

Methods for providing the first prevalence estimates of opiate use in Western Australia

Y.H. Choi^{a,*}, C.M. Comiskey^{b,1}

^a *Biomathematics Unit, Rothamsted Research, Harpenden, Hertfordshire AL5 2JH, UK*

^b *Department of Mathematics, National University of Ireland, Maynooth Co., Kildare, Ireland*

Received 18 October 2002; received in revised form 1 April 2003; accepted 24 April 2003

Abstract

International bodies including the United Nations, the World Health Organisation and the European Monitoring Centre for Drugs and Drug Addiction have highlighted the increasing problem of opiate use and have recommended the truncated Poisson methods and capture–recapture methods, with more than two data sources, for prevalence estimation. This recommendation is difficult to implement however in regions where data sources are limited. The truncated Poisson method, which requires a single data source, is less publicised but is also endorsed by the EMCDDA. Within this paper, we implement both of these methods in Western Australia for the first time. We provide the first multi-source enumeration of opiate prevalence for this region, and in addition develop the two methods for use with in-patient hospital data and police records. We found from the multi-source enumeration that a total of 2826 unique individuals were identified as opiate users in the 2-year study period. In addition, using unstratified capture–recapture models we estimated that the population of hidden opiate users was 17,233 (95% CI of 8960–35,055) during the same period. This gives us a total 2-year prevalence estimate of 20,059 opiate users in Western Australia in the period from the 1st July 1996 to 30th June 1998. This corresponds to an estimated prevalence rate of approximately 18 per 1000 of the population aged between 14 and 55. The need for these methods in the coming decade will increase as we see expansion of the European Union and the increased reporting of drug use in developing and emerging nations.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Prevalence estimation; Opiate use; Truncated Poisson; Capture–recapture; Police data; Hospital inpatient records

Introduction

International bodies including the United Nations (UNDCP, 1997), the World Health Organisation (WHO, 1998) and the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA, 1997) have in recent years highlighted the increasing problem of opiate use in urban and rural communities. Opiate use is a hidden phenomenon and users remain anonymous in the community until they seek treatment or come to the attention of the police. Of these three international organisations, the EMCDDA (2000a) in conjunction with funding from the European Union has endeavoured to develop methodological guidelines that provide a point estimate of the hidden prevalence

of opiate use. The objective of the EMCDDA is the application of this methodology across Europe to enable comparisons and identify nations, localities and cities at high risk, with a view to subsequent targeting with social and medical policy interventions and planning. In a pilot study of five European cities (EMCDDA, 1997), and subsequently in Ireland (Comiskey, 2001) authors report the use of the capture–recapture methodology for the estimation of opiate use at the local level. In each case the use of the methodology is highly recommended, however this recommendation is made with the proviso that a minimum of three data sources be used. This recommendation, whilst both sensible and appropriate is difficult to implement, particularly in regions where data sources are limited or non-existent as in some developing countries or in regions where data protection laws are overly strict.

In this study, we implement both the capture–recapture methodology and truncated Poisson method as suggested by the EMCDDA. However, we develop the capture–recapture method further by using what appears to be four samples

* Corresponding author. Tel.: +44-1582-763133x2402; fax: +44-1582-467116.

E-mail addresses: yoan.choi@bbsrc.ac.uk (Y.H. Choi), cc@maths.may.ie (C.M. Comiskey).

¹ Tel.: +353-1-708-3994; fax: +353-1-708-3913.

but is in effect only two data sources over a 2-year period to provide a 2-year point prevalence estimate. The data requirements for this approach are considerably lower than those required by the three-sample model but the necessary trade-off is the incorporation of additional difficulties with the assumptions of the modelling process. We apply our methodology to the problem of estimating the prevalence of opiate use in Western Australia and in so doing produce the first ever estimates specifically derived for the hidden prevalence of opiate use in this particular region of the world.

Knowing the number of illicit drug users in a community is an essential measure to plan and evaluate the prevention and treatment programs that are available to people using illicit drugs. It is also an important measure to ensure and evaluate the provision of essential law enforcement services and to protect the community against crimes related to illicit drug use.

The aim of the study is two-fold. First, to provide an estimate of the total number of individuals involved in criminal activity related to heroin use in Western Australia (WA) and second, to provide an estimate of the total number of opiate users in the same region.

