The hidden population: Some methodological issues about estimation of problematic drug use

MIKAEL DAHLBERG & MATS ANDERBERG

ABSTRACT
AIM – This article analyses and discusses the estimation of serious or problematic drug use through an empirical example based on a local Swedish study in Gothenburg. METHODS AND DATA – This was a case-finding study with questionnaires directed at organisations which have contact with the target group. The material was supplemented with information from the two documentation systems DOK and ASI. A total of 2,148 reports were collected. Health care data of 1,096 individuals was also collected for analysis with the truncated Poisson method. Analyses with capture-recapture or truncated Poisson were conducted to calculate the size of the hidden population. RESULTS – The statistical analyses resulted in variable numbers for the hidden population, and the total prevalence of serious drug abuse in Gothenburg is estimated to be between 2,200 and 4,400 people. CONCLUSION – The study shows that estimation of the presence and prevalence of problematic drug abuse involves many methodological difficulties and challenges. The significant variation of the size of the hidden population presented in the study raises doubts about the reliability and validity of the different methods. The methods are clearly sensitive to the importance of fulfilling the different basic assumptions.
KEYWORDS – Problematic drug use, estimating prevalence, hidden population, case finding, capture-recapture and truncated Poisson

Submitted 30.05.2012          Final version accepted 10.09.2012

Introduction
Criminalisation and stigmatisation of drug use have led to a situation where problematic drug use is to some extent hidden. This results in a lack of knowledge about the abuse. Some individuals are not interested in reporting their abuse to authorities because they are worried about negative consequences, and others with problematic drug use have no contact whatsoever with the social or health care system. Another explanation could be that the prevalence of drug problems can be difficult to determine. In other words, a hidden population exists.

The question of how many people there are with problematic drug use has no simple or standard answers because there are no records or documentation systems that cover everyone with these kinds of problems. The European definition of “problematic drug use” mainly refers to injecting narcotic substances and long-term or habitual use of opiates, amphetamines and cocaine. The Swedish definition of the term, which is used in this article, implies a broader understanding, which includes drugs such as cannabis and other synthetic drugs (EMCDDA, 2006). Previous Swed-
lish studies have used the term “heavy narcotics abuse” (Olsson, Byqvist, & Gomér, 1994). There are also several different methods of investigating the prevalence of specific populations of abusers (EMCDDA, 1999a). Regardless of the method chosen, the question is ultimately about different kinds of estimates, more or less reliable or valid.

However, estimating or calculating the number of people with problematic drug use in a specific region or nation is urgent. Knowledge about presence or prevalence matters in planning preventive efforts or interventions of a social, medical or psychological nature. In Swedish government reports, such estimates make up the basis for political proposals and decisions (SOU, 2005; SOU, 2011). Still, critical appraisals or analyses of the reliability and relevance of these investigations are rare. This article will draw on a study in Gothenburg to analyse and discuss estimation of problematic drug use. The purpose of the local survey is to obtain an accurate picture of the prevalence of this problem where knowledge is supposed to provide the city with a foundation of planning and development in the field of substance abuse. The idea is to continuously monitor the development by similar surveys conducted every six years.

Several studies have been undertaken about the prevalence of problematic drug use in Gothenburg, both as a detailed part of a national study and as part of the city’s own local estimates. The first national estimate of problematic drug use was given in 1967, based on data from Stockholm (Narkomanvårdskommittén, 1969). The numbers of problematic drug users have since been estimated in three other national studies in 1979, 1992 and 1998 (Olsson, Carlsson, Fant, Olsson, & Roth, 1981; Olsson et al 1994; Olsson, Adamson Wahren, & Byqvist, 2001).

In 1979, an estimated 15,000 people in Sweden had problems with drug use, while the corresponding figure amounted to 19,000 people in 1992. The latest study of this kind, from 1998, estimated that the number of problematic drug users in Sweden and Gothenburg was 26,000 and 2,100, respectively (Olsson et al., 2001). The Swedish National Board of Health and Welfare has published estimates of the number of serious or problematic drug users in 1998-2004 based on information from health care patient registers. For 2004, the estimate for Sweden as a whole was 25,600 persons (SOU, 2011). Using the same methods, the Swedish National Institute of Public Health (FHI, 2010) considered the numbers to total 29,500 in 2007. This report shows the distribution by region (administrative county), and in 2007, about 5,300 people were assessed to suffer

Table 1. Estimates of people with problematic drug use in Gothenburg and in Sweden. The data is based on the Swedish definition of problematic drug use.

| Year | Gothenburg | Sweden |
|------|------------|--------|
| 1979 | 1,800      | 15,000 |
| 1992 | 1,900      | 19,000 |
| 1998 | 2,100      | 26,000 |
| 1999 | 2,100*     | 26,100 |
| 2000 | 2,200*     | 27,700 |
| 2001 | 2,400*     | 27,700 |
| 2002 | 2,500*     | 27,300 |
| 2003 | 2,200*     | 25,700 |
| 2004 | 2,800*     | 25,600 |
| 2007 | –          | 29,500 |

*Indicates local but not published information (Svensson 2005).
from problematic drug use in the Västra Götaland region. Estimates were also produced for Gothenburg, where according to the most recent estimate, from 2004, 2,800 people were suffering from problematic drug use. It should also be noted that the information varies (during 1999-2004) between 2,100 and 2,800 persons.

The increase of problematic drug use in Sweden can also be reflected in many other figures or indicators, such as the number of reported drug offences, health care with drug-related diagnoses, drug-related deaths or compulsory care. The trajectory of problematic drug use for 1998-2009 is contradictory, as some indicators point towards an increase during the last decade while other reports show a stagnation or decrease (CAN, 2010; Socialstyrelsen, 2012).

