E.M.C.D.D.A.

EMCDDA SCIENTIFIC REPORT

European Network to Develop Policy Relevant Models and Socio-Economic Analyses of Drug Use, Consequences and Interventions

Final report: Part 3 – Prevalence of drug use at the local level

EMCDDA / 2002

EMCDDA SCIENTIFIC REPORT

European Network to Develop Policy Relevant Models and Socio-Economic Analyses of Drug Use, Consequences and Interventions

Final Report: Part 3 – Prevalence of drug use at the local level

EMCDDA / P1/ TSER Network/ 2002

This report (Part 3 – Local Prevalence) was prepared by:

Gordon Hay (Work Group coordinator)

Work Group members National Prevalence Estimation:
Gordon Hay, Marcel Buster, Yoon Choi, Catherine Comiskey, Antonia Domingo-Salvany, Päivi Partanen, Filip Smit, Alfred Uhl,

---

Full Network Details:

Project Partners (project and work group coordinators):
Lucas Wiessing, EMCDDA (project coordinator), Gordon Hay, Univ. Glasgow, Carla Rossi, Lucilla Ravà, Univ. Rome 'Tor Vergata', Martin Frischer, Heath Heatlie, Univ. Keele, Hans Jager, Wien Limburg, RIVM, Christine Godfrey, Univ. York, Chloé Carpentier, Monika Blum, Kajsa Mickelsson, Richard Hartnoll, EMCDDA

The important input of all network participants and invited experts is fully acknowledged. For a list of network participants per working group and email contacts see Final Report Part 1, Annex A.

Other Network Participants and Invited Experts:
Erik van Ameijden, Fernando Antoñanzas, Rita Augustin, Ana Maria Bargagli, Massimiliano Bultrini, Maria Fe Caces, Maria Grazia Calvani, John Carnavale, Gloria Crispino O'Connell, Ken Field, Gerald Foster, Maria Gannon, David Goldberg, Peter Hanisch, Toon van der Heijden, Simon Heisterkamp, Matthew Hickman, Neil Hunt, Claude Jeanrenaud, Pierre Kopp, Petra Kümmler, Mirjam Kretzschmar, Marita van de Laar, Nacer Lalam, Fabio Mariani, Linda Nicholls, Alojz Nociar, Deborah Olzewski, Alessandra Nardi, Laetitia Paoli, Päivi Partanen, Paulo Penna, Harold Pollack, Maarten Postma, Thierry Poynard, Jorge Ribeiro, Francis Sartor, Janusz Sieroslawski, Ronald Simeone, Juan Tecco, Alberto Teixeira, Jaap Toet, Gernot Tragler, Giovanni Trovato, Julián Vicente, Katalin Veress, Denise Walckiers, Robert Welte, Ardine de Wit, John Wong, Tomas Zabransky, Terry Zobeck, Brigitta Zuiderma-van Gerwen.

Project funded by the European Commission, DG Research, Targeted Socio-Economic Resarch (TSER). Project n°: ERB 4141 PL 980030, Contract n°: SOE2-CT98-3075
Starting date: 1st December 1998
Duration: 36 months
Date of issue of this report: 31st January 2002

European Monitoring Centre for Drugs and Drug Addiction
Rua da Cruz de Santa Apolónia 23-25
PT - 1149-045 Lisboa
Portugal
Tel: + 351 21 811 30 00
Fax: + 351 21 813 17 11
e-mail: info@emcdda.org
http://www.emcdda.org

Contents

Final Report Part 3: Local Prevalence

For more detail see the full final reports of the six working groups:

Final Report Part 1	General Overview
Final Report Part 2	Work group 1a – National Level Prevalence Estimation
Final Report Part 4	Work group 2a – Modelling Time trends and Incidence
Final Report Part 5	Work group 2b – Modelling Geographic Spread with Geographic Information Systems (GIS)
Final Report Part 6	Work group 3a – Modelling Costs and Cost-effectiveness of Interventions
Final Report Part 7	Work group 3b – Modelling Drug Markets and Policy options

1 Executive Summary

The work group was formed to consolidate EMCDDA-funded and other methodological research into estimating the prevalence of drug use at the local level and also to provide advice and scientific support to researchers wishing to apply the methods in their area. New prevalence estimates have been obtained in areas where research has not previously been carried out and updated estimates have been obtained in existing areas. Advice and scientific support was given to researchers wishing to use the capture-recapture method in Copenhagen, Luxembourg and Matosinhos (Portugal). A new prevalence estimation project was initiated in Scotland (UK) where the capture-recapture method was used to provide an estimate for each of the 32 local government areas.

The results of the prevalence studies have been useful for policy makers and those planning the provision of services. For example, the results of the Scottish research have influence the government's Drug Action Plan including the allocation of funding to local areas of the country. Future prevalence studies will benefit from the experience gained over the course of the project, scientific papers and other reports and from the development of computer routines to assist in analyses.

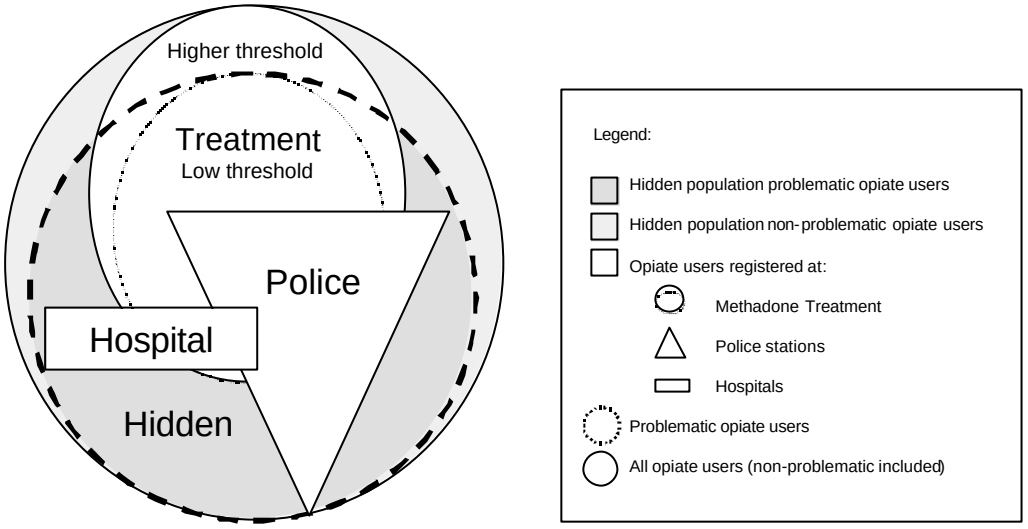
Members of the group are experienced in applying the methods in their geographical areas and have therefore addressed many of the relevant issues within their research. The work group was therefore a vehicle to exchange ideas about prevalence estimation, particularly the methodological aspects of this type of research, and was a resource available for those interested in undertaking prevalence research in other areas within Europe. Links were made between this workgroup and the National Prevalence workgroup, particularly with respect to analyses in Austria, Finland and Ireland where researchers carrying out local prevalence research were also undertaking research at the national level.

Two main methods have been considered; the capture-recapture method and truncated Poisson. The capture-recapture method employs data on drug users identified from two or more sources, such as hospital admissions, methadone registers, police, prison or treatment agency data. Although prevalence estimates can be obtained using two samples, there are certain assumptions that may be violated leading to biased estimates therefore it is preferable to employ three or more samples. The method examines the overlap between sources and uses log-linear regression modelling techniques to provide an estimate of the hidden population, i.e. those not identified from the available sources. The truncated Poisson method employs data from one source and examines the number of times each individual appears within the source. As the number of times a person appears once, twice, three times etc can be assumed to follow a statistical distribution such as the Poisson distribution, the number of people who appear zero times can be estimated. 95% confidence intervals can be found using either method.

In Amsterdam, the Netherlands, the differing case definitions of contributing sources was examined, particularly with respect to the effect on the estimates. Figure 1 shows how the samples used in a typical capture-recapture analysis relate to the estimate populations. Similar issues were faced within research in Dublin, Ireland. Each source which contributes to a capture-recapture analysis will employ a different case definition of what constitutes the population of interest. This is particularly

relevant when considering a definition such as problem drug use. Those who are in contact when hospitals or who have had medical emergencies can perhaps be seen as more problematic than those who are only in contact with low-threshold agencies or have only been identified from police sources. If the capture-recapture analysis is restricted only to high-threshold sources such as hospitals, then the resultant estimates may refer only to a subset of more problematic drug users, rather then quantifying a broader spectrum of drug use in the city.

Figure 1: Samples and estimated populations from analyses in Amsterdam



Methodological advances were also made in Helsinki, Finland where the capture-recapture analysis was undertaken within a Bayesian framework. Thus prior information on the likely size of the drug using population in the city, derived from previous prevalence estimation research, was employed with data from 3 sources to provide new estimates. Using prior information can help in model selection and can make comparisons between estimates in different years more valid. Bayesian estimates can, however, be harder to interpret therefore the application of this method has currently been limited to Helsinki. This methodological development has been worthwhile, and it will be of use in other areas where sufficient data and prior information is available.

In Austria by the use of simulation methods to examine the effect that some of the assumptions inherent in the capture-recapture methodology, such as equal probability of detection in each source, would have on estimates. The use of the truncated Poisson method was compared and contrasted with the capture-recapture method in the Netherlands and in Scotland.

In summary, the workgroup has served two main functions; further contributing to the development of the methodologies and their application in estimating the prevalence of drug misuse and also providing scientific support to new prevalence estimation projects across Europe.