Methods

Truncated Poisson method

Truncated Poisson methods may be applied to a frequency list generated by counts of individuals identified from a single data source. The zero-frequency class consists of individuals who are never observed in the data source. As a result, the frequency list is incomplete and is called 'truncated below one'. The total population size equals the number of individuals ever observed plus the number of individuals never seen: the hidden population. In this case the hidden population may be thought of as the individuals who were criminally active but were not arrested. Consequently, the estimation problem then becomes a problem of estimating the number of individuals in the zero-frequency class from the truncated series of individuals ever seen, by assuming that the frequency list follows a Poisson distribution which is truncated below one. The Poisson distribution is used to mathematically model random behaviour in space or time (Freund, 1992). For example, if data follow a Poisson distribution then we can estimate from the data the probability of one occurrence, two occurrences, three occurrences, and so on. In this study, we wish to know the probability of zero occurrences and hence derive the number of zero occurrences. That is to say we wish to use the truncated Poisson distribution model to estimate the probability of zero occurrences or probability and number of people never seen in the police database.

Wilson and Collins (1992) have reviewed 14 different truncated Poisson methods and their properties. They recommended the use of Chao's (1987, 1989) bias adjusted

estimator for relatively small data and the use of Darroch and Ratcliff's (1980) estimator otherwise. Smit, Toet, and van der Heijden (1997) estimated the prevalence of opiate use in Rotterdam with several truncated Poisson estimators. They proposed that Zelterman's (1988) and Chao's (1989) truncated Poisson estimators, and Smit et al. (1997) truncated Negative Binomial estimator be used.

The data used in this paper have a relatively small number of frequencies and small data sets, therefore Darroch and Ratcliff's estimator (1980) is excluded. Since the data have only two predictors—sex and age, Smit et al. (1997) truncated Negative Binomial regression estimator is also not suitable and therefore it is excluded. We implement Zelterman's (1988) and Chao's (1989) truncated Poisson estimators to complete the series of count data including the zero-frequency class.

Zelterman's (1988) and Chao's (1989) estimators assume the following: (a) the population is closed, (b) the population under study is homogeneous, and (c) the probability of individuals being observed and re-observed is constant during the study period.

Capture–recapture method

The capture–recapture method is an indirect method that generates a prevalence estimate based on the degree of overlap between two or more separate samples of the population under study. It allows the confidence intervals of the population estimate to be calculated. It was originally used in ecological studies to assess the size of animal populations and in human populations to ascertain completion of census data. In epidemiology, different registers of the disease have been used as 'capture' samples, with names or confidential identifiers used to tag unique individuals to assess the overlap between samples. Capture–recapture has been applied to epidemiological studies including heroin use, injecting drug use prevalence and estimates of the size of the population infected with the human immunodeficiency virus (Comiskey, 2001, 2002; Comiskey & Barry, 2001; Domingo-Salvany, Hartnoll, Maguire, Suelves, & Anto, 1995; Frischer, 1992; Hartnoll, Lewis, Mitcheson, & Bryer, 1985; Hay & McKeganey, 1996).

The assumptions of the model are the following: (a) there is no change to the population during the investigation; (b) there is no loss of tags, individuals can be matched from capture to recapture; (c) for each sample, each individual has the same chance of being included in the sample; (d) the samples are independent.

The International Working Group for Disease Monitoring and Forecasting (1995a, 1995b) provides a comprehensive discussion of the method with more than two sample data sets and when all of these underlying assumptions are not met.

When working with three data sources the number of possible models that can be fitted to the data is eight. When changing from three data sources to four, the number of

possible models available changes from 8 to 113 excluding the four-way interaction model. To choose which of these models best fits the data, Hook and Regal (1997) outline several options. To use a non-Bayesian method would involve choosing the method associated with a single model judged optimal. This is usually evaluated by a procedure related to the likelihood ratio statistic G^2 . For any model i and the observed data set we have,

$$G^2 = -2 \sum \text{Obs}_j \log \left(\frac{\text{Obs}_j}{\text{Exp}_{ji}} \right)$$

where Obs_j is the observed number in cell j and Exp_{ji} denotes the expected number in cell j by model i . The lower the value of G^2 , the better the fit of the model. For selecting a model among contenders, Hook and Regal (1997) suggest choosing one with the lowest Akaike Information Criterion (AIC) or else the Bayesian Information Criterion (BIC). This latter criterion has the advantage of picking simpler models than the AIC. AIC is however more widely used than BIC although some theoretical considerations suggest that BIC is preferable. For log-linear models, the values of AIC and BIC for any model i are given by $\text{AIC} = G^2 - 2(\text{d.f.})$ and $\text{BIC} = G^2 - \log(N_{\text{Obs}}/2\pi)(\text{d.f.})$. Generally speaking as the complexity of the model increases the stability of the estimate decreases and the associated confidence interval widens.

