Opiate and crack cocaine use: A new understanding of prevalence

GORDON HAY\textsuperscript{1}, MARIA GANNON\textsuperscript{1}, JANE MACDOUGALL\textsuperscript{1}, CATHERINE EASTWOOD\textsuperscript{2}, KATE WILLIAMS\textsuperscript{2}, & TIM MILLAR\textsuperscript{2}

\textsuperscript{1}University of Glasgow, Centre for Drug Misuse Research, Glasgow, UK and \textsuperscript{2}University of Manchester, National Drug Evidence Centre, Manchester Science Park, Manchester, UK

Abstract

Aims: The aim of this study was to establish prevalence estimates of problem drug use, defined as opiate and/or crack cocaine use by persons aged 15 to 64 years, for England and for each of the 149 administrative areas responsible for commissioning drug interventions.

Methods: Indirect estimation techniques, the capture–recapture and multiple indicator methods, were used to obtain estimates. Information on problem drug users presenting to healthcare settings and/or recorded by the criminal justice system, and drug-related indicator data were used in the analyses.

Findings: There were an estimated 332,090 problem drug users in England during 2005/06 (95\% CI 324,546 to 346,345), equivalent to 9.97 (95\% CI 9.74 to 10.40) problem drug users per thousand population aged 15 to 64 years. Prevalence varied by geographic region and age group: the highest rates were observed in London and for those aged 25 to 34 years.

Conclusions: This study has produced estimates of the prevalence of problem drug use in England that are more robust, more precise, and suggest a higher prevalence than previous studies. The estimates provide a basis on which to formulate policy, plan services, and measure service performance.

Introduction

Problem drug use has an impact on health, crime, and the wellbeing of communities. The social and economic costs of drug misuse in England and Wales have been estimated to be between £10.1 billion and £17.4 billion.
per year, while the costs to the health service, excluding specific addiction
treatment, have been estimated to be between £283 and £509 million per year
(Godfrey, Eaton, McDougall, & Culyer, 2002). Precise, accurate, and
geo-graphically and temporally consistent estimates of the prevalence of problem
drug use have been lacking. However, these should be a key element of the
evidence base that is used to formulate policy, to direct resources effectively, and
to understand the impact of interventions designed to address the problem.

Problem drug use is a stigmatized, covert and often illegal activity. Hence direct
enumeration of the size of this population is very difficult. Estimates based on
population surveys are often unfeasibly small. For example, results from the
2005/06 British Crime Survey suggested that 39,000 people in England and
Wales had used heroin during the previous year (Roe & Man, 2006), but
this figure is smaller than the number of heroin users reported to be in contact
with specialist drug treatment services (108,000), in England alone, during
2005/06 (National Treatment Agency for Substance Misuse, 2006).

Indirect estimation techniques produce more credible results. Capture–
recapture, in particular, has been applied to generate estimates at the city or
sub-regional level (Bello & Chene, 1997; Beynon et al., 2001; Brugha, Swan,
Hayhurst, & Fallon, 1998; Bunter, van Brussel, & van den Brink, 2001; Calkins &
Aktan, 2000; Choi & Comiskey, 2003; Comiskey & Barry, 2001; Davies,
Cormack, & Richardson, 1999; Domingo-Salvany et al., 1998; Frischer et al.,
1993; Hartnoll, Mitcheson, Lewis, & Bryer, 1985; Hay, 2000; Hay &
McKeganey, 1996; Hickman et al., 1999, 2004; Holland et al., 2006; Hope,
Hickman, & Tilling, 2005; Hser, 1993; Larson, Stevens, & Wardlaw, 1994;
Mastro et al., 1994; Platt et al., 2004; Squires, Beeching, Schlecht, & Ruben,
1995), but seldom at the national level (Hay, McKeganey, & Hutchinson, 2001;
McElrath, 2002; Wood, Bloor, & Palmer, 2000). In the United Kingdom, pilot
work has applied a technique known as the multiple indicator method (Wickens,
1993) to extrapolate from sub-national estimates so as to produce estimates for
Great Britain (Frischer, Hickman, Kraus, Mariani, & Wiessing, 2001) and
England (Frischer, Heatlie, & Hickman, 2006). Although these studies
contributed greatly towards methodological developments in the field, they
were hampered by the small number of sub-national estimates available, which
were not geographically representative and which used differing definitions of
problem drug use, and, in the case of the latter study, a lack of verified published
‘indicator’ data. The resulting problem drug use estimate for England was
imprecise: the authors presented a 90% confidence interval for the estimate that
ranged from 174,117 to 401,224 (Frischer et al., 2006).