2 Scientific Overview

2.1 Summary

The main objective for the workgroup were to consolidate the knowledge and experience gained in previous EMCDDA-funded local prevalence estimation projects in order to improve access to practical information about undertaking prevalence estimation research and thus enable prevalence estimates to be obtained in further areas of Europe. A second objective was to specifically target this information to relevant institutions, such as the National Focal Points of the REITOX network in countries which did not have experience of local prevalence estimation. In practise, both of these objectives involved closer working between local prevalence researchers and those involved in the National Prevalence Estimation workgroup.

2.2 Overview

Two main methodologies were considered; the capture-recapture method and the truncated Poisson model. The capture-recapture method employs data from three or more sources to provide an estimate of the size of a drug using population whereas the truncated Poisson model derives a population size estimate from multiple occurrences within a single data source. Methodological advances were made with the capture-recapture method, i.e. the analysis of capture-recapture data (from Helsinki) within a Bayesian framework and the analysis of data from Amsterdam with particular regard for case definitions. The capture-recapture method was also applied in Copenhagen, Dublin, Luxembourg, Matosinhos (Portugal) and across Scotland. Progress was made with the truncated Poisson model by examining applications for estimating problematic heroin use and also drug-related traffic offences. By applying the truncated Poisson model to data from a needle exchange, attendance patterns by individual drug injectors were re-examined to evaluate the efficacy of this harm reduction measure.

The accessibility of both methods was improved by the development of computer methods to assist in the analyses. Descriptions of these computer methods, using the packages Microsoft Access, Microsoft Excel and GLIM, are being developed. The further dissemination of the findings of previous EMCDDA local prevalence research was continued with the production of scientific papers (see section 4).

2.2.1 Capture-recapture methods

The capture-recapture method has been at the centre of European Union comparative local prevalence research since the 1980s. Although there are basic principles underpinning the application of the method in each area it is applied, there are significant local differences between studies, reflecting the diversity of drug misuse across Europe.

Much of the work undertaken at the European level has sought to improve the comparability of the method. To that end, a pilot project which employed a set of common definitions was undertaken in 1997 and following on from that project, a series of methodological guidelines has been produced. Local prevalence estimation, particularly using the capture-recapture method, has been promoted via the dissemination of these guidelines and the operation of a low-threshold helpdesk. A

developing theme over this stage of the work was, however, the need to return to specific local issues rather than to further aim to improve comparability. For example, issues that arise in a particular city or country can often be unimportant in other research. In some countries, the availability of relevant data is an issue, therefore methods which use fewer sources or use a single source over successive time periods can be useful. In other countries, the case definitions of the available sources has been shown to affect the resultant estimates.

Bayesian Methods

The capture-recapture method, in essence, aims to summarise the data on the size of specific sub-populations, such as drug users in treatment or drug users that have been arrested, along with the overlap between sources into a statistical model which can then be used to estimate the size of the total population of drug users. An integral part of the method is therefore the selection of the best statistical model to fit to the available data, bearing in mind that different statistical models can proffer highly differing estimates. In some instances the choice of model is straightforward; statistical reasons for selecting a model, such as significance testing or information criteria like the AIC, often confirm a researcher's ideas about the relationship between data sources. In other cases, what may statistically appear to be a good fit to the data may provide an unrealistic estimate, giving rise to a conflict between the researchers prior knowledge of the subject and the information derived from the available sub-populations.

Bayesian methods offer a way of incorporating a researcher's prior beliefs, such as a previous estimate or knowledge of what constitutes an unrealistic estimate, into the capture-recapture analysis. The use of Bayesian methods within the reanalysis of data from Helsinki, Finland is presented as Appendix 1 of the report of this work group.

Case definition

Within previous research, it has been noticed that the case definitions of the contributing sources were often reflected in the assumed case definition of the resultant estimate, with in some cases, difficulties arising from fitting a valid statistical model to data from sources with markedly different case definitions. Although pertinent across the European Union, this issue was highlighted in the analysis of data from Amsterdam, The Netherlands (see Appendix 2). This issue has also arisen in other research, such as Dublin, Ireland and Dundee, UK. Although no immediate solution for this project was reached, it was agreed that the effect of case definitions of the contributing sources should be discussed when considering the results of prevalence estimation.

Simulation Methods

Many of the different scenarios that occur within a capture-recapture analysis can be explored using simulation methods; for example, dependency between sources and the inherent heterogeneity of the population. In order to do this, a hypothetical population can be created and this population can be split into different cohorts. Members of each cohort can then be assigned a probability of being identified from each source, and this probability can be adjusted to include dependencies between the sources.

As part of research which estimates the prevalence of opiate use in Austria (Uhl and Seidler, 2000), different scenarios which may be relevant within prevalence estimation were simulated. This was done by creating three samples; S_A, S_B and S_C. Each of the samples could be split up into five cohorts C_1 through C_5, each having a certain probability p of being in each of the three samples.

There are two kinds of dependency that can be introduced. One dependency results if the 'catchability' for different cohorts is different in each sample; the other kind happens if the probability of capture in a certain sample is dependent on having been captured in another. The first kind of dependency can be introduced by choosing the probabilities for the different cohorts to reflect the different catchability pattern. The second kind of dependency is introduced by multiplying the probability p, by a factor f, in the samples S_B and S_C in the cases where the person has been captured in sample S_A already.

The size of the cohorts are known, as is the total simulated population, and these simulated samples can be used within an analysis and the resultant estimate can be then compared with the known total population size. The issue of model selection is still pertinent in considering simulated population, and indeed it can be more acute as the artificial heterogeneity or dependencies often result in a difficulty in fitting non-saturated models to the data. Often the saturated model is the only model that fits simulated data. From the different scenarios that had been modelled, it could be shown that heterogeneity constituted a problem but dependency between the samples is not a major problem; indeed, it may help to reduce the bias caused by heterogeneity.

While simulation methods may not completely answer some of the more pertinent questions relating to the application of capture-recapture to drug misuse prevalence estimation, it should be seen as a useful tool in examining different scenarios.

2.2.2 Truncated Poisson model

The truncated Poisson model offers a simple way of obtaining a prevalence estimate by examining multiple occurrences within a single data source. Possible sources that could be used within a truncated Poisson study include episodes of substitute prescribing with methadone, visits to a needle exchange or arrests for drug-related crimes. To justify using a truncated Poisson model, it has to be assumed that the probability of reappearing within a data source does not depend on the probability that an individual has previously appeared. E.g., the probability that a person is arrested for drug offences over a specific time period does not depend on them previously being arrested in that time period. In some circumstances more than others, this assumption is questionable, however in most previous research in the European Union, the validity of fitting a truncated Poisson model to these types of data has not been comprehensively explored.

Within the project, the validity of fitting the truncated Poisson model to data on drug misuse has been explored. Australian data on drug-related traffic offences (see Appendix 3) were used to introduce this issue to the workgroup, and data available to the workgroup members have been re-evaluated in relation to this issue. This work is still in progress, as the diagnostic tests for examining the applicability of the methods

are drawn from research into the use of the Poisson (as opposed to the truncated Poisson) method. These diagnostic tests show how close the observed data fit a poisson model. The initial findings of this work has, however, been to cast doubts on the validity over the use of some types of data within a truncated Poisson study, in particular the observed data used to provide injector prevalence estimates in Aberdeen did not appear to be similar to those expected under a Poisson distribution. The diagnostic tests do not, however, take into account that the specific truncated Poisson models used by the workgroup have been adapted to take account of possible heterogeneity. The truncated Poisson model has also been used in estimating the prevalence of problem drug use in the Grand Duchy of Luxembourg (see Appendix 7).

The truncated Poisson model was used to estimate the prevalence of drug injecting in a UK city using data from a single needle exchange Hay and Smit (submitted). Typically, needle exchanges collated less information on their clients than drug treatment agencies therefore information on the type of drug used or the area of residence of clients was sparse. Therefore although the truncated Poisson model could be used to estimate the number of drug injectors who would potentially use that service, denominator data were not available to provide a prevalence rate, however the total number is important for evaluating the provision of clean injecting equipment in that area and examining the proportion of the area's injectors that are in contact with the service. Re-analysing the data within a truncated Poisson model indicated that a significant minority of clients of the needle exchange only used the service once or twice in a twelve month period. The implications of this finding in terms of potential for disease spread were examined.

2.2.3 New estimates

In Scotland, UK, Gordon Hay has estimated the prevalence of problem drug use in all 32 local government areas which will combine to provide a 'national' prevalence estimate for Scotland (see Appendix 4). Having this complete coverage of local estimates, or anchor points in the context of the multi-variate indicator model, will enable comparisons between the two approaches. In addition, the back-calculation models developed by the time trends / incidence working group will be considered within the Scottish project. Similar projects are being set up in parts of England where complementary approaches will be employed.

Members of the local prevalence working group have adapted the capture-recapture methodology to examine drug use in specific populations. Elaine Hand from the National University of Ireland, in conjunction with Catherine Comiskey, have applied capture-recapture methods using data from two sources (hospital inpatient and methadone data), but using successive years' data from each source to provide 4 distinct samples. The Irish group have also adapted the capture-recapture method to provide information on the number of schoolchildren using drugs in Dublin, Ireland and thus to examine the links between drug use and leaving school at an early age. In the United Kingdom, the prevalence of drug use in the Pakistani / Indian / Bangladeshi ethnic minority population is currently being examined by means of an adaptation of the capture-recapture method by Gordon Hay. Finally, the findings of the local prevalence research group have been adapted into a historical setting when Gordon Hay has helped estimate the number of Americans resident in London

between 1770 and 1775 from data found in diaries and other historical sources. This period has been studied as it has been considered the 'birth' of American culture

2.2.4 Scientific assistance

National prevalence estimates in Luxembourg have been obtained using a range of methods, including the two methods promoted by the local prevalence working group. The results of two-sample, three-sample and four-sample capture-recapture analyses were checked by the co-ordinator of the local working group and advice on the use of the truncated Poisson method was also provided (see appendix 7).