We implement the four-sample capture–recapture method with two data sources each collected over two consecutive time periods. Using the log-linear modelling approach, as discussed above, we can fit the models and using the scaled deviance G^2 and the related P values a suitable model can be chosen. This provides a prevalence estimate for a double period. For example, if data are collected on two opiate prevalence indicators over a 2-year period then it is possible to derive an estimate of the hidden opiate use for this 2-year period.

An approach similar to this was introduced by Wickens (1993), who implemented the capture–recapture method with one data source of cocaine users over two and half years (five periods of 6 months interval) and thus derived a two and half year prevalence estimate.

Data

Western Australia is the largest state within Australia, which consists of two territories and five states. It accounts for approximately one-third of the total Australian landmass and is 2532422.2 km² in size. However, vast tracts of land are uninhabited. The majority of the population is concentrated in the urban district of the capital city, Perth, and the region immediately south west of the city. Eighty-three percent of the state's total population lives within this urban region. There are no other major urban areas or cities within the state. In fact, Perth is known to be the most isolated city in the world as the nearest city, Adelaide, is approximately 4 h flight away.

The Western Australia Police Service (WAPS) provided the data for the study. The data set described persons processed (arrested, summonsed, cautioned) for heroin-related offences during the 1996/1997 and 1997/1998 financial years in Western Australia (WA). The financial year in WA runs from the 1st July to 30th June. The information was extracted from the Offence Information System of the WA Police Service. Six identifiers including drug type, date of birth, gender, initial of surname, full forename, and count of offences were provided. The study used date of birth, gender, and initials of surname and forename as identification variables to build the frequency list. The distinct number of heroin users arrested by the Police Service increased from 299 during the 1996/1997 financial year to 431 (44% growth) during 1997/1998 financial year. The types of heroin identified from the data were Heroin Powder, Heroin Liquid, Heroin Rock, Heroin Homebake Powder, Heroin Gel, Heroin Synthetic, Heroin Crystal, and Heroin Unspecified. The most frequently recorded drug used during the study periods was Heroin Powder. The gender distribution consisted of 28 and 30% females and 72 and 70% males during the study periods, 1996/1997 and 1997/1998, respectively. The most frequent group was 20–24-year-old males. The youngest person in the database was a 7-year-old boy and the oldest person was a 55-year-old male. Since there was only one person younger than 14 years and nobody older than 55 years, the analyses focus on the age group between 14 and 55 years old.

The Health Department of Western Australia has compiled the Hospital Morbidity Database System (HMDS) from 1970 to the current time. The HMDS includes all inpatient episodes for admitted patients to public, private and free-standing day hospitals in Western Australia, including the detoxification program run by the alcohol and Drug Authority, but excluding psychiatric hospitals (Health Information Centre, 1999). Only episodes of treatment requiring an inpatient stay are captured on the HMDS, therefore other drug-caused conditions which received outpatient treatment are not included.

The data extracted from the HMDS for this project contained all opiate-related episodes from 1 July 1996 to 30 June 1998. These patients were identified from the International Code of Diseases (ICD) nine classification codes 304.0* and 304.7*, opioid dependence and combinations of opioid-type drug dependence, code 305.5, opioid abuse and 965.00-02, opiate poisoning (asterisk, * includes all ICD subcodes under the code, for example, 304.0* includes 304.00–304.09). The HMDS database consists of event ID, postcode, gender, date of birth, initials of surname and forename, principal diagnosis, co-diagnoses, extra diagnoses, external causes, principal procedure, and extra procedures. Our data were extracted on the basis of the presence or absence of the appropriate ICD code in any of the diagnoses, i.e. an individual may have been admitted for a chest infection say, as a primary diagnosis but opiate dependence was noted during the patient's medical history.

From our postcode data, we found that approximately 87% of all cases reported in 1996/1997 lived within the urban metropolitan region of Perth and immediately south west of Perth. This percentage rose to 89% in the 1997/1998 data. A similar pattern was observed by Unwin, Codde, and Swensen (1997) who estimated the geographical crude hospital admission rates for opiates in Western Australia and they found that the East Metro Health Zone (203.6 per 100,000) had significantly higher than the state crude rate (137.7 per 100,000). It is interesting to note the cultural difference between WA and parts of Europe. It is common for individuals and families to move house many times within WA. This was evidenced from the multiple postcodes of individuals available in the data. We found that many individuals resided in more than one residence during the study period. For this reason, postcode was not used as a part of the unique identifiers.

To eliminate duplicates and to find the overlaps within the data we choose four identifiers. These were, date of birth, gender, and the initials of the surname and forename of patients. The age range for this part of the study was also 14–55 years old. The number of distinct patients during the 1996/1997 and 1997/1998 financial years were 1176 and 1389, respectively. This represented an 18% increase from one financial year to the next. The ratios of males to females were 46.2–53.8% and 37.9–62.1% during the first and second periods.

