were treatment, police, prison and probation data and these were all held centrally at the national level. The available data allowed analyses at the local (n=149), regional (n=9) and national level. Age group and gender could be used to stratify the data. As a key objective of the study was to obtain comparable estimates in order to assess differences between areas and changes over time a standard analysis plan was developed (Hay et al., 2010). Essentially it restricts model fitting to a set of 22 simple models and seeks the "best estimate" for a particular area by considering different combinations of estimates stratified by age group and/or gender. The 22 simple models were the independence model, six models with only one two-way interaction and the fifteen models that include a pair of two-way interactions. Various methods were used to explore whether the model fitted to the unstratified data was a good fit (in particular by considering the AIC value). If none of those approaches provided what was thought to be a valid estimate then MIM was used instead. Estimates for specific age groups were obtained in a similar manner to that outlined in Hay et al. (2009), which involved deriving an estimate for the proportion in each age group from the best available information (typically mark-recapture estimates that employed model-averaging over the simplest 22 models) and applying those proportions to the "best" available estimate which could be a MIM estimate. Confidence intervals for the regional and national estimates were obtained using the approach outlined in Gemmell et al. (2004). This led to extreme asymmetry in the confidence intervals for the age group estimates, most likely due to the large number of confidence intervals that had been approximated by simulated distributions. Table 2 gives an example of the data for an area of London.

Across the years of the English studies, the national estimates stratified by age group suggest that the number of opiate users aged 15 to 24 decreased over time, as did those in the 25 to 34 age range (Table 3). Analyses based on simulation methods confirmed these significant differences. Interestingly the estimates in the older age group increased over time. As the increase each year is similar to the number of opiate users who would turn 35 and therefore move into the older age range, the increases could perhaps be explained as being due to an ageing cohort.

A consistent finding across the successive years of the English studies is that estimates at lower area levels have relatively wide confidence intervals which do not allow significant changes to be identified across time, but adding estimates to get regional or national estimates reveals more consistent trend information. The level of geographical stratification is important. Stratifying to the lowest level possible maximizes the information provided, although the resulting comparatively large confidence intervals may render over-stratified estimates worthless.

TABLE 2
A typical overlap pattern found in an area of London, Year 1 to Year 8; by age group in Year 8

Source				Year 1	Year 2	Year 3	Year 5	Year 6	Year 7	Year 8		
S1	S2	S3	S4	15-64	15-64	15-64	15-64	15-64	15-64	15-24	25-34	35-64
Yes	Yes	Yes	Yes	0	0	0	2	2	1	0	0	0
Yes	Yes	Yes	No	0	0	1	1	0	0	0	0	0
Yes	Yes	No	Yes	1	3	3	7	5	2	1	0	0
Yes	Yes	No	No	2	5	4	19	8	2	0	0	0
Yes	No	Yes	Yes	0	0	0	0	0	1	0	0	0
Yes	No	Yes	No	0	1	1	0	2	0	0	0	0
Yes	No	No	Yes	0	2	1	1	1	0	0	0	0
Yes	No	No	No	4	10	11	12	17	22	8	10	1
No	Yes	Yes	Yes	3	3	16	15	10	15	1	4	3
No	Yes	Yes	No	9	6	11	30	22	19	0	7	9
No	Yes	No	Yes	10	14	29	26	33	37	2	10	21
No	Yes	No	No	280	342	339	342	348	361	25	128	208
No	No	Yes	Yes	3	0	1	5	4	3	0	0	4
No	No	Yes	No	11	13	16	28	31	25	5	12	6
No	No	No	Yes	4	17	14	27	32	32	1	9	6
No	No	No	No	?	?	?	?	?	?	?	?	?

TABLE 3
Table summarising prevalence estimates by age group in England

Year	Age 15–24		Age 25–34		Age 35–64	
	Estimate	95% CI	Estimate	95% CI	Estimate.	95% CI
1	59583	n.a.	124004	n.a.	97740	n.a.
2	51099	(50311 – 54569)	129128	(125614 – 134005)	106264	(104113 – 111433)
3	44398	(43296 – 47493)	118385	(115969 – 123126)	110340	(107905 – 114727)
5	36546	(35193 – 39879)	109509	(106991 – 111604)	116374	(113981 – 118513)
6	35740	(34204 – 37449)	105770	(103035 – 108132)	122563	(119442 – 125031)
7	30278	(28819 – 31980)	100887	(98739 – 103213)	130628	(127847 – 133506)
8	24942	(23475 – 26488)	95950	(93352 – 97923)	135271	(131740 – 137843)

Also, when summing local estimates to get regional or national estimates, it may be beneficial to keep the number of stratified estimates low in order that the larger area estimates do not have over-large confidence intervals. Questions should, however, be asked about the comparability of estimates when some of the mark-recapture analyses are for relatively large geographical areas covering an urban and rural mix and some for relatively small urban areas.