With regard to the specific targeting of relevant institutions, in particular the National Focal Points of the REITOX network, this has primarily been achieved in four instances. The Spanish Focal Point, *Plan Nacional Sobre Drogas*, participated in a workgroup meeting in April 2000. Discussions about local prevalence estimation with the Irish Focal Point, *The Health Research Board*, were held in June 2000. Previously in Portugal, only one local study and one regional study had been undertaken in Portugal, both in the Setúbal area. Gordon Hay has been assisting a researcher from the University of Porto, Jorge Negreiros, in establishing a study which will estimate the prevalence of problem drug use in Matosinhos, a municipality in the Greater Porto area. Data from four sources were collected by means of a standardised questionnaire. As the data from a prison source only included males, it would not be possible to include that sample in a four-sample capture-recapture analysis and in additional, there appeared to be very few overlaps between the sources. The initial data analysis strategy therefore focussed on providing an estimate of the number of male drug users using combinations of all four sources, or using the three other sources (social services, treatment and harm reduction) to provide an estimate for males and females (see Appendix 4). The initial findings of this research are currently being documented and will inform national prevalence research in Portugal. It is intended that this initial research will form the basis of a scientific paper.

Advice and scientific support were also given to researchers in Copenhagen who had data from two sources (hospital and emergencies) over a three year period. Various analyses were undertaken, starting with the 2-sample method for each year, then various combinations of three-samples and finally a four sample method employing data from the two sources over two successive years. Advice and scientific support were also given to a postgraduate student in Dublin who was carrying out capture-recapture analyses of data on methadone treatment and hospital episodes.

As previously discussed, research projects carried out by members of the local or national prevalence working groups in the United Kingdom have been established drawing on the findings of this TSER network, for example research that has commenced in Brighton, Liverpool and London. In addition, advice and support has been given to other research teams to initiate prevalence research in the United Kingdom, including Manchester and North Somerset in England and Ballymena in Northern Ireland.

2.3 Comparison of planned activities and actual work accomplished

Compared to the planned activities outline in the previous workplans, the actual work accomplished by the group took the form of two main themes; providing scientific support to new and continuing studies and methodological advances. Support to new and continuing studies involved the further dissemination of practical information about undertaking prevalence estimation project, via a set of methodological guidelines and a review of the scientific literature available from the EMCDDA website. More detailed support has been given as responses to e-mail requests to the work group co-ordinator and others within the group. This has led to site visits and analysis of data collected in the bcal areas. Also within this theme, particularly with respect to targeting specific institutions, has been the links with the national prevalence subgroup and a concurrent national prevalence project funded by the EMCDDA. There was a significant overlap between researchers undertaking prevalence estimation at the local level and those involved in producing national prevalence estimates. This is reflected in the composition of the two working groups and this overlap has been useful in supporting existing research and initiating new research in several European Union countries. For example, in Austria, a research project lead by Alfred Uhl who is in both working groups has constructed a national prevalence from 9 regional or city estimates which employed the capture-recapture methodology. In Finland, a capture-recapture estimate for the capital was extrapolated to provide a national estimate by Päivi Partanen, who is also in both groups. Due to the nature of problem drug use in Ireland, the research findings of Catherine Comiskey in Dublin have also formed the basis of national prevalence research in Ireland. The participation of the local prevalence workgroup co-ordinator in national prevalence meetings has led to advice being given to the National Focal Points of the REITOX network and other relevant individuals on local prevalence estimation.

3 List of Project Deliverables

The main project deliverables have been the production of new or updated local prevalence estimates by the members of the workgroups. Some of this research has been published in peer-reviewed journals, however at this stage, some of the findings have yet to be published in the scientific literature and currently take the form of reports. The project deliverables are listed in a following section.

4 Exploitation and dissemination of results

The potential for exploitation of the results of the workgroup are great. The EMCDDA website provides a resource for those interested in undertaking a local prevalence research project. Additionally, it is aimed to publish selected reports as scientific papers. Potential users for the results of the workgroup include the National Focal Points in the REITOX network. This has been done at the workgroup level by specifically targeting the Irish and Spanish Focal Points. The members of the workgroup are also in contact with potential users at a national level, primarily within the framework of Expert Groups established by each Focal Point to improve the scientific quality of the EMCDDA's key indicators, one of which is problem drug misuse prevalence.

The following papers and reports have benefited from the TSER network:

Buster MCA, van den Brink W (2001) Roaming through methodology. XXXI. Estimating partly hidden populations: heroin addicts in Amsterdam Nederlands *Tijdschrift Voor Geneeskunde* 145, 164-166.

Comiskey, C.M and Miller, R (2000) Drug use in young people and its effect on early school leaving. Department of Health, Dublin

Comiskey, C.M. (2001). Methods for estimating prevalence of opiate use as an aid to policy and planning. *Substance Use and Misuse*, 36, 1310150.

Comiskey, C.M. and Barry, J. (2001). A capture recapture study of the prevalence and health implications of opiate use in Dublin. *'European Journal of Public Health'* 11, 198-200.

Flavell, J and Hay, G. (2001) Using Capture-Recapture Methods to Reconstruct the American Population in London in the Late Colonial Period. *Journal of Interdisiplinary History*, XXXII, 37-53.

Hay, G, McKeganey, N and Hutchison, S (2001) Estimating the National and Local Prevalence of Problem Drug Misuse in Scotland. Scottish Executive, Edinburgh

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., Bello, P-Y., Comiskey, C., D'Ippoliti, D., Freire, S, Partenen, P., Seidler, D., Uhl, A., Domingo-Salvany, A., Smit, F., Toet, J and Wiessing, L. 'Estimating the prevalence of opiate use in seven European cities.' (in preparation).

Hay, G and McKeganey, N. (2001) The attendance pattern of clients at a Scottish needle exchange. *Addiction*, **96**, 259-266.

Hay, G and Smit, F 'Estimating the prevalence of drug injecting using needle exchange data' (submitted)

Hay, G. Comiskey, C., Domingo-Salvany, A., Païvi Partanen, P., Smit, F., Uhl, A. and Wiessing, L. 'Estimating Prevalence of Problematic Drug Use at the Local Level in Countries of the European Union.' Abstract submitted to 11th International Conference on the Reduction of Drug Related Harm, Jersey, April 9-13th, 2000.

Hoebe CJPA, Smit F, Vermeulen CMJH, et al. (2001) HIV-positive drug users in South Limburg: number and characteristics; a capture-recapture analysis. Nederlands Tijdschrift Voor Geneeskunde, 145, 1118-1122.

Partanen, P., Kinnunen, A., Leinikki, P., Nylander, O., Seppala, T., Simpura, J., Virtanene, A., Valkki. (1999). Report on the number of amphetamine and opiate users in the Greater Helsinki area and in the whole of Finland in 1997. Aiheita19. States, Helsinki (in finnish).

Partanen P. 'Bayesian estimation of drug use from Finnish register data 1997.' (submitted)

Uhl, A. and Seidler, D. (2000) 'Prevalence Estimation of Opiate Addiction in Austria', LBISucht, Vienna.

Uhl, A. (1998) Capture-Recapture Methodology – How Reliable is the Approach? Paper to EASAR conference, Rome.

5 Managerial and coordination aspects

Over the course of the project, two meetings were held for all work-group members, one in Dublin and the other in Barcelona. Regular contact has been maintained by e-mail. Work-group members were also invited to a meeting of EMCDDA national prevalence estimation representatives and TSER national prevalence work-group members. Smaller meetings were also held between workgroup members, for example Gordon Hay met with Catherine Comiskey and Yoon Choi in Bristol in March 2001 and Gordon Hay visited Porto, Portugal to provide support and assistance to Jorge Negreiros.

APPENDIX 1

Bayesian Estimation of Drug Use from Finnish Register Data 1997

Päivi Partanen, Department of Statistics, University of Jyväskylä

Introduction

Closed population capture-recapture techniques, used in conjunction with log-linear modelling, have become one of the standard methods for estimating the prevalence of problem drug use at the local level (EMCDDA, 1998). The method combines data from several information sources, typically legal or medical databases, producing estimates of the number of missing cases and the total population with the associated confidence intervals.

Models exhibiting different kind of dependencies between data sources are explored in the log-linear analysis (Fienberg, 1972; Bishop et al., 1975; Cormack, 1989,1993). Hook and Regal (1997) have considered the validity of model selection methods and weighting for model uncertainty in capture-recapture estimation. In this classical, non-Bayesian approach, the procedure of finding a single optimal model can be based on the value of the likelihood ratio statistic, G^2 or on information criteria as AIC (Akaike, 1985) and its alterations (i.e. BIC by Schwarz (1978) and DIC by Draper (1995)). The optimal model is the one with the lowest value of the associate information criterion. Moreover, the model selection is guided by investigator's understanding of the data sets and his or her beliefs of possible relationships between registers.

Most of the considerations of developing the capture-recapture method have been made in the field of classical statistics. Still, the methodology involves several limitations and complexities, especially when applied to human populations (Hook and Regal, 1995). Bayesian applications of the technique are few. In some cases there may be substantial prior information and taking it into account would improve the performance of the estimator. Often historical information exists regarding population size. On other occasions there may be subjective notions as to which values of the population size are more plausible than others. In contrast to the classical analysis, Bayesian formulation enables this prior information to be utilised in the modelling.

Bayesian Methods

The steps of Bayesian data analysis can be described as follows:

1. Obtain the prior density distribution. This distribution expresses what is known about the unknown parameters prior to observing the data.
2. Obtain the likelihood function. This is the sampling distribution describing the process giving rise to the data in terms of the parameters. Sample data are used to update the prior information
3. Apply Bayes' theorem to derive the posterior density distribution. This distribution express what is know about the parameters after observing the data. In other

words, uncertain prior belief is updated by sample data to form posterior conclusions.