Results

Truncated Poisson method

The numbers of distinct persons and contacts in the police data during the 1996/1997 financial year were 299 and 343, respectively. Of these 299 individuals, 259 (184 males, 75 females) appeared once, 36 (26 males, 10 females) appeared twice and 4 (1 male, 3 females) appeared three times. For the 1997/1998 financial year, 431 distinct persons and 513 contacts were observed in the data. Of these 431 individuals, 362 (259 males, 103 females) appeared once, 57 (44 males, 13 females) appeared twice, 11 (6 males, 5 females) appeared three times and one male appeared four times.

The estimates using Zelterman's (1988) and Chao's (1989) estimators of the total number of opiate users who are involved in opiate-related criminal activity during the 1996/1997 and 1997/1998 financial years in WA are

presented in Table 1 with 95% confidence intervals. The prevalence and ratios of known to estimated hidden number of the study population are also presented.

According to the Zelterman's estimator with unstratified data, there is a 16% increase from 1241 to 1429 individuals in the point estimate of the total number from year 1 to year 2. Chao's estimator predicts a 28% increase from 1231 to 1581 from year 1 to year 2. However, with Zelterman's estimator the percentage of the hidden population decreased from 75% of the total in year 1 to 70% of the total in year 2. Similarly, Chao's estimator predicts that the hidden population decreases from 76% of the total in year 1 to 73% of the total in year 2. The prevalence results in Table 1 are based on annual age population in Western Australia with the total population of 1,090,660 and 1,112,829 aged between 14 and 55 during the first and second financial years (Australian Bureau of Statistics, 1998). Truncated Poisson estimates were also computed for the gender-stratified data and these are reported in Table 2. In addition, truncated Poisson estimates were computed for the age stratified data and again overall prevalence results were very similar to those presented in Tables 1 and 2 and are not reported here.

Capture–recapture method

Using the 2 years of hospital data and 2 years of police data, we applied the four-sample capture–recapture method to estimate the prevalence of opiate use in Western Australia over the 2-year financial period. The variables that were common to all four of the data sources were initials of surname and forename, date of birth and sex. In order to identify overlaps between the four data sources these variables were used and the results provide us with the first multi-source enumeration of opiate use in Western Australia for the 2-year period. We found that 2826 unique individuals were noted as opiate users from the police and hospital records. Table 3 illustrates these overlaps between these four data sets.

Capture–recapture models incorporating dependencies between the four samples were fitted using GLIM and the results of this analysis are provided in Table 4. Estimates of prevalence are based on the 2-year period ranging from the 1st July 1996 to the 30th June 1998. Included in the table are unstratified and stratified estimates, the best fitting model for that stratification as decided by the scaled deviance or G^2 value, the P value for this deviance and 95% confidence intervals for the estimate of the hidden population.

Table 1
Truncated Poisson estimates of the number of opiate users involved in opiate-related crimes in Western Australia in 1996/1997 and 1997/1998

Unstratified	Year 1		Year 2	
	Zelterman's estimator	Chao's estimator	Zelterman's estimator	Chao's estimator
Total population (95% CI)	1220 (1003–1504)	1231 (1011–1518)	1415 (1206–1712)	1581 (1365–1882)
Prevalence per 1000	1.12	1.13	1.27	1.42
Ratio of hidden:known	3.08:1 (921:299)	3.12:1 (932:299)	2.56:1 (1105:431)	2.95:1 (1271:431)
P value of model fit	0.73	0.72	0.30	0.14

Table 2

Truncated Poisson estimates of the number of opiate users involved in opiate-related crimes in Western Australia in 1996/1997 and 1997/1998 stratified by gender

Stratified by gender	Year 1		Year 2	
	Zelterman's estimator	Chao's estimator	Zelterman's estimator	Chao's estimator
Female (95% CI)	302 (216–504)	369 (246–534)	401 (302–600)	529 (380–765)
Male (95% CI)	939 (712–1378)	862 (659–1078)	1028 (852–1295)	1072 (889–1314)
Total population (95% CI)	1241 (928–1882)	1231 (905–1612)	1429 (1154–1895)	1601 (1269–2079)
Prevalence per 1000	1.14	1.13	1.28	1.44
Ratio of male:female	939:302	862:369	1028:401	1072:529
P value of model fit	0.73	0.74	0.32	0.15

Table 3

Numbers and overlaps between the four data sets

Police year 1	Police year 2	Hospital year 1			
		Present		Absent	
		Hospital year 2			
		Present	Absent	Present	Absent
Present	Present	0	4	5	14
	Absent	8	15	13	240
Absent	Present	21	22	49	316
	Absent	280	826	1013	–