Previous research
Case finding involves acquiring information about known individuals covered by the definition and their present problematic drug use from organisations that may come into contact with these people. Substantial problematic drug use is usually accompanied by many problems and needs in several areas of life. These include medical, psychological, legal and social problems and needs, which is why an individual may have contact with one or more organisations or groups in, for example, health care, social services, correctional services, the police force and voluntary organisations. Information obtained from only a single source may thus provide a limited picture of the population and lead to an underestimate of the total number of individuals. In order to create a higher contribution rate and thus a more certain overall picture of the extent of the problem, it is important to combine several different sources of information. The more sources, the more complete the picture is likely to be (Bloor, 2005; Hay, 1997; Simon, 1997). It can, however, be costly to administer such studies and difficult to obtain relevant information from all of the sources (EMCDDA, 1999a). Olsson (1999) maintains that case finding is far too complicated and unreliable a method to use at the national level, but considers it all the more applicable and relevant in local studies.

Capture-recapture is a method of statistical analysis used, for example, in epidemiology to catch and estimate hidden populations, often applied to populations for which basic information is lacking (Hickman et al., 2002; Hook & Regal, 1995; 2000a; Millar, Domingo-Salvany, Eastwood, & Hay, 2008). The method originated in ecology and zoology, where it is used for estimating the size of different animal populations.

The capture-recapture method is based on overlapping information from at least two sources that can then be used for estimating the number of individuals that do not occur in any of these, which in turn makes possible an overall appraisal of the total population. The technique has been used in several studies to study the prevalence of problematic drug use, HIV, homelessness and similar problems, particularly since the statistical analyses have been improved and refined (Bloor, 2005; EMCDDA, 1999a; 1999b; Hickman et al., 2002; 2006). A number of empirical studies that have specifically examined the prevalence of different kinds of problematic drug use at local and national levels have been conducted over the past dec-
ades (Choi & Comiskey, 2003; Domingo-Salvany et al., 1998; Frischer et al., 2001; Gemmell, Millar, & Hay, 2004; Hay, 1997; Hay & Gannon, 2006; Hickman et al., 1999; Hickman et al., 2006; Hser, 1993; Kimber, Hickman, Degenhardt, Coulson, & van Beek, 2008; Maxwell & Pullum, 2001; Millar, Gemmell, Hay, & Donnall, 2004; Kraus et al., 2003; Olsson et al., 1981; Olsson et al., 1994; Olsson et al., 2001; Uhl & Sieidler, 2001; Vaissade & Legleye, 2008).

The capture-recapture method is based on a number of basic assumptions that must be fulfilled or dealt with in different ways. The five following assumptions should be considered (Hickman et al., 2002):

- The classification of individuals in the population is correct
- Reports from the information sources are representative for the population
- The population is closed and static
- The population is homogenous
- There is no dependence between the sources of information

The method’s limitations primarily arise from that it is difficult to meet all of the basic assumptions (Hickman et al., 2002; Hook & Regal, 1995; Uhl & Seidler, 2001), which indicates that shortcomings in this respect may give rise to results that are uncertain and difficult to interpret (see, for example, Brenner, 1996; Chao, Tsay, Lin, Shau, & Chao, 2001; Cormack, 1999a; 1999b; Corrao, Bagnardi, Vittadini, & Favilli, 2000; Domingo-Salvany et al., 1998; Hook & Regal, 1993; 1995; 1999; 2000b; Korf, 1997; Neugebauer & Witte, 1994; Papoz, Balkau, & Lellouch, 1996; Tilling, 2001).

Truncated Poisson can be used to statistically calculate and estimate the prevalence of problematic drug use with information from a single source, such as existing information from a health care or correctional treatment register. The method is built on an estimate based on repeated events or care occasions during a certain period (Brå, 2003; Choi & Comiskey, 2003; EMCDDA, 1999a; FHI, 2010; Hay, 2003; Hay & Smit, 2003; Hser, 1993; Kimber et al., 2008). The idea is that the distribution among the individuals occurring once or twice in the register is then applied statistically to calculate those who do not occur at all, which means that non-events, or those that have hypothetically not yet been observed, can in this way be calculated (i.e. hidden populations). In this case, some of the assumptions of the capture-recapture method are also the same, for example, that the population is closed and homogenous. In addition, it is assumed that the likelihood of being included in the material is constant and that the number of individuals that occur once is in direct proportion to those who occur twice. The method is also based on the assumption that the individuals who occur three or more times in, for example, a health care register, are not representative of the population and are therefore less relevant for estimating the size of the hidden population.

Researchers seem to employ several different ways of interpreting or relating to the basic formulas that form the foundation of the method. A number of Swedish studies have used a specific application of the truncated Poisson. These studies have had great significance for the assessment of the prevalence of problematic drug use at a national level (for example, Brå, 2003;
One advantage of the truncated Poisson is that it rests on existing information, is relatively simple to use and is thus cost-efficient. For estimations of specific groups, such as injecting problematic drug users and assessment of trends, the method is considered to be particularly useful (Hay & Smit, 2003).

Truncated Poisson as a method for estimating hidden populations is criticised both on theoretical and empirical grounds. Considering that it is based on only one specific register and thus on one specific target group, and given the method’s basic assumptions on constant care opportunities, it is doubtful whether it alone can provide a representative, reliable or valid estimation (Choi & Comiskey, 2003; Hickman et al., 2002; 2006; Hickman & Taylor, 2005; Law, Degenhardt, & McKetin, 2006; Uhl & Seidler, 2001).

Aim
On the basis of an empirical example, this study aims to analyse and discuss the estimation of problematic drug use, drawing on a local Swedish study in Gothenburg which sought to sketch a reliable picture of the prevalence of the problem.

Methods
Design and selection
The data was collected between October 2010 and March 2011 in collaboration with public authorities and organisations that have contact with people who use drugs. This was a case-finding study using questionnaires aimed at those organisations expected to encounter the target group. The questionnaires were filled in by members of staff who had had face-to-face contact at least once during the given period with individuals who met the definition of problematic drug use, and where the staff were aware that drug use had occurred in the past year. Overall, 72 organisations representing approximately 120 local units participated in the study.