There may have been merit in stratifying large areas into geographical sub-areas to see if geographical heterogeneity was present. There did appear to be an issue in the analyses that the simplest 22 models (as described above) were less likely to provide an adequate fit in the larger geographical areas and consistently throughout the repeated years of the study there were areas where all of the deviance values for the simplest 22 models were far in excess of that found in other areas, perhaps indicating that there could be some breach of the mark-recapture assumptions. There is also the issue that, just because it is possible to get models to fit the data, it does not necessarily mean that the assumptions hold and that the estimates are reliable.

5. METHODOLOGICAL ADVANCES

In this section we examine how the "standard approach" has been further developed in two respects to try to address some of the particular methodological issues in using mark-recapture to estimate drug use prevalence.

5.1 Bayesian approach

Naturally, the Bayesian approach to log-linear modelling of mark-recapture data on drug users has been considered. The best examples can be found in a series of papers concerning the estimation of the number of people who inject drugs in Scotland in different years (King et al., 2005, 2009, 2013) and subsequently in England (King et al., 2014). In marked contrast to the innumerable applications of Bayesian methodology throughout the statistical literature in which only uninformative prior distributions are employed, these papers set out in detail their priors based on expert opinion and results obtained elsewhere.

The advantages of the Bayesian approach include the natural way of averaging across models by weighting according to the posterior probabilities of models that have been examined in the Markov chain Monte Carlo procedure. Similarly, inclusion of a specific interaction can be judged on the basis of a probability rather than as the yes-no decision taken in the usual frequentist model selection. Of course, all this takes place within a specific set of models. The Scottish data form a 2^7 contingency table constructed from four sources and three two-level covariates (age, gender, region) with one missing cell for each combination of the covariate values. In order to keep the number of models to manageable proportions, non-zero prior probabilities were assigned only to the two-way interactions. Consequently, models containing three-way and higher order interactions were excluded. This might be reasonable as far as the sources are concerned, but possibly requires examination for interactions that also include covariates.

An interesting feature of the posterior probability densities of total population size in some analyses was the appearance of bimodality. This can be seen most clearly in the Scottish data for 2003 presented in Figure 1 of King et al. (2009) with a low peak at about 22000 and a higher one around 30000. The greater

posterior probability attached to the model that produced the latter estimate resulted in a final estimate of 27357. The difference between the two peaks was traced chiefly to whether or not the interaction between sources S3 and S4 was included. As the coefficient of the corresponding term in the log-linear model had positive posterior mean, its exclusion would tend to reduce the estimated population sizes. This interaction had a lower posterior probability (0.84) for inclusion than the other terms that appeared to be required, indicating some support for its exclusion and hence the presence of the lower peak. The fuller model containing S3.S4 might not have been identifed within the conventional frequentist approach. There are too many possible models to search through all in order to find the "best" (for example, by smallest value of AIC). However, starting as did King et al. (2009) by excluding three-way and higher order interactions from consideration, the usual process of eliminating and reinserting terms as necessary does lead in the end to their model that contains 12 two-way interactions including S3.S4. However, if one were to start from the model that includes all 3-way interactions, then the need for S1.Age.Gender and S4.Age.Gender would be indicated. Neither seems a priori unreasonable. Eliminating superfluous two-way interactions results in a model with good fit (deviance 113.1 on 96 d.f., $p = 0.11$; AIC 583.4 compared to 582.4 from the model with the 12 2-way interactions). Lacking the S3.S4 interaction, however, it gives a relatively low estimate of population size. This example indicates the difficulty of selecting a final model for a large table. However, one factor that may have played a role could be that the table is relatively sparse, with 24 sampling zeros in its 120 cells; only 46 of the 7201 observations represent people who were observed in three sources and only one person appeared in all four.

5.2 Open populations

The typical mark-recapture analysis for estimating the size of a population of drug users is based on a small number of contemporaneous samples (often only three), using anonymised identification codes in order to match people between samples and estimating the size of the closed population over a short period of time. This automatically excludes drawing upon the vast literature on mark-recapture analysis in other fields, which is largely based on successive sampling, possibly over a long period, a structure that enables – and in fact demands – the consideration of open populations. However, if the mark-recapture data on drug users are collected each year, this raises the possibility of studying an open population by matching across years to construct a capture history as in an ecological application. This will not always be feasible for practical reasons. Even with routine administrative sources, coding the identifiers is sometimes only carried out upon special request for the particular purpose of performing one-off mark-recapture estimation. In other cases, because of anxiety over data confidentiality, the codes are used solely for matching within the year and then deleted, so that they are not available for further use. However, if an agency holds its own data and constructs the codes as a matter of course, an open population analysis over years could be carried out in principle, based on the seminal work of Jolly (1965) and Seber (1965). Among other things, this could show the incidence of starting drug use, complementing the prevalence indicated by the existing annual estimates.