Discussion

The use of psychoactive drugs is a cause of considerable morbidity and mortality within the Australian population (Collins & Lapsley, 1992; English et al., 1995). While most of this burden arises from the use of the legally available drugs, alcohol and tobacco, there is considerable and increasing community concern about the use of opiate drugs especially heroin (Australian Institute of Health and Welfare, 1999). In 1998, 2.2% of Australians aged over 14 reported that they had ever used heroin, an increase from 1.4% in 1995. Seven per 1000 of the Australian population aged over 14 reported that they had recently used heroin, an increase from four per 1000 in 1995. It is acknowledged that these data, based as they are on self-report of what is an illegal activity, may have limited validity and should be interpreted cautiously. In addition the definition of heroin users in these surveys could be considered a very broad one and not necessarily problematic users. Clearly, alternative methods of estimating prevalence of drug use are required to clarify both the definition and the size of the hidden population.

Prior to any interpretation of the results, it is imperative that we clearly define the study population and then discuss the validity of the assumptions and any possible violations. Only then can the nature of the prevalence results be appreciated and their possible implications discussed in context. However, we must be cognisant of the fact that this is the first attempt at deriving estimates for this particular region and limitations and inadequacies in the level of detail available in the data are to be expected. Indeed, a secondary

objective of this research is to highlight the nature of the data required to other researchers considering such studies.

Prevalence estimates derived using the truncated Poisson method are based on data arising from police arrests. The study population may then be described as those individuals involved in criminal activity related to opiates. Given this, how do we then define the zero-frequency class in the police arrest database? A review of arrest and reported crime statistics from previous studies indicates that only a small proportion of offences result in arrests. Nation-wide statistics for 1978 (Webster, 1979) reported that 20% of index offences resulted in arrests, indicating that there were approximately four crimes which did not result in arrest for every one arrest. Indeed this figure itself provides one with a historical multiplier for the police data. Unfortunately no such multiplier is available for the West Australian data. It is also argued (Blumstein & Cohen, 1978) that there are far more crimes per arrest than indicated in official statistics. Blumstein and Cohen (1978) pointed out that official statistics neither correct for unreported offences nor count properly the number of individual offenders involved in crimes that are reported. The purpose of their analysis is to indicate that there are far fewer crime events that result in arrests than is indicated by official statistics (Greene & Stollmark, 1981). Therefore, it is likely that many individuals who commit crimes are not arrested. We regard these individuals as the zero-frequency class of the police data. Our prevalence estimates presented in Tables 1 and 2 may then be defined as an estimate of the true number of people involved in opiate-related crimes both detected and undetected. It is interesting to note that both the Zelterman and Chao estimates and model fits are very similar for year 1, with approximately 1200 individuals estimated as involved in opiate-related crimes. In year 2 both estimators predict an increase in the numbers. The Zeltermann estimator provides an estimate of 1415 in year 2 while the Chao estimator provides that there are 1581 individuals in year 1. What is interesting from a policing point of view is the predicted decline in the ratio of hidden to known cases. Both estimators indicate that although overall numbers are increasing, the crime detection rate is improving as the proportion of hidden cases has decreased from year 1 to year 2. While we cannot say that these results are completely reliable due to the nature of the assumptions

Table 4
Illustrating the best fitting model for each of the stratifications and the derived estimates

	Best fitting model	d.f.	Scaled deviance G^2	P value	Estimated hidden population (with 95% CI)	Observed known population	Ratio of hidden:known	Prevalence per 1000
Unstratified	1 + H1 + H2 + P1 + P2 + H1H2 + H1P1 + H1P2 + H2P2 + P1P2	5	8.523	0.13	17233 (8960–35055)	2826	6.10:1	18.03
14–24 years	1 + H1 + H2 + P1 + P2 + H1H2 + H1P1 + H1P2 + H2P2	6	2.844	0.83	5628 (2346–15462)	1155	4.90:1	23.03
25–34 years	1 + H1 + H2 + P1 + P2 + H1H2 + H2P2 + P1P2	7	5.189	0.64	5369 (2322–15237)	991	5.42:1	14.79
35–55 years	1 + H1 + H2 + P1 + P2 + H1H2 + H2P2 + P1P2	7	19.06	0.01	4341 (1372–21749)	680	6.38:1	12.93
Total overall ages					15338 (6040–52448)	2826	5.43:1	16.32
All females	1 + H1 + H2 + P1 + P2 + H1H2 + H1P2 + H2P2 + P1P2	5	3.686	0.60	6008 (2114–19466)	1154	5.21:1	13.11
All males	1 + H1 + H2 + P1 + P2 + H1H2 + H2P2 + P1P2	7	5.758	0.57	10097 (5269–21904)	1672	6.04:1	20.78
Total over both sexes					16105 (7383–41370)	2826	5.70:1	17.01

H1 is hospital data for year 1, H2 is hospital data for year 2, P1 is police data for year 1 and P2 is police data for year 2.

discussed below, it is still interesting from a police point of view and is worth exploring in future studies.