The national prevalence study (Hay et al., 2006) reported here is the first
systematic application of indirect estimation methods to generate a robust and
precise estimate of the prevalence of opiate and/or crack cocaine use in England.
The study has been designed to produce serial estimates over a three-year period.
The focus on opiate and/or crack cocaine use was, in part, at the request of the
commissioning body (the UK Home Office) as those were the drugs that were
considered to cause the most harm to society in general. The remit was also
restricted to those drugs as there may have been difficulty using the specific methods employed by this study to examine a wider case definition. There are, of course, other patterns of drug use that may be problematic to the individual and society, such as the injecting of amphetamines or the use of powder cocaine, however this study only examined opiate and/or crack cocaine use.

The aim of the study was therefore to provide estimates of the prevalence of opiate and/or crack cocaine use at the national and regional for England for the financial year 2005/06.

Method

The national prevalence study of problem drug use used two indirect estimation techniques—multi-sample capture–recapture and the multiple indicator method—to derive estimates for each of the 149 Drug Action Team (DAT) areas in England. These techniques are described in detail elsewhere (European Monitoring Centre for Drugs & Drug Addiction, 1997; Wickens, 1993). Briefly, multi-sample capture–recapture considers the overlap between lists of known individuals, drawn from the population of interest, in order to determine the intensity with which the lists sample from that population. Log linear regression is then used to model the observed pattern of overlap and the ‘best’ fitting model is used to generate an estimate of the number of individuals not included in any list, i.e. the ‘hidden’ part of the population. The multiple indicator method takes prevalence estimates for a set of ‘anchor point’ areas and uses regression analysis to model the relationship between these and a set of indicators of problem drug use. The resulting model is then used to extrapolate to those remaining areas where the indicators are available but prevalence estimates are not.

In a given population, a proportion of drug users are known as a consequence of their contact with health and/or criminal justice services. Four data sources of known problem drug users were used in the capture–recapture analyses for 2005/06: the National Drug Treatment Monitoring System; the National Probation Service Offender Assessment System; drug users convicted under the Misuse of Drugs Act (1971) for offences involving possession (or possession with intent to supply) heroin, methadone and/or crack cocaine recorded in the Police National Computer; and Drug Interventions Programme assessments completed in prisons. Records were retained if the individual had contact with health and/or criminal justice services between 1 April 2005 and 31 March 2006, reported use of opiates and/or crack cocaine, was aged 15 to 64 years, and was known to reside in England. The final four samples included one record for each problem drug user for each DAT area in which they reported to be resident during the year. The overlap between samples was determined by matching forename and surname initials, date of birth and gender within each area.

The 22 simplest models (an independence model with no interactions between samples, six models containing one interaction between single pairs of samples, and 15 models containing interactions between sets of two different pairs
of samples) were tested using log linear regression to determine which best fitted the pattern of overlap. The decision to restrict the consideration of capture–recapture models to the simplest 22 was pragmatic. It would have been less efficient (computationally) to consider all 114 models that can be fitted to a four-source capture–recapture analysis, particularly given the relatively large number of capture–recapture analyses used to obtain the national and regional estimates. In addition, more complex capture–recapture models tend to produce estimates with wider confidence intervals. As there was no reason to believe that the estimates derived using the multiple indicator method were in any way biased, it was felt more appropriate to use a multiple indicator estimate instead of an estimate derived from a complex capture–recapture model.

The analysis (within each area) was carried out for each individual age group strata (15 to 24, 25 to 34 and 35 to 64 years of age) and also separately for males and females. In addition, stratified estimates by age group and gender (i.e. males 15 to 24 years of age) were also derived.

A model was considered to be a good fit when the deviance and Akaike Information Criterion (AIC) values were low, and when the associated estimate was similar to the weighted estimate, calculated as a weighted mean of the available estimates described elsewhere (Hook & Regal, 1997).

The resultant estimate for each area was derived as either the unstratified estimate for that area or a combination of the stratified estimates (on considering the goodness of fit for the various stratified estimates and the desired to produce simpler estimates, such as those derived from analyses involving fewer stratifications).