There have been some studies that include an open-population mark-recapture analysis. In comparing different approaches to estimating the size of a drug-using population in Sydney, Australia, Kimber et al. (2008) use a parameterization of the Jolly-Seber model due to Schwarz and Arnason (1996). More recently, using the same parameterization Van Dam-Bates, Fyfe and Cowen (2015) published an analysis of just three years' data for the city of Victoria, Canada. Other proposals for using the Jolly-Seber model date back to a special journal edition of the *Journal of Drug Issues* where Hser (1993) and Brecht and Wickens (1993) consider the same application of the method to data for Los Angeles County (with the latter also using simulated data to compare methods). Both articles concluded that the Jolly-Seber model underestimated the true population size. A more extensive analysis is currently in progress of 13 years of Greek data, using records from only one of the three treatment sources mentioned above as being employed in the annual estimation procedure (codes from the other two sources are not retained because of confidentiality concerns). Although the first results appear to need considerable refinement, especially because of poor fit to individuals who appear repeatedly, they show a high degree of consistency with the annual three-source estimates, which may not be surprising given the use of the same data in both analyses. Unfortunately, the annual incidence ("birth" in the usual terminology of open population analysis), is rather poorly estimated with typical values around 3000-4000 and standard errors of about 700.

One possibly important advantage of the open population analysis is that it only requires a single source; consequently, it can be applied when the minimum requirement of three sources for the closed population mark-recapture analysis is not met. Of course, as is well known, population size for the latest year cannot be estimated except by imposing equal capture probabilities across years, which may be undesirable. As an aside, we mention here that another well-known method for estimating annual population sizes from repeated captures in a single source is sometimes applied to drug users. This is the analysis based on the truncated Poisson distribution for the number of captures using, for example, Zelterman's estimator, see Böhning and Van der Heijden (2009). This method has been used in a number of settings, such as estimating the number of drug users in Bangkok, Thailand (Böhning et al., 2004) and estimating the number of injectors in Aberdeen using data from a needle exchange (Hay and Smit, 2003).

6. DISCUSSION

We have demonstrated that the mark-recapture method can be used to estimate the prevalence of problem drug use both at the local level and, usually in conjunction with other statistical methods, at the national level. Thus Scotland and England have acquired series of prevalence estimates following the same approach (originally based on the advice and support of Richard Cormack) applied systematically across time. These studies have identified that the prevalence of problem drug use in Scotland has remained relatively constant, whereas in England it appears to be decreasing, particularly amongst younger people. Across Europe various local estimates, again obtained using mark-recapture, are informing the work of the EMCDDA.

It has, however, become increasingly difficult to carry out a mark-recapture study, not because of the methods (or even changes in the nature of the data),

but due to increased concerns about sharing the type of personal data (such as initials and dates of birth) required to identify overlaps. Even in countries where all citizens have unique identification codes – for example Denmark where mark-recapture has been used to estimate the prevalence of hepatitis (Christensen et al., 2012; Hansen et al., 2014) – issues relating to data protection legislation and client or patient confidentiality are becoming more and more prominent. While it is possible to construct unique numeric codes to anonymise or "pseudonymise" personal data, since these are based on personal data the data protection concerns still remain (along with all the attendant data security issues such as safe storage and transfer). A perfect way of anonymising the personal data for matching should not detract from the ability to carry out a mark-recapture study, but often this can introduce an added layer of complexity into the analysis or prevent identification of problems with the data.

Data providers such as the police or treatment services usually consider the data they are asked to provide as personal and confidential and often seek to restrict use to the purpose for which it was requested. While it can be argued that the overlap data are sufficiently removed from personal data (particularly when cell values are large enough to prevent deductive disclosure), data providers sometimes place restrictions on the dissemination and publication of such data. The peer-review process of Hay et al. (2009) requested that the original overlap patterns be supplied in an appendix, however permission to do this was not forthcoming from the contributors of data and even the published information about which models were fitted to the overlap data was not for named areas. Although for a statistician it would be good practice to share source data, it has not always been possible with the data collected during the various United Kingdom mark-recapture studies, the notable exception being the studies that applied Bayesian mark-recapture methods to the same data in England and Scotland; this involved the original study team preparing overlap patterns in the required format (such as aggregating up to regional level).

