Home Office
BUILDING A SAFE, JUST AND TOLERANT SOCIETY

Measuring different aspects of problem drug use: methodological developments

(2nd edition)

Editors:
Nicola Singleton
Rosemary Murray
Louise Tinsley

Home Office Online Report 16/06

Measuring different aspects of problem drug use: methodological developments

Editors: Nicola Singleton, Rosemary Murray, Louise Tinsley

Online Report 16/06

Acknowledgements

Local and national estimates of the prevalence of opiate use and/or crack cocaine use, 2004/05

The authors of chapter 2 of this report would like to thank the following for their support and assistance in making possible the study this chapter describes.

Colin Carroll, Eleanor Gallagher, Jennifer McKell, Chris Thexton and Anne Marie Ure, Centre for Drug Misuse Research, University of Glasgow.

Andrew Jones at National Drug Evidence Centre, University of Manchester.

Home Office staff including Chris May (CARAT data), Warren Evans, Lucy Cuppleditch and members of the PNC data team, and Natasha Garnham, Robin Moore and members of the OASys Data, Evaluation and Analysis Team (O-DEAT), who provided data for the study. Rosemary Murray and Nicola Singleton from the Drugs Analysis and Research Programme, who facilitated the provision of data from central sources and provided useful comment on reports and plans throughout the first sweep of the study.

Members of the Steering Group including representatives from National Treatment Agency for Substance Misuse, Department of Health, the National Probation Service, HM Treasury and the Association of Chief Police Officers and also Patient Information Advisory Group, for their support for the project.

Filip Smit, Trimbos Institute, Utrecht, The Netherlands and Ludwig Kraus, Institute for Therapy Research, Munich, Germany and colleagues in Drug Action Teams who have provided useful comments on the study.

Research Ethics approval for the use of NHS data in this study was granted by the Scottish Multi-centre Research Ethics Committee (ref MRE00/54).

Estimating the size of the UK illicit drug market

The work described in chapter 4 of this report was produced as part of a Home Office funded project “Sizing the UK illicit drugs market” (contract RDS/04/160).

The authors of chapter 4 are extremely grateful to Anne Line Bretteville-Jensen, Robin Fallows, Toni Makkai, Louise Tinsley, Peter Reuter, Nicola Singleton and Nigel South, for help and advice which improved the chapter enormously. Comments from staff of CIDA, FSS and NCIS are also gratefully acknowledged.

The views expressed in this report are those of the authors, not necessarily those of the Home Office (nor do they represent Government policy). Any errors or misinterpretations are the authors’ responsibility.

Contents

List of tables

List of figures

List of boxes

1. Introduction

Overview and policy context

The drug strategy, which was updated in 2002, aims to reduce the harm caused by illegal drugs. It has four key strands:

- reducing drug-related crime;
- reducing the supply of illegal drugs;
- preventing young people from becoming drug users; and
- reducing drug use and drug-related harms through treatment and support.

An understanding of the extent of drug use and its impact on different aspects of society is key to the development of appropriate policies for tackling the problem and monitoring the impact of interventions and services. However, the illegal nature of drug use and supply makes the use of traditional approaches to estimating the extent of the problems inappropriate; for example, the most problematic users lead chaotic lives so that they are unlikely to be represented in the household surveys which provide a measure of drug use in the general population. The Updated Drug Strategy 2002 drew on the information sources that were available at the time. However, these had limitations in terms of their coverage, the time periods covered and their ability to provide any local level information.

Therefore a programme of research was commenced to address these information limitations and a number of feasibility studies investigated alternative methods for measuring the size of the drug use problem from different perspectives. A number of reports have therefore been published that consider suitable estimation methods and give estimates of the number of problematic drug users (Frisher *et al.*, 2004; Hickman *et al.*, 2004; Millar *et al.*, 2004), the size of the market for illicit drugs (Bramley-Harker, 2001) and the economic and social costs of class A drug use (Godfrey *et al.,* 2002). These studies provided estimates which were the best available at the time, but they were conceived as a starting point for further developments and hence also identified a number of ways in which they needed to be improved.

Further work was commissioned to build on these early studies and to provide more robust estimates on which to base and monitor drug interventions. The results of these new studies are brought together in this report. The estimates they provide will allow policy-makers and those involved in providing interventions and services to move away from reliance on the less robust and incomplete estimates from the earlier studies and to move towards more regular production of these basic data for planning and evaluation of the wide range of work being undertaken to tackle the problem of drug use.

Estimating the number of problematic users

Chapter 2, ‘New local and national estimates of problem opiate and/or crack cocaine use (2004/05)’ by Gordon Hay *et al.*, describes the results from the first year of a three-year project. This study provides for the first time a robust national estimate of this important target group which is precise enough to allow monitoring of trends over time in future years (previous estimates had very wide confidence intervals within which the “true” estimate might lie). In addition, the project has provided robust local estimates with age breakdowns, which will be invaluable for treatment planning and to allow the identification of areas with emerging drug problems. These are available in separate regional reports on the NTA website[1].