The multiple indicator method was used to extrapolate estimates for the 39 areas for which suitable capture–recapture estimates were not used. The decision not to use a capture–recapture estimate (and thus opt for an estimate derived from the multiple indicator method) was based on how well the capture–recapture models that were considered fitted the overlap pattern (a poor fit meant that the capture–recapture estimate was not used) and the credibility of the estimate such that if it was more that 10 times the known population then it was not thought credible. The impact of restricting the analyses to the simplest 22 models was examined within a sensitivity analysis, which compared the best estimates (from the 22 simpler models) with the best estimates from all 114 possible models within a random sample of areas. The summed estimates for all areas included in the sensitivity were not significantly different.

The 110 available capture–recapture estimates were used as ‘anchor points’, while stepwise regression was used to select indicators for the best regression model, which overall explained 90% of the variance. This was measured by the adjusted $R^2$ value from the regression analysis. A wide range of possible indicator data were considered, including the aggregate data from the capture–recapture analyses, hospital admissions, published data on drug-related deaths, published data on drug-related and other crime, social indicators such as uptake of benefits and the population density of the area. The indicators chosen, in order of significance, were the number of opiate and/or crack cocaine users recorded by
the National Drug Treatment Monitoring System, population density, the number of opiate and/or crack cocaine users recorded by the Drug Interventions Programme in prisons, burglaries in a dwelling, drug-related deaths, and drug offences for possession of controlled drugs.

Gender and age group specific prevalence estimates were obtained by applying the proportion of problem drug users in each gender and age group strata, derived from weighted capture–recapture estimates, to the total prevalence estimates. The resulting 149 prevalence estimates for DAT areas were then summed to obtain estimates for government office regions and a national estimate. Simulation techniques, as described elsewhere (Gemmell, Millar, & Hay, 2004; Millar, Gemmell, Hay, & Donmall, 2003), were used to obtain 95% confidence intervals.

Results

Table I shows the number of anchor point areas for which capture–recapture estimates were available for use in the multiple indicator method for each government office region. Overall, capture–recapture methods produced prevalence estimates for 110 DAT areas, and the multiple indicator method for the remaining 39 areas.

There were an estimated 332,090 (95% CI 324,546 to 346,345) problem drug users, aged 15 to 64 years, in England during 2005/06. This is equivalent to 9.97 (95% CI 9.74 to 10.40) problem drug users per thousand population aged 15 to 64 years. Estimated population rates varied from 5.32 per thousand in the East of England (95% CI 4.11 to 6.58) to 14.99 in London (95% CI 14.48 to 15.90). Prevalence estimates and rates for England and each government office region are shown in Table I.

Males accounted for over three-quarters (77%) of problem drug users in England, with very little regional variation (from 73% in the South West to 79% in London). In contrast, the age distribution of problem drug users was seen to vary somewhat by region (Figure 1). Nationally, approximately 43% of problem drug users were aged 25 to 34 years, 37% 35 to 64 years, and 20% 15 to 24 years. Both London and the North West had a relatively older population compared to other regions (48% and 42% aged 35 to 64 years respectively), while the East Midlands (27%), Yorkshire and the Humber (26%), and the North East (26%) had the largest proportion of problem drug users aged 15 to 24 years.

Table II presents the estimated prevalence rates by gender and age group. The estimated national prevalence rate for males aged 15 to 64 years (15.32: 95% CI 14.86 to 15.87) per thousand population was much greater than that for females (4.64: 95% CI 4.61 to 4.99). In London, 23.56 (95% CI 22.55 to 24.70) per thousand of the male population aged 15 to 64 years were estimated to be problem drug users. Stratification by age showed the national prevalence rate for the 25 to 34 year age group (21.43: 95% CI 20.76 to 22.24) to be just over twice that of the youngest age group (10.07: 95% CI 9.85 to 10.66). All regions showed a similar pattern for age distribution, whereby the 25 to 34 year age group showed
Table I. Problem drug use estimates and rates per thousand population aged 15 to 64 years, with associated 95% confidence intervals, by government region, 2005/06.