There has been criticism of the "standard approach" used within the United Kingdom studies. In part this criticism is justified as the behaviours of human populations and their involvement with criminal justice services and drug treatment agencies are complex and it would be over-optimistic to believe that the mark-recapture assumptions are fully valid. Since the increase in the use of the method to estimate the size of drug-using populations in the 1980s and 1990s there has been criticism such as Waters (1994) who felt more validation of the method was required in response to LaPorte (1994) who suggested that mark-recapture could bring about a "paradigm shift in how counting is done in all the disciplines that assess human populations".

The validity of the English estimates was also questioned by Frisher and Forsyth (2009) who pointed out that an increase in prevalence between 2001 and 2004 was not matched by other available data such as the number of drug-related deaths or the number of hospital admissions due to drug abuse.

More recently Jones et al. (2014) suggest that standard mark-recapture methods may be fallible when used to estimate the size of covert human populations, particularly in the presence of referrals between data sources such as when the police or a prison refer people into drug treatment. They suggest that this issue is not quite the same as the main assumptions that are routinely considered when

using mark-recapture. Their approach to trying to demonstrate this proposed fallibility is based on highlighting differences between the results of their analyses and the estimates that had been previously published using the same data in Hay et al. (2011), which carried out the analyses at the local area level and often stratified by age group or gender (as well as area of residence) as it was thought unlikely that all drug users across England demonstrated equal "capture probability". Jones et al. (2014) did not consider geographical, age or gender heterogeneity as they carry out their analyses only with data aggregated up to the national level. Their simulated datasets are similarly only analysed at the national level.

The use of a second method in some areas was not noted in the comparisons. This is an important omission as MIM was used in the English analyses when the mark-recapture method failed to provide a valid estimate when fitting the 22 simplest models. This could be because these models do not fit the data due to remaining heterogeneity, dependencies between data sources or indeed the referral issue they are proposing. It could also be due to data quality issues at the local level, something that is not unknown when analysing data on covert populations such as drug users. It is often seen that fitting more complicated models increases the estimated population size, which could be a true reflection of prevalence or an inappropriate attempt at dealing with unreliable data. Fitting over-complicated models at the national level could therefore be one of many reasons why the comparisons they make may not be valid. The same can be said about the Bayesian analyses that always fit mark-recapture models, albeit at the regional level, even when it has been difficult to fit models at the lower area level.

Finally the criticism of the "standard approach" is based on various three-sample analyses, some of which involve fitting less commonly used non-hierarchical log-linear models. They cite Richard Cormack's description of fitting a three-way interaction when only three sources are available, as a 'leap of faith' (Cormack, 1999). In the three-source scenario it usually is indeed a leap of faith, which is one of the reasons why the English series of annual analyses avoided using only three sources. However when using four sources, fitting a four-way interaction is less of a leap of faith when the available data suggest that the leap from two-way to three-way interactions is not required.

We have presented in some detail the use of mark-recapture methodology to estimate the sizes of populations of drug users. This is possibly the most common application of the method in the social sciences. Much of this work is carried out at a relatively unsophisticated level by the standards of modern statistical modelling, a situation which could and should be improved. However, it is our impression that there is a general feeling that the quality of the data that are available in these applications is not adequate to support more than basic analyses. Mark-recapture samples often comprise what already exists or can be extracted from overworked and understaffed services, rather than what would be employed in an ideal world. Practical problems often dominate. Some of these may be the same in nature, although not necessarily in degree, as in other areas of application. For example, imperfect matching across samples may occur in other epidemiological analyses or in ecological studies, but may be more prominent when the use of anonymised identification codes is a necessity and some of the matching information is provided by people who may have low trust in services and could be

under the influence of drugs.

The major difficulty of case definition has been mentioned earlier, as well as the difficulty of ensuring that the sources employ the same definition and thus can be regarded as drawn from the same population. Often, this will not be true. Police records and interviews with trained psychologists in a treatment service are unlikely to agree on the details of an individual's status as a drug user (substances used, injecting behaviour). The hope is that they agree sufficiently, although assessing the validity of this pious hope is difficult and perhaps rarely undertaken. It is worth mentioning, though, some methodological innovations that address the issue of incomplete agreement between sources. Van der Heijden, Zwane and Hessen (2009) treat the problem of lists that refer to non-identical (though overlapping) regions, time periods and age or other groups by applying the EM algorithm to the analysis of the incomplete tables formed from these lists. The same authors, and also Zwane and Van der Heijden (2007) and Zwane and Van der Heijden (2008) consider the similar issue raised by non-identical sets of covariates associated with each source. Overstall et al. (2014), continuing their analyses of Scottish data, presented a method of allowing for a broader definition in one of the four sources than in the other three. Specifically, the hepatitis C virus diagnosis database did not distinguish past from current drug injectors which made its use in the estimation of the number of people who currently inject drugs problematic. This broader definition implies that the contingency table cell counts referring to people who were present only in that database and not the others are left-censored. An extension of previous models to include this feature allowed this source's inclusion in the modelling.