4. Derive inference statements. The posterior density distribution contains all that is known about the parameters following the observation of the data. One illustrative summary is to plot the density function showing the density and range of values. Other useful summaries include the mode, median or mean values of this distribution, or quantiles.

Data

The Bayesian methodology can be illustrated by analysing three sample capture-recapture data to estimate the total number of amphetamine and opiate users in the greater Helsinki area in 1997. The data includes cases from three official sources: the hospital patient discharge register (HILMO), criminal report register (RIKI) and substance abuse database of persons suspected of driving under the influence of drugs (DUID). A total of 833 users were identified by the registers. The data are given below

	DUID				
	Present		absent		
	RIKI				
HILMO	present	Absent	present	absent	total
present	4	7	24	273	308
absent	56	101	418	X	575
total	60	108	442	273	883

Before 1997, the prevalence of heavy drug use in Helsinki was estimated using classical methods in 1995 to be 3109 (95% CI 2278-4449). All indicators are showing an increase in the use of drugs in Finland, and the classical estimate for 1997 was 5368 (95% CI 4037-7412).

Three different prior distributions were selected for the analysis:

1. A uniform prior with the total population size, N, lying equally likely in the interval (3500,6000), based on the assumption that the number of drug users has increased.
2. An informative prior based on the estimate for 1994 and letting the prior distribution be a negative binomial distribution with mean 3000 and a standard deviation of a) 550 or b) 1000.
3. An informative prior based on the researcher's subjective conceptions

Results

The table below summarises the results of the Bayesian analyses

Posterior Distribution	Bayes estimate (model average)	Quantiles 2.5%-97.5%
1. Uniform	4312	3552-5723
2a. Neg. Bin. (3000,550)	2508	1614-3965
2b. Neg. Bin. (3000,1000)	1354	1014-3945
3. Subjective	3401	2238-5157

These estimates can be compared with the point estimate from the classical analysis where the population of drug users was 5368 (95% CI 4037-7412). This estimate is larger than those obtained by the Bayesian analysis. When the values of N are let to be equally likely (the uniform prior) the Bayesian results are nearest to the classical results. In the classical analysis, a log-linear model representing interaction between RIKI and DUID data sources was clearly dominating. The Bayesian analyses also showed that this model makes a considerable contribution to the averaged posterior, however the Bayesian analyses included information from other models which reduced the size of the estimates.

Assuming that all log-linear model are equally likely may not be realistic. If models which were clearly unrealistic were eliminated then the results would be more reliable. Reducing the number of models in the analysis would diminish the significance of the prior distribution, however this is not possible by examining the information included in the data, rather information should be elicited from experts working in the drug field.

APPENDIX 2

Estimating the Number of Opiate Users in Amsterdam by Capture-Recapture: the Importance of Case Definition

Marcel Buster, Municipal Health Service, Amsterdam

At the meeting of the local prevalence estimations working group I addressed two problems concerning the three sample capture recapture method that I encountered when applying this method in Amsterdam. The first problem is related to the violation of the closed population assumption. The second to the definition of the population of interest in relation to the samples that are used.

I performed this Capture –Recapture estimation with a hospital, police and treatment sample of opiate users. All registers are available at the Municipal Health Service (MHS). The hospital sample consists of opiate users that are admitted in an Amsterdam Hospital. If an opiate user is admitted, the MHS is informed and assistance is given in regard to methadone dose, handling personality disorders etc. The police sample consists of opiate users that received methadone at a police station while arrested. When a drug user is arrested the physician of the MHS visits the police station. If necessary methadone is prescribed in order to prevent withdrawal symptoms. The treatment sample consists of opiate users that are treated at the MHS, the general practitioner (GP) or an abstinence oriented treatment centre. GP treatment requires a more socially stabilised life style and discipline in taking methadone. For them methadone is distributed at the pharmacist once a week. The abstinence oriented treatment centre requires motivation to stop the use of drugs. No requirements (except a opiate addiction, registration, periodical medical check up and X-thorax screening) are necessary at the MHS. However there is a limited accessibility for non-residents, they can only receive treatment in case of prostitution or severe co-morbidity. Treatment at the MHS is called low-threshold treatment and treatment at the GP and abstinence oriented treatment higher threshold treatment. When these three samples are combined eight possible combinations of appearance within different samples exist. Varying from registration in all three samples (police (P), hospital (H) and treatment (T)) to registration in none of the samples. The latter however is unknown and has to be estimated.

The first problem is the special vulnerability to the violation of the closed population assumption of one cell when applying the capture recapture method, the *relative* number of opiate users appearing at the Police and Hospital samples but not at the Treatment sample (P+,H+,T-) decreases as the study period increases. Opiate users that are temporary residing in Amsterdam have limited access to treatment. During their stay in Amsterdam they are however, at risk to be arrested by the police or admitted at a hospital. As the study period increased the proportion of opiate users that is not at risk during the whole study period will increase as well. Those opiate users who visit Amsterdam for a short period (and do not participate in treatment) are at low risk to appear both at the police station and hospital sample.

Treatment participants generally live in Amsterdam. When the study period increases the proportion that is arrested increases and the proportion admitted in a hospital will increase as well. Among treatment participants the risk to be both arrested and admitted is the product of these two proportions. When applying a loglinear model in

a situation with a long observation (registration) period, there appears to be a positive dependency between Treatment and Hospital and Treatment and Police (the model is focussed on the low number of cases within the T-,H+,P+ cell). The model that includes these dependencies will show the best fit but will overestimate the contents of the zero cell (the hidden population). The saturated model will overestimate its size as well.

When the treatment period is limited to three months, the best fitting model (as evaluated with the Aikaike's and Bayesian Information Criterion) changes, the best model has only interaction; the positive relation between Hospital and Treatment sample. In this example the effect of the violation of the closed population assumption is not limited to the overestimation of the hidden population (this would happen when applying a two sample C-RC) but while a specific subgroup (those not in treatment) is more dynamic (less closed) than the other (those in treatment) the violation of this assumption alters the dependencies between the different subgroups. These dependencies however, are biased and adjustment for these dependencies lead to a more severe overestimation.

There are two ways to reduce this bias, the first is the limitation of the sampling period. The second is a limitation of the target population. For example to limit the sampled opiate users to those that are registered in the local population register.

Table 1: the 3 – sample capture recapture; results of the loglinear analysis

		N=3141		N=2380		N=1814	
		1 year		½ year		¼ year	
	df	Llh	N	llh	N	llh	N
Crude	3	60.2	5242	24.0	4481	6.6	3976
H*T	2	34.1	5411	12.6	4650	1.6	4130
H*T + P*T	1	0.4	9665	0.0	8904	0.2	5549
Saturated	0	0.0	9102	0.0	8341	0.0	6115

Llh= loglikelihood

The Second problem has to do with the correct matching between the target population and the sampled population. I started the study with the idea of estimating two populations, a population of problematic opiate users and the total population of opiate users. In the first situation all sampled opiate users should be problematic opiate users. Opiate users arrested at the police station or admitted at a hospital were defined as problematic (with judicial or medical problems at least). Moreover, opiate users that participated at the low threshold treatment were considered to be problematic (otherwise they would have been treated at their GP or abstinence oriented treatment). When estimation the total number of opiate users I included both low threshold and higher threshold treatment participants. The samples and estimated populations are shown at figure 1. I used a three months observation period and the contents of the different cells are shown at table 2.

Figure 1: the samples and estimated populations

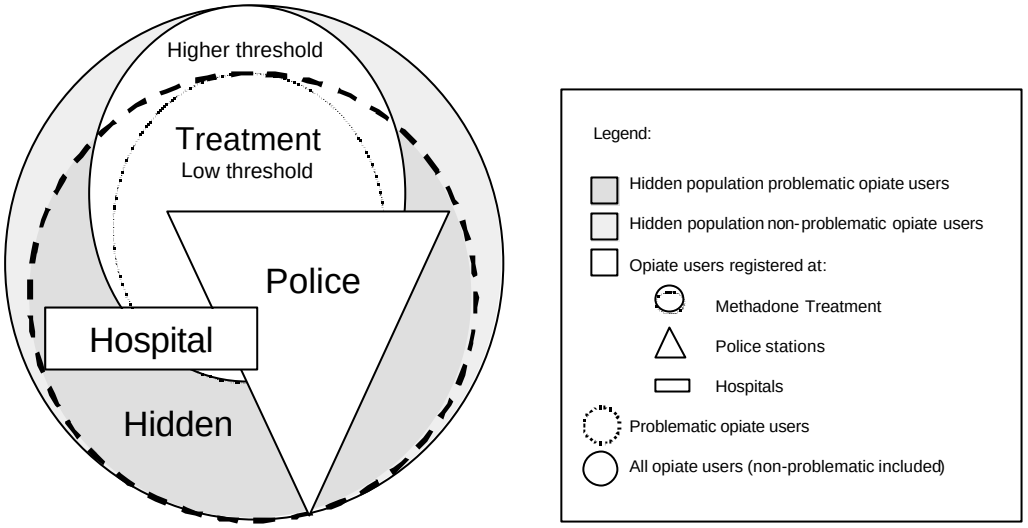


Table 2: Description of the study-population registered during the first three months of 1997

	H + T + P +	H+ T + P -	H + T - P +	H - T + P +	H + T - P -	H - T + P -	H - T - P +	*N*
Treatment								
Low threshold	8	36	6	203	52	1,078	431	1,814
Total Treatment	9	50	5	245	38	2,125	389	2,861

P = Police, H= Hospital; T= treatment ; N= total

The results of the loglineair analysis of both populations are shown at table 3. Both analyses are comparable regarding the model and information criteria. However, the estimated numbers are different. Not taking into account the 95% Confidence interval we could conclude that the estimated number of non problematic opiate users is approximately 6,185 – 4,130 = 2,055.