<table>
<thead>
<tr>
<th>Government office region</th>
<th>Estimate</th>
<th>95% Confidence interval</th>
<th>Rate</th>
<th>95% Confidence interval</th>
<th>Number of anchor points</th>
<th>Total number of DAT areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>East of England</td>
<td>19,174</td>
<td>14,806–23,699</td>
<td>5.32</td>
<td>4.11–6.58</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>East Midlands</td>
<td>24,845</td>
<td>21,862–28,091</td>
<td>8.75</td>
<td>7.70–9.89</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>North West</td>
<td>54,953</td>
<td>51,683–59,897</td>
<td>12.22</td>
<td>11.49–13.32</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>South East</td>
<td>30,533</td>
<td>25,570–36,223</td>
<td>5.72</td>
<td>4.79–6.79</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>South West</td>
<td>29,491</td>
<td>27,782–32,114</td>
<td>9.07</td>
<td>8.54–9.87</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>West Midlands</td>
<td>37,311</td>
<td>34,233–41,245</td>
<td>10.66</td>
<td>9.78–11.78</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Yorkshire &amp; the Humber</td>
<td>41,064</td>
<td>38,211–44,623</td>
<td>12.35</td>
<td>11.49–13.42</td>
<td>10</td>
<td>15</td>
</tr>
</tbody>
</table>
the highest prevalence rates and, with the exception of London, the 35 to 64 year age group the lowest. London had the highest prevalence rate of problem drug use in the oldest age group compared with other regions, a rate that was higher than for those in the youngest age group in the region, while Yorkshire and the Humber had the highest prevalence rate for the 15 to 24 year age group.

**Discussion**

Attempts to estimate the national prevalence of problem drug use using survey methods produces estimates that are unfeasibly small. Previous studies have estimated the prevalence of problem drug use in English regions, and smaller areas, using indirect techniques (Frischer et al., 2001, 2006). However, these studies have extrapolated from estimates that lack a consistent case definition and that relate to a small number of potentially unrepresentative areas, and the resulting estimates have also lacked precision.

The national prevalence study (Hay et al., 2006) presented here is the first rigorous application of indirect estimation methods to estimate the prevalence of problem drug use in England. It is the first study to produce estimates for all DAT areas based on a consistent case definition and that extrapolated from a majority to a minority of areas. This paper reports on the second sweep of estimates from the study, the first sweep being for 2004/05, which produced

![Figure 1. Estimated age breakdown of problem drug users by government office region.](image)
Table II. Prevalence estimates (rate per 1000 population) for problem drug users by gender, age group, and government region 2005/06.

<table>
<thead>
<tr>
<th>Government office region</th>
<th>Females Estimate</th>
<th>Females 95% CI</th>
<th>Males Estimate</th>
<th>Males 95% CI</th>
<th>15–24 years Estimate</th>
<th>15–24 years 95% CI</th>
<th>25–34 years Estimate</th>
<th>25–34 years 95% CI</th>
<th>35–64 years Estimate</th>
<th>35–64 years 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>South East</td>
<td>2.84</td>
<td>2.39–3.49</td>
<td>8.61</td>
<td>7.15–10.09</td>
<td>7.04</td>
<td>6.02–8.35</td>
<td>12.35</td>
<td>10.10–14.56</td>
<td>3.28</td>
<td>2.73–3.94</td>
</tr>
</tbody>
</table>
a national estimate of 327,466 (95% CI 325,945 to 343,424) problem drug users, or 9.93 (95% CI 9.88 to 10.41) per thousand population aged 15 to 64 years (Hay et al., 2006).

There has been minimal and insignificant change in the prevalence of problem drug use between 2004/05 and 2005/06 at the national and government office region level. However, a significant increase has been observed for some DAT areas, particularly in the North East of England. As this paper has shown, a large proportion of problem drug users in the North East are from the younger age range. Previous research suggests this age distribution is characteristic of a recent increase in incidence whereby young people have joined the using population, and an increase in prevalence is an expected outcome (Millar, Gemmell, Hay, Heller, & Donmall, 2006). It is acknowledged that insufficient time has lapsed between the two study periods to report confidently on trends, and any significant changes should be interpreted as a possible change in prevalence rather than evidence of a definite change. However, building on this existing work will enable trends to be tracked in the future.

One of the major strengths of this study is the availability of a very large number of anchor points, obtained via capture–recapture, on which to base the multiple indicator model estimates. Anchor point estimates for problem drug use were available for 110 of the 149 areas. Uniquely, these anchor estimates were based on a consistent case definition and consistent data sources. This, combined with the use of simulation methods to derive 95% confidence intervals, has resulted in a much more precise national estimate than those previously derived via indirect techniques.