Prevalence estimates derived using police arrests, inpatient hospital records and the capture–recapture methodology must also be adequately defined. Within this part of the study we have employed two very different data sources. Although both refer to opiate users, it is important that we highlight the fact that the police data also include the possibility of counting opiate dealers who are not themselves users. Unfortunately, data supplied to this study did not include detailed information on the nature of the opiate-related crime. However, we do know that drug offences accounted for 4.9% of all offences reported and recorded in the Offences Information System of the WAPS, in 1997. Possession or use offences were the most common drug offences reported (53.0%). Manufacturing or growing drugs (11.2%), dealing and trafficking (4.8%), and other miscellaneous drug offences including possession of instruments for consuming drugs (31.0%) made up the remaining offences (Ferrante, Loh, & Fernandez, 1998). In addition, recent statistics published (Valuri, Indermaur, & Ferrante, 2002) show that in 2000, 15,581 drug offences were recorded by the WAPS. These offences consisted of possession/use of drugs (89.2%), manufacture/grow drugs (6.1%) and deal/trafficking drugs (4.6%). It is reasonable to assume that the overwhelming majority of opiate drug offences during our study period are also for possession and use of opiates.

The definition of opiate user in these circumstances is very broad and inclusive and reflects the nature of the data. These individuals may be defined as general opiate users including individuals involved in opiate crimes as opposed to more problematic users. This distinction between a more general opiate user and a problematic user as also used within the EMCDDA report on national prevalence estimation (EMCDDA, 2000b). Problematic opiate users may be thought of as being at an increased risk of overdose or death. It should also be noted that the mixing of medical and police data sources is very common in the application of the capture–recapture methodology as can be seen from German, Dutch, Scottish and Irish estimates (EMCDDA, 1997).

Both methodologies assume that the population is closed and there is no change to the size of the population during the period of the investigation. This first assumption is clearly violated. It is well documented that opiate users progress through a so-called drug using career. In Ireland, it was suggested (Comiskey, Ruskin, & Wood, 1992) that drug users on average use drugs for a period of 10 years. It is however known that drug users start and stop using drugs several times during this career. For this reason in Ireland, Comiskey (2002) estimated the hidden prevalence of opiate use in young people at city level collected over a time period of 1 year. This approach is endorsed elsewhere (EMCDDA, 1997). It is also suggested that if the time frame of sampling is not excessively long, it can be confidently assumed that the first assumption is not influential. This paper focuses on two 1-year periods, the 1996/1997 and 1997/1998 financial

years. For the reasons stated above we are reasonably confident that the first assumption holds at least for the truncated Poisson results. However, as the capture–recapture method is applied to the 2-year period we are less confident in this case and this must be remembered when interpreting results.

In addition the capture–recapture method assumes that there is no loss of tags, that is, individuals can be matched from capture to recapture. We have no way of assessing this assumption other than by looking closer at the nature of the data sources used. Both police arrests and hospital in-patient records are reliable sources of good quality demographic data. Although one can give a false name to police within the community or on the street, once formally arrested some form of official identity is required. Similarly with hospital in-patient records, formal identification would be of the utmost importance for treatment records. Again, and for these reasons, we are reasonably confident that this assumption holds.

Both the truncated Poisson and the capture–recapture method assume that the study population is homogeneous and that for each of the samples, each individual has the same chance of being included in the sample. Clearly given police data this is not the case. Chao's and Zelterman's truncated Poisson estimators are however known to be fairly robust in the sense that both will underestimate the true population size in the presence of heterogeneity (Chao, 1989; Wilson & Collins, 1992). So, if heterogeneity is suspected, then the estimates from both estimators are assumed to be the lower bounds of the true population size. Alternatively, one may try to remove some of the heterogeneity by stratifying the data upon the basis of gender or age and then carry out subgroup analysis on groups that are more homogeneous. Finally one pools these estimates into a single estimate as we have done here.

The constant capture probability assumption also causes problems when the probability of subsequent arrests may increase or decrease with the number of prior appearances in the database. Decreases occur when the disciplinary nature of arrest results in decreases in the level of an individual's criminal activity (the theory of special deterrence). Increases in the recapture probability may occur once the police gain more knowledge of individuals who have been arrested and those individuals become more vulnerable to police arrest. It has been asserted (Greene & Stollmark, 1981) that the number of arrests is strongly related to the number of arrests in the previous year, this suggestion clearly opposes the theory of special deterrence. Again, one way of dealing with this assumption is to keep the study period short as is practically possible.