ACKNOWLEDGEMENTS

The authors would like to thank our colleagues within the EMCDDA and the network of National Focal Points for their continuing contribution to the debate surrounding the application of the mark-recapture method to estimating drug prevalence, particularly Lucas Wiessing, Richard Hartnoll, Antònia Domingo-Salvany, Ludwig Kraus, Catherine Comiskey, Filip Smit and Colin Taylor. We thank the EMCDDA for occasional financial support in order to execute research projects in this area and our respective National Focal Points for their continued support over many years. Gordon Hay would like to thank Maria Gannon, Jane Casey, Sharon Hutchinson, Islay Gemmell, Tim Millar, Ruth King and Sheila Bird for their contribution to various studies and to the wider development and application of the method. Gordon Hay would especially like to thank Richard Cormack for his support in the early Scottish studies, particularly allowing the use of the macros he developed.

REFERENCES

Armstrong, J.R.M. and Hayes, R.J. (1992). Estimating prevalence of injecting drug use in an urban population: limitations of the three-sample estimation procedure. *Int J. Epidemiol.* **21** (3) 613.

Bello, P-Y. and Chêne, G. (1997). A capture-recapture study to estimate the size of the addict population in Toulouse, France. In *Estimating the Prevalence of Problem Drug Use in Europe*. EMCDDA, Lisbon.

Beynon, C., Bellis, M., Millar, T., Meier, P., Thomson, R. and Jones, K. (2001). Hidden need for drug treatment services, measuring levels of problematic drug use in the North West of England. *J. Public Health Med.* **23**(4) 286–291.

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

Bloor, M., Wood, F. and Palmer, S. (2000). Use of mark-recapture techniques to estimate the size of hard-to-reach populations. *J. Health Serv. Res. Policy* **5**(2) 89–95.

Böhning, D., Suppawattanabodee, B., Kusolvisitkul, W. and Viwatwongkasem, C. (2004). Estimating the number of drug users in Bangkok 2001: A capture-recapture approach using repeated entries in one list. *Eur. J. Epidemiol.* **19**, 1075–1083.

Böhning, D. and Van der Heijden, P.G.M. (2009). A covariate adjustment for zero-truncated approaches to estimating the size of hidden and elusive populations. *Ann. Appl. Stat.* **3** 595–610.

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

Brugha, R.F., Swan, A.V., Hayhurst, G.K. and Fallon, M.P (1998). A drug misuser prevalence study in a rural English district. *Eur. J. Public Health* **8** 34–36.

Buckland, S.T., Goudie, I.B.J and Borchers, D.L. (2000). Wildlife population assessment: Past developments and future directions. *Biometrics* **56** 1–12.

Choi, Y.H. and Comiskey, C.M. (2011). Methods for providing the first prevalence estimates of opiate use in Western Australia. *Int. J. Drug Policy* **14** 297–305.

Christensen, P.B., Hay, G., Jepsen, P., Omland, L.H., Just, S.A., Karup, H.B., Weis, N., Obel, N. and Cowan, S. (2012). Hepatitis C prevalence in Denmark-an estimate based on multiple national registers. *BMC Infect. Dis.* **12** 178.

Claeskens, G. and Hjort, N.L. (2008). *Model Selection and Model Averaging*. Cambridge University Press, Cambridge.

Comiskey, C.M. (2001). Methods for estimating prevalence of opiate use as an aid to policy and planning. *Subst. Use Misuse* **36** 131–151.

Comiskey, C.M. and Barry, J.M. (2001). A capture recapture study of the prevalence and implications of opiate use in Dublin. *Eur. J. Public Health* **11**(2) 198–200.

Cormack, R.M. (1985). Examples of the use of GLIM to analyse capture-recapture studies. In *Statistics in Ornithology*. (B.J.T. Morgan and P.M. North, eds.) 243–273, Springer-Verlag, New York.

Cormack, R.M. (1992). Interval estimation for mark-recapture studies of closed populations. *Biometrics* **48**(2) 567–576.