Table 3: C-RC Estimations of problematic opiate users and all opiate users

	Problematic opiate users	**All opiate users**
Best fitting model	H+P+T+(H*T)	H+P+T+(H*T)
Degrees of freedom	2	2
Likelihood Ratio	1.62	1.39
AIC	-2.38	-2.61
BIC	-13.4	-14.5
Estimated Number	4,130 (95% CI: 3,753-4,566)	6,185 (95% CI: 5,697-4,766)

This however is a misleading notion, the first estimation is based on samples that are restricted to problematic opiate users. Part of the higher threshold clients are non-

problematic methadone clients however, are at low risk of appearing in hospital or police sample.

When this proportion of these clients among the total treatment sample is large, the estimation of the total number will be large as well. However, if stable methadone clients become stabilised by methadone treatment, this proportion of non-problematic methadone users will not be related to the proportion of non-problematic heroin users outside treatment

The estimated number of 6,185 opiate users is considered to be invalid because it is based on an analysis without a correct matching of the sampled population and target population. Although statistical criteria indicate that both capture recapture estimates are correct, we can only know whether they are correct when the sampled population matches correctly with the target population. If only high threshold treatment is available, the use of this treatment population may result in over-estimations. In this case it is not possible to estimate the non-problematic heroin users (and thus the total population of opiate users) with a capture recapture estimation.

C-RC is popular because existing registers can be used. Data are registered during a long period of time. In Amsterdam a large numbers of opiate users are registered and we were able to restrict the observation period. Moreover, because of the differentiated treatment system we were able to exclude part of the treatment population. This enabled us to show the different (false) results that can be obtained when a long time period or the total treatment population is used.

APPENDIX 3

One Sample Capture/Recapture Methods for Estimating the Prevalence of Hidden Problematic Heroin Use and Heroin Traffic offence in Western Australia

Dr. Yoon Choi
Department of Mathematics, National University of Ireland, Maynooth

The Capture/Recapture methodology has been used for estimating the prevalence of heroin use. The European Union has endorsed this methodology in the European Monitoring Centre for Drugs and Drug Addiction guidelines (EMCDDA, 1997) for estimating prevalence at local level. One key requirement of the method is the availability of three data sources, a requirement which restricts the use of the method. The less publicised but also EMCDDA endorsed, truncated Poisson methods with a single data source, is developed and critically analysed with data from the Western Australia Police Service.

Data on people arrested, summoned, and cautioned for heroin related offences, were extracted from the Offence Information System (OIS) of the Western Australian Police Service, for the periods of the financial years, 1996/97 and 1997/98. The study population to consider for the study is special subgroups of opiate users consisting of the heavy problematic users and heroin traffic offenders. Using the truncated Poisson method, the true numbers of the study population during the study periods.

The dilemma of the law enforcement is to interpret the numbers of offenders during the period when a new policy was introduced. Upon the change of the number of notified offenders to the authority whether increased or decreased, the new policy can not be evaluated without revealing the change in the true number of offenders. For example, the number of notified offenders could be increased while the true number of offenders decreased. Hence, an advantage of applying the truncated Poisson method with a multiple time frame is to estimate this change and hence to aid the law enforcement to evaluate their policy changes during study period.

The Capture/Recapture methods with a single data source, truncated Poisson methods, are applied to a frequency distribution generated by the number of appearances of individuals within the data. That is, we count the number of individuals who have been seen once, twice, three times, etc, in the data. The zero frequency class consists of individuals who were 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 observed plus the number of individuals never seen, that is the hidden population. The estimation problem then becomes a problem of estimating the number of individuals in the zero-frequency class from the truncated series of individuals seen. We assume that this frequency distribution follows a Poisson distribution which is truncated below one. Using the truncate Poisson distribution, observed, the size of the total population is estimated using the Chao's (1987 and 1989) and Zelterman's (1988) truncated Poisson methods. Using the estimated total population the expected Poisson distribution (with the frequency zero class) can be constructed. The null hypothesis that these observed and expected truncated Poisson distributions have the same mean has to

be examined using the Chi-square goodness of fit test with an alpha value 0.05 to validate the estimates.

The Chao's (1987 and 1989) and Zelterman's (1988) truncated Poisson methods emphasise the frequency classes one and two since people observed once or twice are more resemble to the people who were never observed. These estimators are known to be fairly robust in the sense that both will underestimate the true population size in the presence of heterogeneity. Which means that they give at least the minimum bound of the study population. When the truncated Poisson distribution fits data well, these estimators would produce similar results. This can be proved theoretically.

There are three assumptions for these truncated Poisson estimators. The population is closed, the population is homogeneous, and the probability of individuals being observed and re-observed is constant during the study period. Stephen (1996) suggested that if the time frame of sampling is not excessively long, therefore we may assume that the first assumption holds. To reduce the violation of the second assumption we stratify the data according to the age and gender. The third assumption is that there is no behavioural response to the previous capture. In reality, 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). Increase in the recapture probability may occur since the police gain more knowledge on individuals who were previously arrested and the individuals may become more vulnerable to police arrest.
Goodness of fit tests, c^2, and their associated *p* values indicate that all the models fit the data well during the first financial year. During the second financial year two estimators with the stratified population by gender and age are not in the 95% acceptance region, hence these estimates are excluded. Another finding from this phenomenon is that if the model with unstratified data does not fit data well then models with stratified data would not fit the data either. If the model with unstratified data does not fit the data, then the models with one or more stratified groups do not fit the data and cause the unfitness of the model with the unstratified data.

The estimates of the total study population are 1,175 ~ 1,377 and 1,415 ~ 1,619 during the 1996/97 and 1997/98 financial years respectively. The total ranges of the 95% confidence intervals are 783 to 2,372 for the first financial year and 1,065 to 2,508 for the second financial year. However all 95% confidence intervals for the estimates have a common overlapping range of values there are 1,012 to 1,423 for the first financial year and 1,365 to 1,712 for the second financial year. Prevalence rates indicate that there were between 1.08 to 1.26 offenders per 1000 population who were problematic heroin users or heroin traffic offenders during the first financial year and between 1.27 to 1.45 offenders per 1000 population during the second financial year in WA.

The most hidden groups during the first period was females aged between 35 and 55 years and males aged between 25 and 34 years. The most hidden group during the second period changed to the group of females aged between 25 and 34 years. The most notified group during the second period was males aged between 35 and 55 years. During the second period, this group changed to the group of females aged between 35 and 55 years. These findings indicate the change of the implemented

policy effectiveness on different age and gender groups from the first to second periods.

The total study population has increased between 15.2% and 30.1% from the first period to the second period, while the proportion of the hidden population has decreased between 9.2% and 26.7% during the two periods. This indicates that the 45% increase in the known population do not tell the authority the change of the total population. These findings also indicate the limitation of the multiplier method.

The truncated Poisson methods may be used to estimate the true number of offenders in a community since it is acknowledged that there are far more offenders than notified (Webster, 1979; Blumstein & Cohen, 1978).

References

Blumstein, A. and 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** pp783-791.

Chao, A. (1989). Estimating Population Size for Sparse Data in Capture-Recapture Experiments. *Biometrics*. **45** pp427-438.

EMCDDA (European Monitoring Centre for Drugs and Drug Addiction). 1997. *EMCDDA Project (Final Report): Methodological Pilot Study of Local Level Prevalence Estimates*. Lisbon.

Webster, W.H. (1979). Crime in the United States, 1978. U.S. Department of Justice, Federal Bureau of Investigation, Washington, D.C.

Zelterman, D. (1988). Robust estimation in truncated discrete distributions with application to capture-recapture experiments. *Journal of Statistical Planning and Inference.* **18**, pp225-237.

APPENDIX 4

The prevalence of problem drug use in Scotland

Gordon Hay, Maria Gannon and Neil McKeganey
Centre for Drug Misuse Research
University of Glasgow

Information on the prevalence of problem drug use is vital for planning the provision of drug treatment agencies and to serve as baseline data when examining drug-related morbidity and mortality. Prevalence information is of most worth at the local level where decisions are made about the response to drug misuse. In Scotland, as in the rest of the United Kingdom, Drug Action Teams provide a co-ordinated response to drug misuse at the local level. Information on the prevalence of drug misuse at the national level is also important, particularly in assessing government strategies. Although the crime surveys can examine drug use, few respondents report heroin use and information is not presented at the Drug Action Team level. We have used the capture-recapture method to provide estimates of the prevalence of problem drug use in of the 22 Drug Action Team areas of Scotland.

Methods and results

The application of the capture-recapture method to estimate drug use prevalence has been described elsewhere [1]. In most areas we have collated data on drug users in contact with drug treatment agencies and general practitioners and those identified from the Police for possession of drugs. We have also employed data from Social Work Departments on drug users committing crimes, this being the Scottish equivalent of probation data. While data on new contacts to drug treatment services could be collated directly from the Scottish Drug Misuse Database, this source was augmented by data on all current clients collected directly from over 70 agencies across Scotland. Within this study we defined problem drug use to be the misuse of opiates or benzodiazepines, including methadone. Although increasing, the use of crack cocaine in Scotland is comparatively rare. While there are other forms of drug use that can be considered problematic, such as amphetamine injecting, this definition has been used in order to make comparisons with previous research in Scotland. In the Orkney Isles and the Western Isles, there were insufficient data available to use the capture-recapture method therefore the prevalence of problem drug use was estimated by comparing the ratio of known to hidden problem drug users found elsewhere in Scotland.