This study represents the first attempt at estimating the prevalence of hidden opiate use in Western Australia and as such is a significant step forward in spite of the possible violations of the assumptions. Comparing results we see that a total of 2826 unique individuals were identified from the multi-source enumeration over the 2-year period. This gives a point estimate of 2.54 per 1000 of population and may

be viewed as the absolute minimum prevalence rate for the 2-year period 1st July 1996 to 30th of June 1998. What is also interesting to note was the degree of overlap between the two different data sources. From Table 3 we can see that 2119 individuals were present in the hospital database only, 570 were present in the police data base only, leaving only 137 present in both. Clearly the two services are encountering different subgroups of the overall opiate using population and have major implications for these services. This is the first time that the degree of overlap for these two sources of opiate data has been brought together to produce a true picture of the number of unique opiate users attending these services. Estimates derived from the truncated Poisson method both stratified and unstratified appear very low, but we must bear in mind the nature of the data. This data involves only those who were arrested, summoned or cautioned for opiate-related offences. It does not capture those who may have been arrested for other crimes, theft say, and who subsequently had opiates noted in their files. If this were the case the numbers would be considerably inflated as was seen by Comiskey (2000). In addition prevalence estimates are for a single year. We found that approximately 1000–1500 individuals were involved in opiate crimes in year 1 and this rose to approximately 1200–1900 in year 2. Estimated prevalence rates for the 2-year period ranged from 1.12/1000 to 1.42/1000. Turning to the 2-year point prevalence estimates derived from the capture recapture methodology we see that estimates ranged from 16.32/1000 to 18.03/1000. It is interesting to endeavour to compare this with recent survey data (Australian Institute of Health and Welfare, 1999), where 7/1000 of those survey reported recent heroin use. Clearly self-reports produce underestimates, and our estimate refers to a 2-year period, however we are encouraged by the similarities of scale of the two estimates given especially that this study is the first of its kind in the region.

Finally, it is the opinion of the authors that no one study will provide a comprehensive prevalence estimate and that a range of approaches, definitions and methods need to be applied before one can provide policy planners with a clear picture of the profile, nature and extent of opiate use. In this first study of prevalence estimation in Western Australia it is hoped that debate over the question on the nature and extent opiate use and related crime has been initiated, inadequacies in the data provided highlighted and the problems with adherence to model assumptions illustrated. In addition, it is hoped that the discussion of the methods and the illustration of the applications will provide assistance and insight for other similar cities world wide initiating their first ever opiate prevalence estimation.

Acknowledgements

The authors are grateful to the Western Australia Police Service and Health Department of Western Australia for providing the data for the study. The authors would like

to thank to Dr. James Cross and Mr. David Ryder, Edith Cowan University, Western Australia, for their insightful suggestions and discussions, and Mr. Filip Smit, Trimbos Institute, The Netherlands, for his insights and suggestions on the truncated Poisson methods. This work was supported by the Data Modelling and Analysis Research Group (DMARG), Edith Cowan University, Western Australia, and the National University of Ireland, Maynooth, Ireland. Special acknowledgement goes to the reviewers who provided us valuable insights of our paper.