Cormack, R.M. (1989). Log-linear models for capture-recapture. *Biometrics* **45** 395–413.

Cormack, R.M. (1999). Problems with using capture-recapture in epidemiology: an example of a measles epidemic. *J. Clin. Epidemiol.* **52**(10) 909–914.

Davies A.G., Cormack, R.M. and Richardson, A.M. (1999). Estimation of injecting drug users in the City of Edinburgh, Scotland, and number infected with human immunodeficiency virus. *Int J. Epidemiol.* **28**(1) 117–121.

Des Jarlais, D.C. (1994). Cross-national studies of AIDS among injecting drug users. *Addiction* **89** 383–392.

Domingo-Salvany, A. (1997). Estimating the prevalence of drug use using the capture-recapture method: an overview. In *Estimating the Prevalence of Problem Drug Use in Europe.* EMCDDA, Lisbon.

Domingo-Salvany, A., Hartnoll, R.L., Maguire, A., Suelves, J.M. and Anto, J.M. (1995). Use of capture-recapture to estimate the prevalence of opiate addiction in Barcelona, Spain, 1989. *Amer. J. Epidemiol.* **141**(6) 567–574.

Domingo-Salvany, A., Hartnoll, R.L., Maguire, A., Brugal, M.T., Albertin, P., Caylà, J.A., Casabona, J. and Suelves, J.M. (1998). Analytical consideration in the use of capture-recapture to estimate prevalence: Case studies of the estimation of opiate use in the metropolitan area of Barcelona, Spain. *Amer. J. Epidemiol.* **148**(8) 732–740.

Doscher, M.L. and Woodward, J.A. (1983). Estimating the size of subpopulations of heroin users–applications of log-linear models to capture-recapture sampling. *Int. J. Addict.* **18**(2) 167–182.

European Monitoring Centre for Drugs and Drug Addiction. (1999). *Methodological guidelines to estimate the prevalence of problem drug use on the local level.* EMCDDA, Lisbon.

European Monitoring Centre for Drugs and Drug Addiction. (2013). *PDU (problem drug use) revision summary.* EMCDDA, Lisbon.

Francis, B.J., Green, M. and Payne, C.P. (1992). *The GLIM4 Manual*. Oxford University Press, Oxford.

Fienberg, S.E. (1972). The multiple recapture census for closed populations and incomplete 2^k contingency tables. *Biometrika* **59**(3) 591–603.

Fienberg, S.E. (1992). Bibliography on capture-recapture modelling with application to census undercount adjustment. *Surv. Methodology.* **18** 143–154.

Frischer, M. (1992a). Estimated prevalence of injecting drug-use in Glasgow. *Br. J.Addict.* **87**(2) 235–243.

Frischer, M. (1992b). Estimating prevalence of injecting drug use in an urban population: limitations of the three-sample estimation procedure – reply. *Int J. Epidemiol.* **21** (3) 613–614.

Frischer, M. (1997). More complex capture-recapture models: an illustrative case study using data from Glasgow, Scotland. In *Estimating the Prevalence of Problem Drug Use in Europe.* EMCDDA, Lisbon.

Frischer, M. and Leyland, A. (1992). Reliability of population and prevalence estimates. *Lancet* **339**(8799) 995.

Frischer, M., Bloor, M., Finlay, A., Goldberg, D., Green, S., Haw, S., McKeganey, N. and Platt, S. (1991). A new method of estimating prevalence of injecting drug use in an urban population: results from a Scottish city. *Int. J. Epidemiol.* **20**(4) 997–1000.

Frischer, M., Leyland, A., Cormack, R., Goldberg, D., Bloor, M., Green, S., Taylor, A., Covell., R, McKeganey, N. and Platt, S. (1993). Estimating the population prevalence of injection drug use and infection with human immunodeficiency virus among injection drug users in Glasgow, Scotland. *Amer. J. Epidemiol.* **138**(3) 170–181.

Frisher, M. and Forsyth, A. (2009). Assessing the validity of recent estimates of problematic drug use in England. *J. Epidemiol. Commun. Health* **63**(1) 87–88.

Frisher, M., Heatlie, H. and Hickman, M. (2007). Validating estimates of problematic drug use in England. *BMC Public Health* **7:286**.

Gemmell, I., Millar, T. and Hay, G. (2004). Capture-recapture estimates of problem drug use and the use of simulation based confidence intervals in a stratified analysis. *J. Epidemiol. Commun. Health* **58** 758–765.