Some external dropouts occurred by units which did not take part in the study. The police authority in Gothenburg was not able to participate in the actual study, but they did contribute some special information that was converted into anonymous material. This information showed how many people had been arrested on suspicion of drug offences. As notations regarding actual drug use are absent, this material is not used for estimation of the unobserved population (see Choi & Comiskey, 2003; Kraus et al., 2003).

Health care services participated in the study through outpatient and institutional care units. Other units within the health care system, such as emergency treatment, the infectious diseases clinic and community health centre for homeless people, judged that it was not possible for them to participate. Also, information was not available from one correctional institution in Gothenburg.

The term “heavy narcotics abuse” is vague and can be defined or interpreted in many different ways (SOU, 2005). Internationally, this kind of term is seldom used; instead, the designation “problematic drug use” is much more common.
In this study, the definition of serious substance abuse is based on previous Swedish studies (see Olsson et al., 2001) and is as follows: “individuals, regardless of age, who during the past 12 months have at some time injected narcotics, or who have used narcotics daily or almost daily (4 days or more per week) during the past 30 days”.

Materials and procedure
The data collection was conducted by a special project organisation and was completed in June 2011. A total of 1,924 questionnaires were turned in, and of these, 176 were incomplete or did not fulfil the criteria for problematic drug use. The material was supplemented with 400 reports from the two documentation systems, DOK¹ and ASI², that also met the criteria. Questionnaire-based reports made up 81 percent of all reports, while the remaining 19 percent came from the documentation systems.

The participating units were categorised into five sources: social welfare offices, outpatient care services, health care, correctional institutions and voluntary organisations.

A total of 2,148 reports were collected. The table above shows the distribution of the different categories.

The questionnaire, which was the primary mode of data collection for the study, included 20 questions. In addition to certain basic information on the reporting organisations and individuals (such as sex, age and city district), questions covered drug usage (primary drug and four other possible substances), care and treatment contacts as well as living conditions, such as housing and livelihood. An individual code (birth year, birthday and initials in the first name and surname such as 6513MD) was used primarily as a sorting variable and to control for double reporting and for determining individuals who occurred in one or more sources. In this manner, we avoided working with names, personal data or addresses.

The decisive inclusion criterion was whether a reported individual fulfilled the definition of problematic drug use. Other data used for including/excluding were study criteria for the time period, municipality affiliation as well as incomplete questionnaires/forms.

After a first review of the material, reports were excluded that did not fall under the definition of problematic drug use. Next, duplicate reports were identified with the aid of the individual codes. In some cases, identification and control of individual reports also took place with the help of information about sex, age, city district and the indicated primary drug of choice and other drug. Duplicate reports were removed so that only one individual per source occurred. This review was performed separately for both the questionnaire and the data from DOK and ASI. After the initial review of each individual report, the various materials were compiled.

| Informant                      | Number | Percentage |
|-------------------------------|--------|------------|
| Social welfare offices        | 632    | 29         |
| Outpatient care services      | 396    | 18         |
| Health care                   | 600    | 28         |
| Correctional institutions     | 307    | 14         |
| Voluntary organisations       | 213    | 10         |
| **Total**                     | 2,148  | 100        |

Table 2. Reports of individuals with serious substance abuse in Gothenburg, October 2010 – March 2011.
into a jointly formed data structure where duplicate reporting in relation to the five different sources was again controlled.

In addition to the material from the actual survey, we also collected local health care data (for Gothenburg) for analysis with the truncated Poisson method, including 1,096 individuals during 2010. The information covered all institutional care for any kind of drug-related diagnosis (according to ICD-10, including F, X and T codes), and the criteria otherwise followed the previously conducted national estimates (see FHI, 2010).

Validation of data
Quality regarding the information from the 1,748 included questionnaires was tested for reliability. A test was conducted on the questionnaire’s reliability for the 299 individuals who occurred twice or more times in the material. On the information about the primary drug of choice, the percentage agreement was 0.70 (where 1.00 is maximum agreement), which is considered acceptable (Bordens & Abbot, 2005). When the reported drug substance occurred in both of the compared questionnaires, agreement equalled 0.98, which is very high. The third variable tested – injection of drugs during the last 12 months – showed a percentage agreement of 0.84, which is also considered high. The interview instruments DOK and ASI have in previous reliability and validity tests been judged as having an acceptable quality (Anderberg & Dahlberg, 2009; McLellan, Cacciola, Alterman, Rikoon, & Carise, 2006).

Statistical analysis
Two different analyses were conducted using the capture-recapture method and three different analyses of the material from the health care registers used the truncated Poisson method. The first analysis with the capture-recapture method used all five sources, and in the second, the material was categorised in three sources: social welfare (including social welfare offices, outpatient care services and voluntary organisations), health care and correctional institutions. The pooled material from the study was processed and analysed in PASW Statistics 18, using the capture-recapture method to estimate the undetected population. The analyses were conducted in compliance with conventional procedures (for example, EMCDDA, 1999b; 2002; Uhl & Seidler, 2001). Below are the formulas used in the truncated Poisson analyses:

1. \[ N = S + f_1 * f_1 / (2 * f_2) \]
   This formula has been used in Swedish national estimates (FHI, 2010).

2. \[ N = 2 * S + f_1 * f_1 / (2 * f_2) \]
   Estimate according to Chao (Chao, 1989).

3. \[ N = S + S / [1 - \exp(-2 * f_2 / f_1)] \]
   Estimate according to Zelterman (Zelterman, 1988).

\[ N = \] total number for estimate including those not observed.
\[ f_1 = \] number of people who received care once.
\[ f_2 = \] number of people who received care twice.
\[ S = \] total number of people who have at some time received care.
Results
A total of 1,708 individuals were identified from the 2,148 reports submitted. Table 3 shows the number of individuals reported in more than one source.