There are, however, specific issues in applying a method that is more frequently used to estimate prevalence at the local level in a more systematic fashion across the whole country. In a capture–recapture analysis for a single area, a wide range of different models can be fitted to the overlap data, including those that include dependencies between three separate data sources. The study took a pragmatic approach in only fitting the simplest 22 models to the available overlap data. The reasons for this were twofold; first to simplify the analysis by reducing the number of different competing models/estimates that need to be considered. Second, the more complex models tended to offer estimates with wider confidence intervals. It was decided that an estimate derived using multiple indicators methods would be more appropriate. Sensitivity analyses were undertaken to establish (in a randomly selected number of DAT areas) whether fitting more complex models would impact on the size of national or regional estimates. From those sensitivity analyses it was established that restricting the capture–recapture analyses to the simplest 22 models did not introduce any significant bias into the estimates.

As in other epidemiological applications of capture–recapture, the extent to which the study has entirely met the assumptions that underpin the method cannot be known (Cormack, 1999). In common with previous applications to drug-user populations, inaccuracies in matching individuals and migration or mortality during the study period are factors that may have reduced the observed
overlap between lists. This would usually have the effect of inflating the resulting estimates.

Of course, the definition of problem drug use employed in this study does not adhere to diagnostic criteria for addiction. This is, in part, driven by the availability of suitable data as it is inconceivable that non-medical sources would record, for example, DSM-IV diagnostic criteria (American Psychiatric Association, 2003). However, it is desirable that the definition of problem drug use should reflect both medical and social/legal consequences (Hickman et al., 1999).

It is, however, difficult to corroborate the estimates derived within this study as, without another method for accurately establishing prevalence, direct comparisons with competing estimates cannot be made. Approaches such as the DELPHI method have been used elsewhere (Hutchinson, Bird, Taylor, & Goldberg, 2006). It may be possible to try to assess the estimates against local opinions and that had been done in a previous English study (Frisher, Heatlie, & Hickman, 2007) although it is unclear as to how such an exercise can scientifically validate or dispute estimates derived from the statistical methods used in this study.

As expected, the estimate derived from this study is much larger than that suggested by population surveys (Roe & Man, 2006), but is consistent with the order of magnitude suggested by treatment surveillance data from the National Drug Treatment Monitoring System. It is also of a similar order of magnitude to an estimate produced by an earlier application of indirect estimation techniques for England in 2001 (287,670; 90% confidence interval 174,177 to 401,224; Frischer et al., 2006). However, the estimate for 2005/06 is based on a narrower, more specific, definition of problem drug use than the earlier estimate. Hence, notwithstanding the imprecision of the earlier estimate, the current estimate may suggest a somewhat higher prevalence of problem drug use in England than was previously thought.

Clearly these estimates have implications in relation to policy, prevention and future surveillance, including the surveillance of blood-borne viruses such as HIV and hepatitis C. With estimates systematically derived at the local and regional level for the entire country, those charged with the planning and provision of services now have information that they can use to more effectively target treatment services to where there is greatest need. There is an opportunity for the results of this study to be built upon by repeating the exercise to obtain information on trends in drug prevalence, which can then perhaps be used to gauge the success of local and national strategies to reduce numbers of drug users and increase the proportions of drug users in treatment.

Acknowledgments

The National Drug Evidence Centre receives funding from the National Treatment Agency for substance use and the Home Office. Funding for this
study was received from the Home Office, UK Government. The views expressed
in this article are those of the authors, not necessarily those of the Home Office
(nor do they reflect Government Policy).

**Ethical approval:** Use of National Health Service data was granted by the
Scottish Multi-centre Research Ethics Committee (ref. MRE00/54). The use of
data from the National Drug Treatment Monitoring System was subject to the
approval of the National Treatment Agency for Substance Misuse, the use of data
from the Drug Interventions Programme was subject to the agreement of the
Home Office, and the use of data from the Offender Assessment System was
subject to the agreement of the Home Office and the Chief Probation Officer for
England and Wales. The use of data from the Police National Computer was
subject to the approval of the Home Office and the Association of Chief Police
Officers.

**Declaration of interest:** The authors report no conflicts of interest. The authors
alone are responsible for the content and writing of the paper.

**References**