Hansen, N., Hay, G., Cowan, S., Jepsen, P., Krarup, H.B., Obel, N., Weis, N. and Christensen, P.B. (2014). Hepatitis B prevalence in Denmark an estimate based on nationwide registers and a national screening project. *Eurosurveillance* **18**(47) 9–16.

Hartnoll, R.L., Lewis, R., Mitcheson, M. and Bryer, S. (1985). Estimating the prevalence of opioid dependence. *Lancet* **325**(8422) 203–205.

Hartnoll, R.L., Daviaud, E., Lewis, R. and Mitcheson, M. (1987). *Drug problems: assessing local needs. A practical manual for assessing the nature and extent of problematic drug use in a community.* London, Drug Indicators Project, Birkbeck College.

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

Hay G. (2000). Capture-recapture estimates of drug misuse in urban and non-urban settings in the north east of Scotland. *Addiction* **95**. 1795–1803.

Hay, G. and Gannon, M. (2006). Capture-recapture estimates of the local and national prevalence of problem drug use in Scotland. *Int. J. Drug Policy* **17** 203–210.

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

Hay, G. and Smit, F. (2003) Estimating the number of drug injectors from needle exchange data. *Addiction Res. Theory* **11**, 235–243.

Hay, G., Gannon, M., Casey, J. and McKeganey, N. (2009). *Estimating the national and local prevalence of problem drug use in Scotland.* Information Services Division, Edinburgh.

Hay, G., Gannon, M., Casey, J. and Millar, T. (2011). *Estimates of the prevalence of opiate and / or crack cocaine use 2009/2010 Sweep 6 Report.* National Treatment Agency for Substance Use, London.

Hay, G., Higgins, K., Gannon, M, and Carroll, C. (2006). *Prevalence of problem opiate and problem cocaine use in Northern Ireland.* DHSSPS, Belfast.

Hay, G., Gannon, M., MacDougall, J., Eastwood, C., Williams, K. and Millar, T. (2009). Capture-recapture and anchored prevalence estimation of injecting drug users in England: National and regional estimates. *Stat. Methods Med. Res.* **18**(4) 323–339.

Hay, G., Gannon, M., MacDougall, J., Eastwood, C., Williams, K. and Millar, T. (2010). Opiate and crack cocaine use: A new understanding of prevalence. *Drugs: Educ. Prev.*

Polic. **17**(2) 135–147.

Hickman, M., Cox, S., Harvey, J., Howes, S., Farrell, M., Frischer, M., Stimson, G., Taylor, C. and Tilling, K. (1992). Estimating the prevalence of problem drug use in inner London: a discussion of three capture-recapture studies. *Addiction* **94**(11), 1653–1662.

Hook, E.B. and Regal, R.R. (1995). Capture-recapture methods in epidemiology–methods and limitations. *Epidemiol. Rev.* **17** 243–264.

Hook, E.B. and Regal, R.R. (1997). Validity of methods for model selection, weighting for model uncertainty, and small sample adjustment in capture-recapture estimation. *Amer. J. Epidemiol.* **145**(12) 1138–1144.

Hser, Y.I. (1993). Population estimation of illicit drug users in Los Angeles County. *J. Drug Issues* **23** 323–334.

Information Services Division (2014). *Estimating the national and local prevalence of problem drug use in Scotland 2012/13.* Information Services Division, Edinburgh.

International Working Group for Disease Monitoring and Forecasting (1995a). Capture-recapture and multiple-record systems estimation I: History and theoretical development. *Amer. J. Epidemiol.* **142** 1047–1058.

International Working Group for Disease Monitoring and Forecasting (1995b). Capture-recapture and multiple-record systems estimation II: Applications in human diseases. *Amer. J. Epidemiol.* **142** 1059–1068.

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

Jones, H.E., Hickman, M. Welton, N.J. De Angelis, D. Harris, R.J. and Ades, A.E (2014). Recapture or precapture? Fallibility of standard capture-recapture methods in the presence of referrals between sources. *Amer. J. Epidemiol.* **179**(11) 1383–1393.

Kelly, A., Carvalho, M. and Teljeur, C. (2003).*Prevalence of opiate use in Ireland 2000-2001. A 3–source capture-recapture study* National Advisory Committee on Drugs and Alcohol, Dublin.

Kimber, J, Hickman, M., Degenhardt, L., Coulson, T. and van Beek, I. (2008). Estimating the size and dynamics of an injecting drug user population and implications for health service coverage: comparison of indirect prevalence estimation methods. *Addiction* **103** 1604–1613.

King, R., Bird, S.M., Brooks, S.P., Hutchinson, S.J. and Hay, G.(2005). Prior information in behavioural capture-recapture methods: Demographic influences on drug injectors' propensity to be listed in data sources and their drug-related mortality. *Amer. J Epidemiol.* **162** 694–703.