The table shows that 81 percent of the 1,708 unique reports occur in only one source and accordingly, the remaining 19 percent are found in two or more of the five sources.

Table 4 describes the 1,708 individuals in relation to some of the study’s basic variables.

Table 3. Numbers of individuals with serious substance abuse in Gothenburg, October 2010 – March 2011.

| Individual reporting | Number | Percentage |
|----------------------|--------|------------|
| One source           | 1,385  | 81         |
| Two sources          | 234    | 14         |
| Three sources        | 74     | 4          |
| Four sources         | 13     | 1          |
| Five sources         | 1      | –          |
| **Total**            | **1,708** | **100**    |

When compared with individuals reported from other sources, it is evident that individuals reported by the social welfare offices are often between 20-39 years old, have amphetamines as their primary drug and inject drugs more often. Their incomes are to a greater degree provided through income support or the Swedish Social Insurance Agency, and they more often than not live in some kind of institutional housing.

The reports from outpatient care services highlight both a younger group and an older group. The primary drug is largely cannabis or amphetamines, and income is primarily wage-based and comes from parents. Members of this group more commonly live with their parents, have temporary accommodation or are homeless.

Reports from health care include a larger number of women, and the individuals are primarily in the age group 30-39. Their primary drug is heroin and other opiates or alcohol and pharmaceuticals. Individuals in this group largely support themselves with wage-based income or benefits through the Swedish Social Insurance Agency. More than in other groups, the reports indicate that they often have stable accommodation.

Individuals primarily reported by the correctional institutions are often men between 20-29 years old who have amphetamines or cannabis as their primary drug. They largely receive their income from other earnings or criminal activities. Many individuals in this group have unstable accommodation or live with their parents or in temporary accommodation.

Individuals reported by voluntary organisations are often men who are 40 years old or more. They largely report amphetamines and alcohol as their primary drug and they inject drugs more often than do other groups. These individuals more often receive their income through income support or other earnings, or from criminal activities. Their accommodation status is mostly unstable.

In general, the table shows the heterogeneity of the material. The five sources represent different groups, and problematic drug use can be expressed in different manners.

We used local health care data for analysis by the truncated Poisson method. Figure 1 (see page 10) shows the distribution of the number of care/treatment occasions.
Table 4. Individuals with serious substance abuse in Gothenburg, October 2010 – March 2011, reported from different sources by gender, age, primary drug use, labour status and housing. Percent.

| Variables | Social welfare offices | Outpatient care services | Health care institutions | Voluntary organisations | Total |
|-----------|------------------------|--------------------------|--------------------------|-------------------------|-------|
| **Sex (N=1708)** | | | | | |
| Male | 76 | 77 | 71 | 86 | 80 | 76 |
| **Age (N=1708)** | | | | | |
| 19 years and younger | 3 | 11 | 4 | 5 | - | 5 |
| 20-29 years | 30 | 23 | 29 | 39 | 19 | 29 |
| 30-39 years | 27 | 18 | 28 | 25 | 24 | 25 |
| 40-49 years | 23 | 23 | 23 | 22 | 33 | 24 |
| 50 years and older | 18 | 25 | 16 | 10 | 24 | 18 |
| **Primary drug (N=1708)** | | | | | |
| Amphetamines | 40 | 46 | 26 | 39 | 46 | 37 |
| Heroin and other opiates | 19 | 11 | 27 | 18 | 14 | 19 |
| Cannabis¹ | 15 | 27 | 14 | 28 | 18 | 19 |
| Alcohol | 11 | 7 | 15 | 4 | 23 | 11 |
| Pharmaceuticals² | 10 | 5 | 15 | 4 | 4 | 9 |
| Other substances³ | 4 | 4 | 3 | 7 | 1 | 4 |
| Injection of primary drug of choice (N=1493) | 33 | 31 | 24 | 29 | 44 | 31 |
| **Labour status during the past 6 months (N=1568)** | | | | | |
| Wage-based income⁴ | 10 | 23 | 24 | 15 | 11 | 18 |
| Benefits from the Swedish Social Insurance Agency⁵ | 21 | 12 | 28 | 12 | 8 | 19 |
| Income support | 63 | 51 | 42 | 52 | 72 | 53 |
| Supported by parents | 4 | 10 | 4 | 4 | 2 | 5 |
| Other support⁶ | 3 | 5 | 4 | 17 | 7 | 5 |
| **Living status during the past 6 months (N=1619)** | | | | | |
| Stable accommodation | 23 | 14 | 44 | 22 | 32 | 28 |
| With parents | 10 | 20 | 15 | 16 | - | 14 |
| Temporary accommodation | 11 | 17 | 9 | 13 | 3 | 11 |
| Institutions⁷ | 42 | 26 | 17 | 18 | 32 | 27 |
| Unstable accommodation⁸ | 13 | 23 | 15 | 32 | 34 | 20 |

**Notes**
1 Cannabis includes hashish and marijuana.
2 Pharmaceuticals include sedative and pain-killing medications.
3 Other substances include GHB/GBL, other substances, cocaine, anabolic androgenic steroids as well as LSD and other hallucinogens.
4 Own job, unemployment benefits, medical benefits, retirement pension, study allowance, study grant, parental allowance.
5 Sickness benefits, activity allowance.
6 Criminal activities, prostitution, other earnings.
7 Trail or training apartment, institution, family care, prison, jail.
8 Hotel, with relatives, homeless, other kind of housing.
for the 1,096 people who on one or more occasions were placed in inpatient care within the health care services in Gothenburg.

Thus, 61 percent have been treated in inpatient care on one occasion, and 19 percent on two occasions. During 2010, a total of 2,078 treatment occasions were associated with some kind of diagnosis that involved use of drugs.