After eliminating multiple occurrences between data sources, 22,975 individuals were identified from treatment agencies, general practitioners, the Police and criminal justice sources as misusing opiates or benzodiazepines. The data on this known population was used to estimate that there were 55,800 problem drug users in Scotland in 2000. The 95% confidence interval for this estimate is 43,591-77,697, as found by following the method of Cormack [2] and summing the lower and upper bounds of each local estimate. As seen in Table 1, the prevalence estimates vary across the Drug Action Team areas, with the lowest being found in the Islands and other rural areas such as Highland. The cities of Dundee, Aberdeen and Edinburgh

also have high levels of problem drug use, where to the West of Glasgow, the prevalence of problem drug use in Argyll & Clyde is higher than the Scottish average.

Table 1 Estimates of the number of problem drug users by Drug Action Team Area

Area	Known	Total Estimate		*Population* (age 15-54)	Prevalence	
		n	95% CI		%	95% CI
Aberdeen City	1,194	3,645	2,659-5,965	123,240	3.0	2.2-4.8
Aberdeenshire	639	1,372	1,091-1,808	128,371	1.1	0.8-1.4
Argyll & Clyde	2,324	5,405	4,383-7,631	231,543	2.3	1.8-3.2
Angus	291	702	558-916	58,395	1.2	1.0-1.6
Ayrshire & Arran	1,171	3,058	2,451-3,932	200,291	1.5	1.2-2.0
Borders	106	585	297-1,503	54,800	1.1	0.5-2.7
Dumfries & Galloway	651	1,179	1,034-1,368	73,642	1.6	1.4-1.9
Dundee City	899	2,700	1,828-4,523	76,509	3.5	2.4-5.9
East Lothian	239	779	551-1,180	49,422	1.6	1.1-2.4
Edinburgh City	2,536	5,872	4,754-7,573	271,103	2.2	1.8-2.8
Fife	1,348	2,867	2,355-3,636	192,389	1.5	1.2-1.9
Forth Valley	730	2,208	1,363-4,516	154,777	1.4	0.9-2.9
Greater Glasgow	7,248	15,975	13,797-19,069	519,332	3.1	2.7-3.7
Highland	358	1,029	571-1,327	111,033	0.9	0.5-1.2
Lanarkshire	1,828	5,076	3,782-7,373	317,855	1.6	1.2-2.3
Midlothian	246	729	452-1,450	46,673	1.6	1.0-3.1
Moray	107	398	247-731	45,533	0.9	0.5-1.6
Orkney Isles	12	29		10,163	0.3	
Perth & Kinross	385	902	623-1,486	69,965	1.3	0.9-2.1
Shetland Isles	54	109	66-237	12,359	0.9	0.5-1.9
West Lothian	402	1,116	835-1,579	92,512	1.2	0.9-1.7
Western Isles	27	65		13,775	0.5	
SCOTLAND	**22,795**	**55,800**	**43,591-77,697**	**2,853,682**	**2.0**	**1.5-2.7**

In the Orkney Isles and Western Isles Drug Action Team areas, there were insufficient data to perform capture-recapture analyses. Therefore the prevalence estimates in these areas were obtained by applying the known to unknown ratio of problem drug users found elsewhere in Scotland to the number of known drug users in those areas.

Discussion

The capture-recapture method, when applied in a systematic manner, can inform the work of Drug Action Teams by providing estimates of the prevalence of problem drug use in their area. Local estimates can be combined to provide a national estimate.

References

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

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

APPENDIX 5

Estimating the Prevalence of Problem Drug Use in Matosinhos, Portugal

Jorge Negreiros, University of Porto
Gordon Hay, University of Glasgow

Data on 402 problem drug users were collated by means of a standardised questionnaire sent to four services; a Treatment centre, a Social service, a Prison and a Harm Reduction centre. There were 448 records. 13 were discarded as they did not have a date of birth. There appeared to be 3 people who were in the Social Service data twice and 1 person in the Harm Reduction data twice. These duplicates were discarded. The following overlap table was produced.

Overlap between 4 data sources

		Source 1	Present	Present	Absent	Absent
		Source 2	Present	Absent	Present	Absent
Source 3	Source 4					
Present	Present		0	0	0	3
Present	Absent		0	2	0	37
Absent	Present		1	0	13	58
Absent	Absent		9	71	208	0

Source 1 = Social Services
Source 2 = Treatment Centre
Source 3 = Prison
Source 4 = Harm Reduction Centre.

In total, there were 402 individuals identified from the 4 data sources. On initial inspection of the table, many cells are zero therefore a four-sample analysis may not be succesful, however when this analysis was undertaken, it was possible to fit log-linear models to the data. The best 6 models, along with the independence model, are shown below.

Results from 4 sample analyses

Model	Deviance	Df	Estimate	Total	P	AIC
Independence	20.71	10	1579	1981	0.023	0.71
S2xS3	9.15	9	1265	1667	0.424	-8.85
S2xS3+S1xS2	8.51	8	1323	1725	0.385	-7.49
S2xS3+S3xS4	8.77	8	1144	1546	0.012	-7.23
S2xS3+S1xS3	6.93	8	1631	2033	0.362	-9.07
S2xS3+S1xS4	**5.86**	**8**	**1129**	**1531**	**0.544**	**-10.14**
S2xS3+S2xS4	9.15	8	1260	1662	0.663	-6.85

From this table, there are 1.129 'hidden' drug users and thus 1.531 in total. The 95% confidence interval for the hidden population is [737-1.755]. The 95% confidence interval for the total population is [1.139-2.157].

Because of the many zero cells, sources were combined within a 3 sample capture-recapture. The best way to combine the data sources was to combine the Social Service and the Treatment data into a single source. The following table can therefore be produced:

Overlap between 3 data sources

	Source 1	Present	Present	Absent	Absent
	Source 2	Present	Absent	Present	Absent
Source 3					
Present		0	14	3	58
Absent		2	288	37	-

Source 1 = Combined Social Services and Treatment
Source 2 = Prison
Source 3 = Harm Reduction

The known population is 402.

Results from 3 sample analyses

Model	Deviance	Df	Estimate	Total	P	AIC
Independence	6.99	3	1611	2013	0.072	0.99
S1xS2	5.76	2	1764	2166	0.056	1.76
S2xS3	**0.75**	**2**	**1116**	**1518**	**0.687**	**-3.25**
S1xS3	5.02	2	2664	3066	0.081	1.02
S1xS2+S2xS3	0.30	1	1193	1595	0.584	-1.70
S1xS2+S1xS3	1.27	1	5328	5730	0.260	-0.07
S1xS3xS2xS3	0.19	1	715	1117	0.663	-1.81
Saturated	0.00	0	0	402		0.00

From this table, the estimated hidden population is 1.116 giving a total estimate of 1.518. The confidence interval for this hidden population estimate is [646-1.977] and the confidence interval for the total population is [1.048-2.379]. From this 3 sample analysis, The weighted Bayesian estimate suggests that the total population size is 1.878.

As the Prison data only included males, the analysis were replicated for males only.

Overlap between 4 data sources- Males only

		Source 1	Present	Present	Absent	Absent
		Source 2	Present	Absent	Present	Absent
Source 3	Source 4					
Present	Present		0	0	0	3
Present	Absent		0	2	0	37
Absent	Present		1	0	11	48
Absent	Absent		7	66	193	0

Source 1 = Social Services
Source 2 = Treatment Centre
Source 3 = Prison
Source 4 = Harm Reduction Centre.

In total, there were 368 males identified from the 4 data sources. Again, as many cells are zero therefore a four-sample analysis may not be succesful, however when this analysis was undertaken, it was possible to fit log-linear models to the data. The best 6 models, along with the independence model, are shown below.

Results from 4 sample analyses - Males only

Model	Deviance	Df	Estimate	Total	P	AIC
Independence	20.32	10	1541	1909	0.026	0.32
S2xS3	9.22	9	1205	1573	0.417	-8.78
S2xS3+S1xS2	8.25	8	1283	1651	0.409	-7.75
S2xS3+S3xS4	8.22	8	1015	1383	0.017	-7.78
S2xS3+S1xS3	**6.76**	**8**	**1593**	**1961**	**0.412**	**-9.24**
S2xS3+S1xS4	7.06	8	1093	1461	0.563	-8.94
S2xS3+S2xS4	9.22	8	1203	1571	0.530	-6.78

From this table, there are 1.593 'hidden' drug users and thus 1.961 in total. The 95% confidence interval for the hidden population is [894-3.007]. The 95% confidence interval for the total population is [1.262-3.375].

Again, the 3 sample analysis by combining the Social Service data with the treatment data to get the following table for males

Overlap between 3 data sources- Males only

	Source 1	Present	Present	Absent	Absent
	Source 2	Present	Absent	Present	Absent
Source 3					
Present		0	12	3	48
Absent		2	266	37	-

Source 1 = Combined Social Services and Treatment
Source 2 = Prison
Source 3 = Harm Reduction

The known population is 368.

Results from 3 sample analyses - Males only

Model	Deviance	Df	Estimate	Total	P	AIC
Independence	7.34	3	1494	1862	0.062	1.34
S1xS2	5.76	2	1672	2040	0.056	1.76
S2xS3	0.88	2	976	1344	0.644	-3.12
S1xS3	5.46	2	2412	2780	0.065	1.46
S1xS2+S2xS3	0.30	1	1064	1432	0.584	-1.70
S1xS2+S1xS3	1.30	1	4921	5289	0.254	-0.70
S1xS3xS2xS3	0.18	1	592	960	0.671	-1.82
Saturated	0.00	0	0	368		0.00

From this table, the estimated hidden population is 976 giving a total estimate of 1.344. The confidence interval for this hidden population estimate is [543-1.802] and the confidence interval for the total population is [911-2.170]. From this 3 sample analysis, The weighted Bayesian estimate suggests that the total population size is 1.690. The associated 95% confidence interval for this weighted estimate for the total population is [1.079-3.333], which is quite wide.