King, R., Bird, S.M., Hay, G. and Hutchinson, S.J. (2009). Estimating current injectors in Scotland and their drug-related death rate by sex, region and age-group via Bayesian capture-recapture methods. *Stat. Methods Med. Res.* **18** 341–359.

King, R., Bird, S.M., Overstall, A., Hay, G. and Hutchinson, S.J. (2013). Injecting drug users in Scotland, 2006: Listing, number, demography and opiate-related death rates. *Addiction Res. Theory* **21** 235–246.

King, R., Bird, S.M., Overstall, A., Hay, G. and Hutchinson, S.J. (2014). Estimating prevalence of injecting drug users and associated heroin-related death rates in England by using regional data and incorporating prior information. *J. Roy. Statist. Soc. Ser. A* **177** 209–236.

Kraus, L., Augustin, R., Frischer., M, Kümmler, P., Uhl, A. and Wiessing, L. (2003). Estimating prevalence of problem drug use at national level in countries of the European Union and Norway. *Addiction* **98**(4) 471–485.

Kraus, L., Hay, G., Richardson, C., Yarğic, Í., Ílhan, M., Ay, P., Civil, F., Pinarci, M., Tunçoğlu, T. and Piontek, D. (2011). *Estimating problem drug use in Ankara, Istanbul and Ízmir. Final Report on Contract CC.11 IPA3.080* EMCDDA, Lisbon.

LaPorte R.E. (1994). Assessing the human condition – capture-recapture techniques. *Br. Med. J.* **308**(6920) 5–6.

McElrath, K. (2002). *Prevalence of problem heroin use in Northern Ireland.* DHSSPS, Belfast.

McKeganey N., Barnard M., Leyland A., Coote, I. and Follet, E. (1992). Female streetworking prostitution and HIV infection in Glasgow. *Br. Med. J.* **305** 801–804.

Overstall, A. and King, R. (2014). conting: An R package for Bayesian analysis of complete and incomplete contingency tables. *J. Statist. Software* **58**(7).

Overstall, A., King, R., Bird, S.M., Hutchinson, S.J. and Hay, G (2014). Incomplete

contingency tables with censored cells with application to estimating the number of people who inject drugs in Scotland. *Stat. Med.* **33**(9) 1564–1579.

Regal, R.R. and Hook, E.B. (1984). Goodness-of-fit based confidence intervals for estimates of the size of a closed population. *Stat. Med.* **3**(3) 287–291.

Richardson, C. (1997). Capture-recapture methodology: lessons from studies on animal populations. In *Estimating the Prevalence of Problem Drug Use in Europe*. EMCDDA, Lisbon.

Richardson, C. and Antaraki, A. (2015). High risk drug users in Greece: estimating the size of the older population. *Int J Aging Soc.* **5** 71–77.

Robertson, R. and Richardson, A. (2007). Heroin injecting and the introduction of HIV/AIDS into a Scottish city. *J. Roy. Soc. Med.* **100**(11) 491–494.

Schwarz, C.J. and Arnason, A.N. (1996). A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics* **52** 860–873.

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

Sekar, C.C. and Deming, W.E. (1949). On a method of estimating birth and death rates and the extent of registration. *J. Amer. Stat. Assoc.* **44** 101–115.

Squires, N.F., Beeching, N.J., Schlecht, B.J.M. and Ruben, S.M. (1995). An estimate of the prevalence of drug misuse in Liverpool and a spatial analysis of known addiction. *J. Public Health* **17**(1) 103–109.

Van Dam-Bates, P., Fyfe, M. and Cowen, L.E. (2015). Applying open population capture-recapture models to estimate the abundance of injection drug users in Victoria, Canada. *J. Subst. Use* in press.

Van der Heijden, P.G.M., Zwane, E. and Hessen, D. (2009). Structurally missing data problems in multiple list capture-recapture data. *AStA Adv. Stat. Anal.* **93** 5–21.

Waters. W.E. (1994). Capture–recapture techniques – more unreliable in humans than birds. *Br. Med. J.* **308**(6927) 531.

Zwane, E.N. and Van der Heijden, P.G.M (2007). Analysing capture-recapture data when some variables of heterogeneous catchability are not collected or asked in all registrations. *Stat. Med.* **26**(5) 1069–1089.

Zwane, E.N. and Van der Heijden, P.G.M (2008). Capture-recapture studies with incomplete mixed categorical and continuous covariates. *J. Data Sci.* **6** 557–572.