A few different accepted analytical models for estimating the unobserved population and prevalence of problematic drug use in Gothenburg have been tested with varied results. The study included 1,708 identified individuals reported by the participating units. Results from the hidden population analyses with capture-recapture and truncated Poisson are reported in table 5.

As is evident from this table, the statistical models used provide an inconsistent estimation of the hidden population, ranging between 1,100 and 2,700 people. When we add the actual number of individuals reported to these numbers, the total prevalence of problematic drug use in Gothenburg is estimated, according to the analyses, at between 2,200 and 4,400 individuals.

Discussion

With this study as a starting point, it is possible to ascertain that estimation of the presence and prevalence of problematic drug use involves significant methodological difficulties and challenges. As a result, cross-validation by using a combination of multiple data sources and methods of estimation is recommended for studies of prevalence of problematic drug use (Bloor, 2005; Choi & Comiskey, 2003; EMCDDA, 1999a; 1999b; 2002; Hartnoll, 1997; Hickman et al., 2002; 2006; Kimber et al., 2008; Kraus et al., 2003; Law et al., 2006):

The use of multiple data sources and different approaches to analysis, together with an honest assessment of the uncertainties involved, provide a number of estimates, each with vary-

Table 5. Estimates of problematic drug users in Gothenburg, October 2010 – March 2011, by five methods.

| Method                          | Registered users | Estimate of non-registered users | Estimate of problematic drug users |
|---------------------------------|------------------|----------------------------------|-----------------------------------|
| Truncated Poisson*              | 1,096            | 1,100                            | 2,200                             |
| Truncated Poisson (Chao)        | 1,096            | 2,200                            | 3,300                             |
| Truncated Poisson (Zelterman)   | 1,096            | 2,400                            | 3,500                             |
| Capture-recapture (five sources)| 1,708            | 2,000                            | 3,700                             |
| Capture-recapture (three sources)| 1,708          | 2,700                            | 4,400                             |

* Following the formula used in previous Swedish estimates (Brå 2003 and FHI 2010).
The extent to which differing estimates agree lends robustness to the results (Law et al., 2006, p. 156).

In the empirical study that serves as the basis for this article, we have therefore applied and tested several different approaches for estimating the number of individuals with problematic drug use in Gothenburg.

The case-finding study identified 1,708 individuals among 2,148 received reports. Seventy-two different public authorities and organisations turned in information on people they had encountered with a problematic drug use. Information is not available from the units that did not participate. For example, information is lacking from the police authority and some divisions of health care services.

As this study shows, the analyses indicate highly variable estimates of the unobserved population and prevalence of problematic drug use in Gothenburg. A number of different aspects may contribute to the explanation of this large variation. One such feature concerns the basic assumptions applicable in the different methods of analysis.

All of the five basic assumptions that apply for capture-recapture studies are to a greater or lesser extent difficult to fulfil. With these assumptions as a starting point, we will discuss the extent to which the assumptions have been followed or conformed to.

Correct classification means that the individuals are covered by the definition of the population and that the same individuals can be identified in the different information sources through an operationalisation and construction of individual codes. It is important that the definition of the target group is as clear and exact as possible in order to make it simple for the informants to determine who should be included and studied. Here, the knowledge on the specific individual may differ between the different sources; one organisation may have extensive information about an individual, while another organisation may have more uncertain or fragmentary information to report (Olsson, 1997). It is therefore extremely important that the definition is very detailed and operationalised in measurable terms. In order to minimise erroneous classification, we have used a simple but accepted definition (see Bloor, 2005).

The second assumption – that the individuals reported by the informants are representative for the population – is also difficult to live up to. It is nevertheless important that there are several sources of information, that all informants come in contact with the relevant target group of people with problematic drug use, and that the study strives to include sources that represent both treatment and the criminal justice system (Hickman et al., 1999; 2002). In this case, several organisations and public authority divisions have contributed with information. These informants include both treatment systems and correctional treatment. However, since the police authority was not able to answer the questionnaire, the judicial system is underrepresented.

That the population is assumed to be closed would entail that no individuals are lost or added during the study’s observation period, and that this is equally likely to be reported by all sources. Prob-
lematic drug use can, however, escalate or stop, people can move to or from the municipality or they may even die due to problematic drug use. One way of dealing with such unavoidable mobility is to limit the study period based on the characteristics of the specific problem, usually to 6 or 12 months when dealing with problematic drug use (Bloor, 2005). In this case, and in agreement with previous Swedish studies, we limited the sampling period to 6 months.

The fourth assumption regarding homogeneity is significantly more complicated to fulfil, because it is well-established that problematic drug use can be expressed in many different ways. These include substance and consumption patterns, evident in this study. As can be seen in the results (see Table 4), the five sources partially represent different groups and illustrate that as a comprehensive definition, problematic drug use shows much individual variation.

Regarding the fifth assumption in studies using the capture-recapture method in the area of problematic drug use, it is common that a relationship exists between certain sources. For example, a person reported by the homecare and outpatient care services can also be found in the information reported by the social welfare service. In this case, the correlation seems to be weak or non-existent, as it shows that overlaps between the different sources are relatively few (see Table 3).

As the results of the capture-recapture analyses make clear, the estimates of the hidden population are highly variable and unreasonably large. In addition to the difficulties in fulfilling the basic assumptions of the method of analysis, one other aspect is significant. The major reason for the overestimation is that the basic data contains several special characteristics that affect the results.