From these different tables, the 3-sample estimates for males only appears to be the preferred analysis; this suggests that the hidden population of males is 976, giving a total estimate of 1.344 [95% CI 911-2.170]. An associated weighted Bayesian estimate for that overlap data is 1.690.

APPENDIX 6

An Estimate of the Prevalence of Opiate Use in Dublin, 1997 and a look at the progression to Treatment for Opiate Use

Elaine Hand
Mathematics Department
National University of Ireland, Maynooth

Objective

To estimate the prevalence of opiate drug misuse in Dublin, Ireland in 1997. By implementing the capture-recapture method with three data sources, figures from drug related deaths, drug treatment agencies and hospital admissions will be utilised to derive estimates of this hidden population. Estimates will be based on age and sex of drug misuser. To estimate the latency period of opiate use, and to look at the factors which affect this latency period by implementing some survival analysis techniques.

Background/Significance

Despite considerable advances in the provision of health care services to drug misusers, medical, public health and financial planning continues to be plagued by the uncertainty of the extent of the hidden drug misuse problem. The popular media of television, radio and press speak of the 'raging epidemic', the 'scale of the drugs haul' and more recently, 'people power: pushing out the pushers'. Similarly within the national and international medical and scientific press many questions on the properties and the prevalence of drug misuse have been addressed.

Johnson et al (1994) in a study of the risk behaviour in attendees at a Dublin needle exchange program speaks of the high level of unsafe injecting and sexual activity. The authors point out the need for more effective health promotion among drug users in Dublin. Comiskey (1991) and Comiskey et al (1992) in a 2 year survey of drug users estimated that a total of 375 people enter the drug using population each year with 198 of these being in the Dublin region. In addition the survey found that the age the respondants first used intraveneous drugs, ranged from 11 to 28 years of age for males and 10 to 30 years of age for females. O'Higgins (1995) in a 4 year study found that the numbers seeking treatment for the first time had almost doubled from 624 in 1990 to 1150 in 1994. This study also identified the importance of the age and sex profile of the drug misuser. While these studies provide a significant and valuable contribution to our understanding of the drug misuse profile in Dublin there is to date no comprehensive study on the prevalence of drug misuse in our capital city.

"Neither gardai nor community workers believe that Dublin's heroin supply will dry up as a result of the seizures. ... no immediate effect on the price of a street deal ... stable at £20..."

Catherine Cleary (The Irish Times, 9-Nov-1998)

"A heroin addict who used a blood-filled syringe to steal £30 from a taxi driver ..."

(The Irish Times, 30-July-1998)

The above quotes illustrate the heroin problem in Dublin, they give just a sample of the sort of news that is found everyday in the papers, and media in general. A major issue in dealing with the heroin problem in Dublin is the number of individuals who are using heroin. Another issue is the length of time individuals are using opiates for, before they seek treatment.

Comiskey [1] conducted a study to estimate the prevalence of opiate use in 1996, she estimated that there was between 12,037 and 15,306 opiate users in Dublin in 1996, however another prevalence estimate is important to support that estimate and to see how much opiate prevalence is changing from year to year.

Dean et al [3] observed the length of time patients were on drugs prior to first treatment. He observed that *"... in 1979, 68\% had been on opiates for four years or longer ... this had fallen to 40% by 1983."* Dean et al [3] also make the suggestion that those attending treatment in 1983 were doing so at an earlier stage, in their opiate using careers, than those who attended treatment in 1979. It is necessary therefore to estimate the latency period of opiate use.

Detailed description of results

The prevalence of opiate use in Dublin.

Using the capture-recapture method we got an estimate of between 18,972 and 26,708 individuals who were using opiates in the two year period 1996-1997. Using the capture-recapture method and the method of known percentages we were able to get an estimate of the number of opiate users in Dublin in 1997, which we estimate to be between 10,971 and 19,337. A summary of all the results we obtained using the capture-recapture method is shown in table 1, given below.

Year	Source	Method	Definition	Known	Estimated Total
1996	M96, H96	2-sample Capture-Recapture	Problematic User	3423	6183
1996	M96, M97, H96, H97	4-sample Capture-Recapture & known Percentage	Problematic User	3423	11270
1996	M96, H96, P96	3-sample Capture-Recapture	User	6264	13460
1997	M97, H97	2-sample Capture-Recapture	Problematic User	4525	10960
1997	M96, M97, H96, H97	4-sample Capture-Recapture & known Percentage	Problematic User	4527	14889
1996/1997 combined	H96, H97, M96, M97	4-sample Capture-Recapture	Problematic User	6250	19340
1996/1997 combined	P96, H97, M97	3-sample Capture-Recapture	User	7289	21916

M96=Methadone 1996, M97=Methadone 1997, P96=Police 1996, H96= Hospital 1996, H97=Hospital1997

Table 1: Summary of Results from Capture-Recapture Estimates

When we compare like with like, the two sample capture-recapture estimate of prevalance of problematic opiate use, we see that the prevalence has risen by over 4,000 individuals in 1997. We get the same result when we compare the estimates for the prevalence of problematic users, using the four sample capture-recapture method and the method of percentages, again there is an increase of just less than 4,000 people in 1997. Comiskeys [1] three sample capture-recapture estimate gave us an estimated prevalence of opiate use, of 13,460 for 1996, we could not obtain a comparable three sample estimate for 1997 as the police data set was not available in 1997. If we use the estimates from 1996 and 1997 that are comparable, namely the estimates derived using the ``method of known percentages", we can say that the prevalence of opiate use in Dublin did rise slightly in 1997 from the prevalence in 1996. However, all the estimates do confirm that the opiate prevalence problem in Dublin is well over 10,000 people.

The latency period of opiate use.

Using survival anaylsis and the Kaplan-Meier method we analysed the latency period of opiate use. The Kaplan-Meier fiunction was estimated using the equations given below.

$$S(t_j) = P(t_j) + P(t_{j-1}) + P(t_{j-2}) + \ldots$$

$$P(t_j) = \frac{Number\ not\ seeking\ treatment\ at\ time\ t_j}{Number\ using\ opiates\ at\ time\ t_j}$$

where $S(t_j)$is the survival function and $P(t_j)$ is the probability of seeking treatment at time t.

The average length of time that an individuals will use opiates before seeking treatment is two and a half years. Females tended to seek treatment earlier on in their drug using career than males, see figure 1. Frequency of use of an opiate also affected the duration of use of that opiate, those individuals who used opiates daily also used opiates the longest, see figure 2.

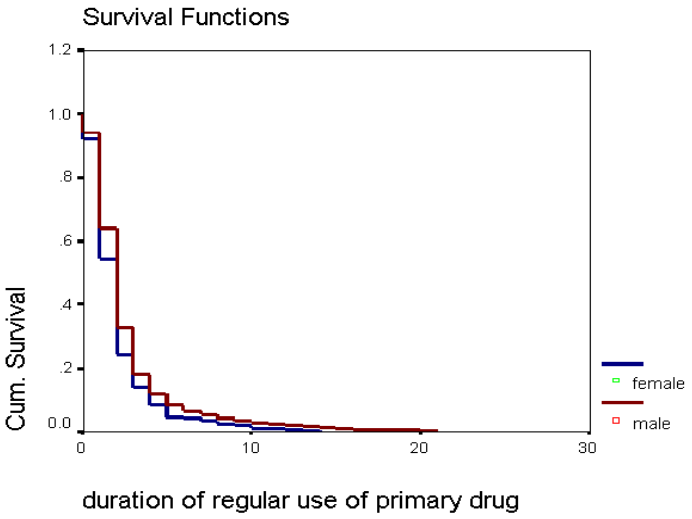


Figure 1

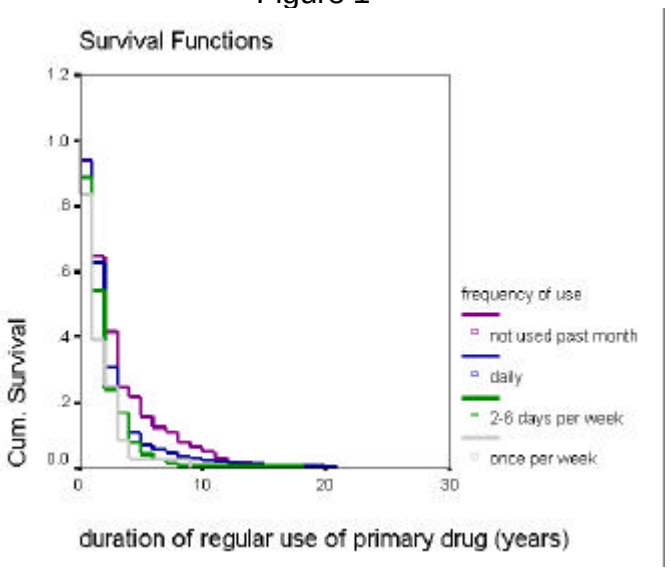


Figure 2

Discussion of Results

	1996 only (not 1997)	1996 & 1997	1997 only (not 1996)
Known prevalence of opiate use	1725	1698	2827
Estimated total prevalence of opiate use.	5338	5261	8741

Table 2: Summary of prevalence estimation results.

Using table 2, if we subtract the known number of opiate users in 1997 from the estimated total prevalence of opiate use in 1997 we can conclude that there were almost 6,000 individuals using opiates in 1997 who were not using opiates in 1996, and who had not been known to the treatment centers and hospitals in 1997. It is therefore possible to make the assumption, that these 6,000 individuals had yet to receive their first treatment.