References

- Australian Bureau of Statistics. (1998). *Population, Western Australia*. Australian Bureau of Statistics, Western Australian Office.
- Australian Institute of Health and Welfare. (1999). *1998 national drug strategy household survey*. Canberra: Author.
- Blumstein, A., & Cohen, J. (1978). *Estimation of individual crime rates from arrest records*. Urban Systems Institute, Carnegie-Mellon University, Pittsburgh, Pennsylvania.
- Chao, A. (1987). Estimating the population size for capture–recapture data with unequal catchability. *Biometrics*, 43, 783–791.
- Chao, A. (1989). Estimating population size for sparse data in capture–recapture experiments. *Biometrics*, 45, 427–438.
- Collins, D. J., & Lapsley, H. M. (1992). Drug abuse economics: Cost estimates and policy implications. *Drug and Alcohol Review*, 11, 379–387.
- Comiskey, C. M. (2001). Methods for estimating prevalence of opiate use as an aid to policy and planning. *Substance use and misuse: Emerging issues for the 21st century*.
- Comiskey, C. M. (2002). Young people, drug use and early school leaving—estimating the prevalence, assessing the impact and assisting policy and planning. *Drugs: Education, Prevention and Policy*.
- Comiskey, C. M., & Barry, J. (2001). The prevalence and health implications of opiate use in Dublin. *European Journal of Public Health*, 11(2), 198–201.
- Comiskey, C. M., Ruskin, H. J., Wood, A. D. (1992). *Modelling for the transmission dynamics of HIV and AIDS in drug users in Ireland*. Report for the Irish Society of Actuaries and the Irish Insurance Federation, Dublin City University, Ireland.
- Darroch, J. N., & Ratcliff, D. (1980). A note on capture–recapture estimation. *Biometrics*, 36, 149–153.
- Domingo-Salvany, A., Hartnoll, R., Maguire, A., Suelves, J., & Anto, J. M. (1995). Use of capture–recapture to estimate the prevalence of opiate addiction in Barcelona Spain 1989. *American Journal of Epidemiology*, 141, 567–574.
- EMCDDA. (1997). (European Monitoring Centre for Drugs and Drug Addiction). *Methodological pilot study of local prevalence estimates*. Report on project CT.96.EP.07. Lisbon: EMCDDA.
- EMCDDA. (2000a). (European Monitoring Centre for Drugs and Drug Addiction). *Methodological guidelines to estimate the prevalence of problem drug use on the local level*. Report on project CT.97.EP.05. Lisbon: EMCDDA.
- EMCDDA. (2000b). (European Monitoring Centre for Drugs and Drug Addiction). *National prevalence estimates: Improvement of comparability of national estimates of addiction prevalence final report*. Report on project CT.96.EP.06. Lisbon: EMCDDA.
- English, D., Holman, D., Milne, E., Winter, M., Hulse, G., Codde, G., 1995. The quantification of drug caused morbidity and mortality in Australia. Canberra: Australian Government Publishing Service.
- Ferrante, A., Loh, N., & Fernandez, J. (1998). *Crime and justice statistics for Western Australia for January–December 1997*. Crime Research Centre, The University of Western Australia.

- Frischer, M. (1992). Estimated prevalence of injecting drug use in Glasgow. *British Journal of addiction*, 87, 235–243.
- Greene, M. A., & Stollmark, S. (1981). Estimating the number of criminals. In J. A. Fox (Ed.), *Models in quantitative criminology* (pp. 1–24). New York: Academic Press.
- Hartnoll, R., Lewis, R., Mitcheson, M., & Bryer, S. (1985). Estimating the prevalence of opiate dependence. *The Lancet*, 26, 203–205.
- Hay, G., & McKeganey, N. (1996). Estimating the prevalence of drug misuse in Dundee Scotland: An application of capture–recapture methods. *Journal of Epidemiology and Public Health*, 50, 469–472.
- Health Information Centre. (1999). *Hospital morbidity data system reference manual*. Health Statistics Branch, Health Department of Western Australia.
- Hook, E. B., & Regal, R. R. (1997). Validity of methods for model selection weighting for the model uncertainty and small sample adjustment in capture–recapture estimation. *American Journal of Epidemiology*, 145, 1138–1144.
- 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.
- Smit, F., Toet, J., & van der Heijden, P. (1997). Estimating the number of opiate users in Rotterdam using statistical models for incomplete count data. *EMCDDA project (final report): Methodological pilot study of local level prevalence estimates*. Lisbon: European Monitoring Centre for Drugs and Drug Addiction.
- UNDCP. (1997). *United Nations Drug Control Programs World Drug Report*, Oxford University Press.
- Unwin, E., Codde, J., Swensen, G. (1997). *Hospitalisation due to drugs other than tobacco or alcohol in Western Australia 1991–1995*. Western Australian Drug Abuse Strategy Office. Health Information Centre. Health Department of Western Australia.
- Valuri, G. M., Indermaur, D., & Ferrante, A. M. (2002). *The criminal careers of drug offenders in WA: A study of the recidivism and criminal history of those arrested for a drug offence in WA between 1989 and 1999*. Crime Research Centre, The University of Western Australia.
- Webster, W. H. (1979). *Crime in the United States, 1978*. Washington, DC: US Department of Justice, Federal Bureau of Investigation.
- WHO. (1998). *World Health Organisation programme on substance abuse, linking dynamic modelling to activities*. Presentation made to the EMCDDA meeting on dynamic modelling of drug use problems, Lisbon.
- Wickens, T. D. (1993). Quantitative methods for estimating the size of a drug using population. *Journal of Drug Issues*, 23, 185–216.
- Wilson, R. M., & Collins, M. F. (1992). Capture–recapture estimation with samples of size one using frequency data. *Biometrika*, 79, 185–216.
- Zelterman, D. (1988). Robust estimation in truncated discrete distributions with application to capture–recapture experiments. *Journal of Statistical Planning and Inference*, 18, 225–237.