As already mentioned, capture-recapture studies are based on the assumption that individuals in the defined population have an equal chance of being reported by any of the reporting organisations. This assumes that there is no dependence or correlation between the various sources, either positive or negative. If this requirement is not fulfilled, a situation of “variable catchability” occurs, which affects the analysis and the estimation of the total population (Brenner, 1996; Chao et al., 2001; Cormack, 1999a; 1999b; Corrao et al., 2000; Domingo-Salvany et al., 1998; Hook & Regal, 1993; 1995; 1999; 2000b; Korf, 1997; Neugebauer & Wittes, 1994; Papoz et al., 1996; Tilling, 2001). If the overlapping is limited and there is a weak correlation between two or more sources, this can lead to an overestimation of the population (see Korf, 1997). Such is the case here: the individuals do not have the same opportunity of being reported by all of the sources.

One explanation for the disproportionate distribution found in this study between those who were reported by one or more sources is likely the heterogeneity in the material, i.e. that different subgroups clearly occur with differences based on sex, age, socioeconomic factors, type of abuse or how serious the abuse is (see Table 4).

It should be noticed that source dependence and variable catchability are, in fact, intertwined concepts because the heterogeneous probability
of capture might be the underlying explanation for an observed dependence among sources (Corrao et al., 2000, p. 605).

The distinctive characteristics of the different sources are probably also significant for the kind of people an organisation or public authority concentrates on. In turn, this gives larger or smaller opportunities for being “captured” by a specific source, which in turn may influence the interaction between sources and thus the estimation of the total population (Tilling, 2001).

Another possible explanation for the unreasonably high estimates of the total population is that the reports from the various sources together and to a great extent approach the actual total number of individuals who have serious problems with abuse. The hidden population is therefore limited (Hook & Regal, 1995). Thanks to carefully conducted data collection, this study has likely managed to “capture” a significant portion of the population, and dropout cases seem to be assessable. The capture-recapture method is based on there being an unknown number of individuals, but it seems to be too sensitive for use if the unknown number of individuals is relatively small already at the start (Brenner, 1996; Cormack, 1999a).

There are also limitations on the statistical assumptions connected to analyses using truncated Poisson. The assumption that the population is closed is difficult to assess, as the data was obtained from the health care services; however, considering the nature of the problem, certain mobility over time most likely also occurs here.

Also, while the information was taken from a specific register, this group cannot be considered homogeneous, either. In addition, problems with abuse can take on different forms, and patients have different kinds of drug-related diagnoses based on both substance and consumption patterns. For example, some patients have problems with drugs as their main diagnosis, while others have this as a sub-diagnosis alongside other more salient concerns.

Just as problematic is the assumption that the likelihood of being included in the material is constant and that the number of individuals who have been in care once is in proportion to patients who have been in care twice, three times or not at all. As shown above, the distribution is variable in the health care data used (see Figure 1). The probability of individuals seeking health care for their problems with substance abuse can be affected by age and social situation, for example, not to mention the availability and selection of health care services (Law et al., 2006).

In addition, there are many doubts regarding whether patients from a health care register can be said to represent the total population. Based on the heterogeneity of the case-finding material and the different kinds of problematic drug use in the information reported from the different sources, it is obviously risky when only one data source is considered to be representative for the entire group (Choi & Comiskey, 2003; Hickman et al., 2002; 2006; Hickman & Taylor, 2005; Law et al., 2006; Uhl & Seidler, 2001).

Previous research within the field indicates several problematic relationships for the various methods used in this study as well as some uncertainty about estimations of unidentified populations in general. The significant variation on hidden
populations presented in this study does questions the reliability and validity of the methods. The sensitivity of the methods regarding the importance of fulfilling the various basic assumptions is apparent, which a number of other studies also confirm (Brenner, 1996; Choi & Comiskey, 2003; Cormack, 1999a; 1999b; Corrao et al., 2000; Domingo-Salvany et al., 1998; Hickman et al., 2002; 2006; Hickman & Taylor, 2005; Korf, 1997; Law et al., 2006; Neugebauer & Wittes, 1994; Papoz et al., 1996; Tilling, 2001; Uhl & Seidler, 2001).

Our conclusions also lead to some uncertainty about how earlier and more relevant estimates of the prevalence of problematic drug use on the national and local level should be perceived and interpreted. Despite the limitations of the methods, great importance has been attached to results from such studies, and in Sweden, they have formed the basis for political proposals, deliberations and decisions (SOU, 2005; SOU, 2011).

Implications for future studies
We argue that estimates of hidden populations within the area of substance abuse based on the capture-recapture method or the truncated Poisson method should be viewed with both caution and scepticism. What should be conducted instead are thorough local studies in order to be able to estimate the prevalence of serious or problematic drug use as close to the actual number as possible. One such way of proceeding is also in line with the prevalence studies which are regularly conducted in health care research and which are based on objective information in registers or specific studies.

The heterogeneity described for the target group of people with problematic drug use also argues in favour of conducting a similar study in the future using case-finding studies, including informants from several different organisations and public authorities. It would probably be difficult to find one single source that would be representative of this target group. In conclusion, we recommend a total assessment based on cross-validation of different material and factual information.

Declaration of Interest None.

Mikael Dahlberg, Researcher
Linnaeus University
E-mail: mikael.dahlberg@lnu.se

Mats Anderberg, Researcher
Linnaeus University
E-mail: mats.anderberg@lnu.se

NOTES
1 DOK stands for Documentation of clients.
2 Addiction Severity Index.
REFERENCES

Anderberg, M., & Dahlberg, M. (2009). Strukturerade intervjuer inom missbruksvården – som en grund för kunskapsutveckling. [Structured interviews in substance abuse treatment – as a foundation for the development of knowledge.] Växjö: Växjö University Press.

Bloor, M. (2005). Population estimation without censuses or surveys: A discussion of mark-recapture methods illustrated by results from three studies. Sociology, 39(1), 121–138.

Brenner, H. (1996). Effects of misdiagnoses on disease monitoring with capture-recapture methods. Journal of Clinical Epidemiology, 49(11), 1303–1307.

Bordens, K.S., & Abbott, B.B. (2005). Research, design and methods: A process approach. 6th edition. New York: McGraw Hill Higher Education.