We know that 69.9% of all opiate users who sought treatment, sought treatment, for the first time, in the first two years of opiate use. Assuming that the 6,000 individuals who did not seek treatment in 1996 or 1997 are opiate users who have not sought treatment, but who will eventually seek treatment, we can say that in the years 1998-2000 at least 4,200 individuals could be expected for first treatment, the figure will be higher when those who started using opiates in 1998 or 1999 are considered.

We know that the duration of opiate use, until first treatment, is over two years, then it is fair to say that those 6,500 individuals should be expected into treatment, on average, sometime in the year 1999 or 2000.

In our analysis of the hospital 1997 data we were able to examine the number of individuals from different postal areas of Dublin. It is clear that there are certain areas in Dublin were there are clusters of opiate users, perhaps this should also be taken into account when treatment clinics and funding is provided to local boroughs.

When we estimated the prevalence of opiate use in 1997 and stratified the results by age and sex, we discovered that the ratio of known to unknown individuals was highest for females aged 15-24. For every one female aged 15-24 years that either hospital or treatment services knew as an opiate user there was four other females in the same age group who were opiate users and unknown to the authorities. This is quite a serious result, it could be that young females are afraid of losing children or of their problem
been discovered, it is an issue which should be addressed when drug treatment services are planned.

Conclusion

This project has brought forward some interesting results. We estimated that there was approximately 14,000 opiate users in Dublin in 1997. We also discovered that an opiate user will use opiates regularly for approximately 2.5 years before seeking treatment for the first time. Some issues which it is felt should be addressed in order to try reduce the opiate problem in Dublin, would be to seriously tackle the heavy prevalence of opiate use in the city centre of Dublin and also to try and encourage young females to seek treatment.

In the past few years since the study carried out by Dean et al [3] we have seen the latency period of opiate use decrease, however the prevalence of opiate use would appear to have increased. Some research in the area of post treatment return to opiate use may prove useful. Also, given the high numbers of opiate users living in the city centre, and other deprived areas of the city it may prove useful to see if there is a relationship between drug use and poverty or unemployment. Some preliminary work in this area has been initiated by Comiskey (private communication).

References

1. C.M. Comiskey. Estimating the prevalence of opiate drug use in Dublin, Ireland during 1996. The Department of Health, Hawkins House, Dublin 2., 1998.
2. C. Rossi. Monitoring drug control strategies: hidden phenomena, observable events, observable times. *International Journal of Drug Policy,* to be published.
3. G. Dean, A. O'Hare, A. O'Connor, M. Kelly, and G. Kelly. The "opiate epidemic" in Dublin: 1979-1983. *Irish Medical Journal,* 78(4), April 1985.

APPENDIX 7

Estimating the prevalence of problem drug use in the Grand Duchy of Luxembourg

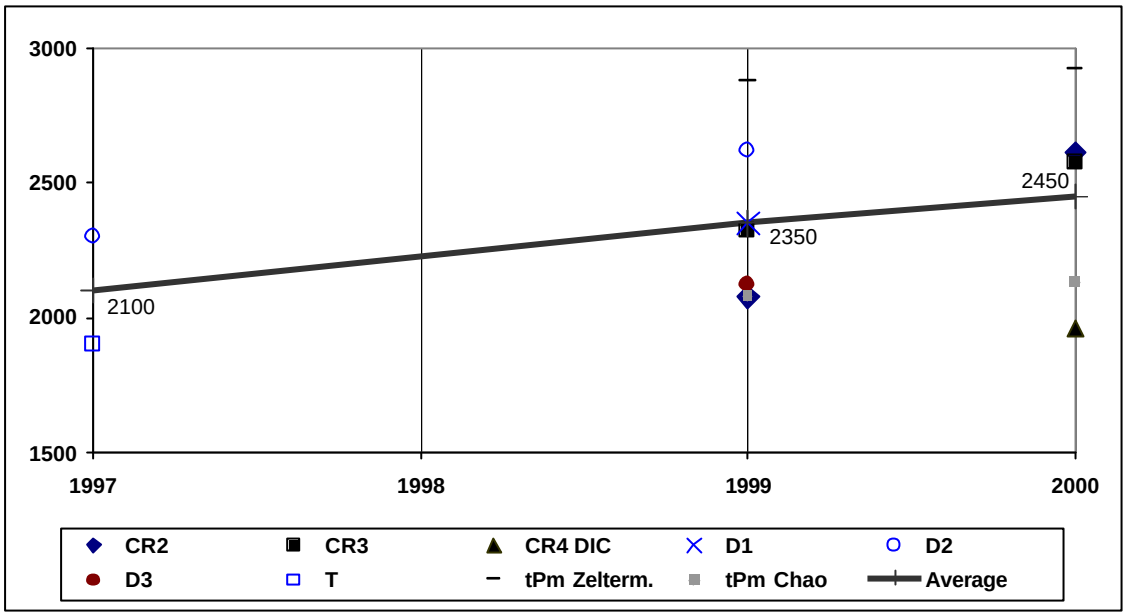
Alain Origer

A first national drug prevalence study has been conduced in 1997 by the national EMCDDA focal point. The 1997 study relied on a limited number of estimation methods and primarily aimed at the analysis of the national context with a view to the application of a multi-methods approach in coming years. Data presented in the present report have been provided by the latest drug prevalence study (hereinafter referred to "2001 study") conduced by the focal point between 1999 and 2001 (Origer, 2001) and refers to the years 1999 and 2000. The 2001 study pursued two primary objectives. It is the first comparative multi-methods drug prevalence study at the national level. Furthermore the evaluation of estimation methods in the light of national data availability and quality has allowed to define an overall methodology with respect to the follow up of national drug prevalence and incidence parameters in the future.

The research strategy relies on the methodological framework of the Luxembourgish Information System on Drugs an Drug Addiction (RELIS), set up in 1995 by the national focal point of the EMCDDA. RELIS stands for a nationwide multi-sectorial information network, including specialised drug treatment institutions, general hospitals, counselling centres and competent law enforcement agencies. As such, it provides for the most comprehensive and reliable data on problem drug users indexed by national institutions. In compliance with RELIS case definitions, the present study specifically aims at the prevalence estimation of problem use of illicitly acquired high risk drugs (HRC) in the national population aged between 15 and 54 years. The chosen terminology defines the target population with regard to the observable consequences of drug use, the nature of consumed substances as well as the context (legal or illegal) of their acquisition.

Data from 1999 and 2000 have been considered in comparison with first national drug prevalence figures from 1997. The following methods have been applied : Case finding (CF), capture-recapture on 2,3 and 4 sources (CR 2,3,4), truncated Poisson model associated to Zelterman's and Chao's estimators (tPm), and four different multiplier methods using data from law enforcement sources, drug mortality registers (D1,2,3) and treatment agencies (T).

Chart. 1. Prevalence estimations of problem HRC drug use (1997 – 2000)



Tab. 1. Prevalence and prevalence rates according to selected sub-groups (1997 – 2000)

	1997	**1999**	**2000**
GENERAL POPULATION			
National population on 1st July	421,000	432,450	438,500
National population aged between 15 and 54 years on 1st July	239,818	245,308	248,440
HRC USERS IN CONTACT WITH THE NATIONAL INSTITUTIONAL NETWORK (low threshold agencies not included)			
Total number of indexed users (multiple counts excluded)	/	1,198	1,024
Number of treatment demanders in specialised institutions		757	637
outpatient	/	*624*	*557*
inpatient	/	*218*	*178*
Number of drug law offenders (ad minima consume of HRC drug(s))	/	551	510
PROBLEM USE : HRC DRUGS			
Average prevalence	2,100	2,350	2,450
Total prevalence rate	5 /1000	5.43 /1000	5.59 /1000
Total prevalence rate - age: 15-54	8.8 /1000	9.58 /1000	9.86 /1000
PROBLEM USE : MAIN DRUG - HEROIN			
Prevalence heroin	1,680	1,975	2,010
Total prevalence rate - heroin	4 /1000	4.57 /1000	4.58 /1000
Total prevalence rate - heroin - age:15-54	7 /1000	8.05 /1000	8.09 /1000
INTRAVENOUS DRUG USE (IDU)			
Prevalence IDU	1,370	1,780	1,715
Total prevalence rate - IDU	3.25 /1000	4.12 /1000	3.91
Total prevalence rate - IDU - age:15-54	5.71 /1000	7.26 /1000	6.90

Absolute prevalence and prevalence rates of problem HRC drug use have shown a growing tendency over the past four years. The increase rate observed between 1999 and 2000 is less pronounced than the one observed during the period 1997 to 1999. The observed figures comply with the stability of heroin use and intravenous drug use prevalence between 1999 and 2000. Although the total drug use prevalence shows an upwards tendency, heroin use does not contribute significantly to the referred progression. Intravenous drug use prevalence has even shown a slight decrease in 2000.

In 1999, the average prevalence of problem HRC drugs use (2,350) and the related prevalence rate of 5.59/1000 (9.58/1000 in national population aged between 15 and 54) are considered to show good validity according to the non-contradicting estimates obtained by the multi-methods approach and the evolution of indirect indicators such as the number of fatal overdose cases, the number of distributed syringes through the national needle exchange programme, the number of HRC drug law offences and seizures and admission data of low threshold agencies. Prevalence figures calculated for 2000 (N: 2,450, total rate: 5,59/1000, rate 15-54: 9,58/1000) also fit the curve of indirect indicators. Since the 2000 figures have been obtained by a limited number of estimation methods, observed tendencies should be confirmed by further

research based on the evaluation outcome of multi-source methodologies in the light of national specificities. According to experience gathered in the framework of the present research project, the application of a routine set of methods including CR3, tPm and D1, currently appears to be a highly valuable option.