Brå (2003). Polisens insatser mot narkotikabrotsligheten – Omfattning, karakter och effekter. [Police efforts to combat drug offending: Their extent, character and effects]. Rapport 2003:12. Stockholm: BRÅ (Brottsförebyggande rådet).

CAN (2010). Droguutvecklingen i Sverige 2010. [Drug trends in Sweden 2010]. CAN Rapport 125. Stockholm: CAN (Centralförbundet för alkohol- och narkotikaupplysning).

Chao, A. (1989). Estimating population size for sparse data in capture-recapture experiments. Biometrics, 45(2), 427–438.

Chao, A., Tsay, P.K., Lin, S-H., Shau, W-Y., & Chao, D-Y. (2001). The application of capture-recapture models to epidemiological data. Statistics in Medicine, 20(3), 3123–3157.

Choi, Y.H., & Comiskey, C.M. (2003). Methods for providing the first prevalence estimates of opiate use in Western Australia. International Journal of Drug Policy, 14(4), 297–305.

Cormack, R.M. (1999a). Problems with using capture-recapture in epidemiology: An example of a measles epidemic. Journal of Clinical Epidemiology, 52(10), 909–914.

Cormack, R.M. (1999b). Reply to preceding comments. Journal of Clinical Epidemiology, 52(10), 929–933.

Corrao, G., Bagnardi, V., Vittadini, G., & Favilli, S. (2000). Capture-recapture methods to size alcohol related problems in a population. Journal of Epidemiology and Community Health, 54(8), 603–610.

Domingo-Salvany, A., Hartnoll, R.L., Maguire, A., Brugal, T., Albertin, P., Caylà, J.A., Casabona, J., & Suelves, J.M. (1998). Analytical considerations in the use of capture-recapture to estimate prevalence: Studies of the estimation of opiate use in the metropolitan area of Barcelona, Spain. American Journal of Epidemiology, 148(8), 732–740.

EMCDDA (1999a). Scientific review of the literature on estimating the prevalence of drug misuse on the local level. Lisbon: EMCDDA (European Monitoring Centre for Drugs and Drug Addiction).

EMCDDA (1999b). Methodological guidelines to estimate the prevalence of problem drug use on the local level. Lisbon: EMCDDA.

EMCDDA (2002). European network to develop policy relevant models and socio-economic analyses of drug use, consequences and interventions. Lisbon: EMCDDA.

EMCDDA (2006). Annual report: The state of the drugs problem in Europe. Lisbon: EMCDDA.

FHI (2010). Skattning av problematisk narkotikaanvändning 2007. In Narkotikabrukiet i Sverige (p. 305). [Drug use in Sweden]. Östersund: Statens folkhälsoinstitut.

Frischer, M., Hickman, M., Kraus, L., Mariani, F., & Wiessing, L. (2001). A comparison of different methods for estimating the prevalence of problematic drug misuse in Great Britain. Addiction, 96(10), 1465–1476.

Gemmel, I., Millar, T., & Hay, G. (2004). Capture-recapture estimates of problem drug use and the use of simulation based confidence intervals in a stratified analysis. Journal of Epidemiology and Community Health, 58(9), 758–765.
Hartnoll, R. (1997). General conclusions. In G.V. Stimson, M. Hickman, A. Quirk, M. Frischer & C. Taylor (Eds.), *Estimating the prevalence of problem drug use in Europe* (pp. 263–264). Lisbon: EMCDDA.

Hay, G. (1997). The selection from multiple data sources in epidemiological capture-recapture studies. *The Statistician, 46*(4), 515–520.

Hay, G. (2003). *Truncated Poisson analyses of data from the Riga Needle Exchange*. Washington: The Synergy Project.

Hay, G., & Gannon, M. (2006). Capture-recapture estimates of the local and national prevalence of problem drug use in Scotland. *International Journal of Drug Policy, 17*(3), 203–210.

Hickman, M., & Taylor, C. (2005). Indirect methods to estimate prevalence. In Z. Sloboda (Ed.), *Epidemiology of drug abuse* (pp. ). New York: Springer.

Hickman, M., Cox, S., Harvey, J., Howes, S., Farrell, M., Frischer, M., Stimson, G., Taylor, C., & Tilling, K. (1999). Estimating the prevalence of problem drug use in inner London: A discussion of three capture-recapture studies. *Addiction, 94*(11), 1653–1662.

Hickman, M., Taylor, C., Chatterjee, A., Degenhardt, L., Frischer, M., Hay, G., Tilling, K., Wiessing, L., Griffiths, P., & McKetin, R. (2002). Estimating the prevalence of problematic drug use: A review of methods and their application. *Bulletin on Narcotics, 54*(1–2), 15–32.

Hickman, M., Hope, V., Platt, L., Higgins, V., Bellis, M., Rhodes, T., Taylor, C., & Tilling, K. (2006). Estimating prevalence of injecting drug use: A comparison of multiplier and capture-recapture methods in cities in England and Russia. *Drug and Alcohol Review, 25*(2), 131–140.

Hook, E.B., & Regal, R.R. (1993). Effect of variation in probability of ascertainment by sources (“variable catchability”) upon capture-recapture estimates of prevalence. *American Journal of Epidemiology, 137*(10), 1148–1166.

Hook, E.B., & Regal, R.R. (1995). Capture-recapture methods in epidemiology: Methods and limitations. *Epidemiological Reviews, 17*(2), 243–264.

Hook, E.B., & Regal, R.R. (1999). On the need for a 16th and 17th recommendation for capture-recapture analysis. *Journal of Clinical Epidemiology, 53*(12), 1275–1277.

Hook, E.B., & Regal, R.R. (2000a). Recommendations for presentations and evaluations of capture-recapture estimates in epidemiology. *Journal of Clinical Epidemiology, 52*(10), 917–826.

Hook, E.B., & Regal, R.R. (2000b). Accuracy of alternative approaches to capture-recapture estimates of disease frequency: Internal validity analysis of data from five sources. *American Journal of Epidemiology, 152*(8), 771–778.

Kimber, J., Hickman, M., Degenhardt, L., Coulson, T., & van Beek, I. (2008). Estimating the size and dynamics of an injecting drug user population and implications for health service coverage: Comparison of indirect prevalence estimation methods. *Addiction, 103*(10), 1604–1613.

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

Korf, D.J. (1997). Comparison of different estimation methods in the Netherlands. In G.V. Stimson, M. Hickman, A. Quirk, M. Frischer & C. Taylor (Eds.), *Estimating the prevalence of problem drug use in Europe* (pp.). Lisbon: EMCDDA (European Monitoring Centre for Drugs and Drug Addiction).

Law, M.G., Degenhardt, L., & McKetin, R. (2006). Methods estimating the prevalence of problem drug use. *Journal of Drug Policy, 17*(3), 154–158.

McLellan, A.T., Cacciola, J.S. & Alterman, A.I., Rikoon, S.H., & Carise, D. (2006).
The Addiction Severity Index at 25: Origins, contributions and transitions. *The American Journal on Addictions, 15*(2), 113–124.

Maxwell, J.C., & Pullum, T.W. (2001). Using a modification of the capture-recapture model to estimate the need for substance abuse treatment. *Evaluation and Program Planning, 24*(3), 257–265.

Millar, T., Gemmell, I., Hay, G., & Donmall, M. (2004). *The dynamics of drug misuse: Assessing changes in prevalence*. Glasgow: The Centre for Drug Misuse Research, University of Glasgow.

Millar, T., Domingo-Salvany, A., Eastwood, C., & Hay, G. (2008). *Glossary of terms relating to capture-recapture methods*. *Journal of Epidemiology & Community Health, 62*(8), 677–681.

Narkomanvårdskommittén (1969). *Narkotikaproblemet. Samordnade åtgärder. Del III. [The drug problem. Co-ordinated actions. Part III]*. SOU 1969:52. Stockholm: Esselte.

Neugebauer, R., & Wittes, J. (1994). Annotation: Voluntary and involuntary capture-recapture samples – Problems in the estimation of hidden and elusive populations. *American Journal of Public Health, 84*(7), 1068–1069.

Olsson, B. (1997). Problems of definition and other aspects of case-finding: The Swedish nationwide studies. In G.V. Stimson, M. Hickman, A. Quirk, M. Frischer & C. Taylor (Eds.), *Estimating the prevalence of problem drug use in Europe*. Lisbon: EMCDDA (European Monitoring Centre for Drugs and Drug Addiction).

Olsson, B. (1999). Metoder att uppskatta antalet narkotikamissbrukare. In *Narkotikastatistik – om samhällets behov av information om narkotikastatistik* (pp.). [Drug statistics: On society’s need for information about narcotics]. SOU 1999:90. Stockholm: Socialdepartementet.

Olsson, B., Carlsson, G., Fant, M., Olsson, O., & Roth, C. (1981). Heavy drug abuse in Sweden 1979 – A national case-finding study. *Drug and Alcohol Dependence, 7*(3), 273–283.

Olsson, B., Adamson Wahren, C., & Byqvist, S. (2001). *Det tunga narkotikamissbrukets omfattning i Sverige 1998. MAX-projektet, delrapport 3. [The extent of heavy drug abuse in Sweden 1998]*. Stockholm: CAN (Centralförbundet för alkohol- och narkotikaupplysning).

Olsson, O., Byqvist, S., & Gomér, G. (1994). The prevalence of heavy narcotics abuse in Sweden in 1992. *Scandinavian Journal of Social Welfare, 3*(2), 81–84.

Papoz, L., Balkau, B., & Lellouch, J. (1996). Case counting in epidemiology: Limitations of methods based on multiple data sources. *International Journal of Epidemiology, 25*(3), 474–478.

Simon, R. (1997). Estimating prevalence using the case-finding method: An overview. In G.V. Stimson, M. Hickman, A. Quirk, M. Frischer & C. Taylor (Eds.), *Estimating the prevalence of problem drug use in Europe*. Lisbon: EMCDDA (European Monitoring Centre for Drugs and Drug Addiction).

Socialstyrelsen (2012). *Tillståndet och utvecklingen inom hälso- och sjukvården och socialtjänsten. Lägesrapport 2012. [The status and development in health care and social services]*. Stockholm: Socialstyrelsen.

SOU (2005). *Personer med tungt missbruk. Stimulus till bättre vård och behandling. [People with heavy drug abuse. Incentives for better care and treatment]*. Stockholm: Fritzes offentliga publikationer.

SOU (2011). *Missbruket, Kunskapen, Vården. Missbruksutredningens forskningsbilaga. [Drug abuse, knowledge, care]. Governmental report, research appendix]*. Stockholm: Fritzes offentliga publikationer.

Svensson, D. (2005). *Tungt missbruk av alkohol och narkotika. [Problematic use of alcohol and drugs]*. Unpublished paper. Göteborg: Göteborgs stad.

Tilling, K. (2001). Capture-recapture methods – Useful or misleading? *International Journal of Epidemiology, 30*(1), 12–14.

Uhl, A., & Seidler, D. (2001). *Prevalence estimate of problematic opiate consumption in Austria, second revised edition*. Vienna: Ludwig Boltzmann Institute for Addiction Research.

Vaissade, L., & Legleye, S. (2008). Capture-
recapture estimates of the local prevalence of problem drug use in six French cities. 
*European Journal of Public Health, 19*(1), 32–37.

Zelterman, D. (1988). Robust estimation in truncated discrete distributions with application to capture-recapture experiments. *Journal of Statistical Planning and Inference, 18*(2), 225–237.