Drugs consumption and its social impact

The new national estimate of the number of problematic users has been used as a basis for work to provide updated estimates of the economic and social costs of Class A drug use. This

[1] http://www.nta.nhs.uk/

is described in Chapter 3. These cost estimates provide a measure of the total costs to society of Class A use, both by problematic and recreational users, for the year 2003/04. As well as using the new estimates of the number of problem drug users, which has resulted in a much tighter estimate of costs, this new estimate has taken advantage of new sources of data to provide a more comprehensive assessment of the cost of drug use. These improvements have resulted in a more robust estimate of current costs but also mean that the results presented here for 2003/04 are not comparable with those relating to 2000.

Chapter 4 contains a report of the study estimating the size of the UK market for six key illicit drugs: cannabis, amphetamines, ecstasy, powder cocaine, crack cocaine and heroin. The analysis uses a survey-based demand side approach which builds on the previous study (Bramley-Harker, 2001) but differs from it in a number of ways which makes direct comparison of the estimates inappropriate. The new approach offers the potential to produce updated estimates of market size at regular intervals to provide a broad indicator of long-term trends in the size of the illicit drug market.

All of the figures in the new reports are more robust than the ones that came from the earlier feasibility studies. They contain significant improvements and the changes incorporated mean that they are not comparable to the previous estimates. They have also been designed to allow updating in the future and to serve as a baseline for monitoring trends. The information they contain is essential for planning and monitoring interventions and to inform future development of the drug strategy.

2. Local and national estimates of the prevalence of opiate use and/or crack cocaine use (2004/05)

Dr Gordon Hay[2], Maria Gannon[2], Jane MacDougall[2], Tim Millar[3], Catherine Eastwood[3] and Professor Neil McKeganey[2]

Summary

Overview

Information about the prevalence of problem drug misuse should be an essential part of the evidence base used to formulate policy, inform service provision, and assess the wider population impact of interventions. Although direct enumeration is not possible, indirect techniques can provide estimates of drug misuse prevalence. This research uses data sources that are available at the local and national level to estimate the prevalence of problem drug misuse.

Estimates are provided for the 149 Drug Action Team (DAT) areas and nine Government Office Regions in England. Two established prevalence estimation methods are used: the capture-recapture method and the multiple indicator method.

The capture-recapture method has been used to estimate the prevalence of problem drug use in the majority of DAT areas in England. The multiple indicator method provided local estimates in the remaining DAT areas. The national estimate for problem drug use was derived as the sum of the 149 DAT area estimates.

Data sources

Four sources of data were available within which individual problem drug users (defined as those who use opiates and/or crack cocaine), opiate users and crack cocaine users could be identified. These sources of data are drug treatment, probation, police and prison data. There were three sources of data from which drug injectors could be identified (drug treatment, probation and prisons).

Persons resident in each DAT area, in contact with these sources during 2004/05, known to be using heroin, methadone, other opiate drugs, or crack cocaine were included in the analysis. Only those aged from 15 to 64 were included. The overlap between data sources was determined via comparison of initials, date of birth and gender within each DAT area. Established statistical modelling techniques were used to examine this overlap and to produce prevalence estimates stratified by age group, gender, and DAT area of residence.

[2] The Centre for Drug Misuse Research, University of Glasgow, 89 Dumbarton Road, Glasgow, G11 6PW
In collaboration with
[3] The National Drug Evidence Centre, University of Manchester, Rutherford House, Manchester Science Park, Manchester, M15 6GG

Address for correspondence: Dr Gordon Hay at the Centre for Drug Misuse Research: Telephone: 0141 330 5413; Fax: 0141 330 2820;
E-mail: prevalence@gla.ac.uk

Methods

Two methods have been used to estimate the local and national prevalence: the capture-recapture method, which was used in 117 out of the 149 DAT areas (79%) to obtain problem drug use prevalence estimates; and the multiple indicator method, which was used in the remaining 32 DAT areas. The capture-recapture method uses information on the overlap between data sources that are available at the local level (i.e. information on the number of individuals appearing in more than one data source) to provide estimates of the size of the hidden population (i.e. problem drug users not identified from any data source). The multiple indicator method models the relationship between the prevalence of problem drug use and readily available indicators such as aggregate numbers of drug users in treatment or committing drug-related crimes in those areas where these prevalence estimates are available. It can therefore provide prevalence estimates for areas where capture-recapture estimates are not available.

Results

Table S2.1 presents the national estimates and their associated 95 per cent confidence intervals. Total estimates for problem drug use (opiate and/or crack cocaine use), opiate use, crack use and drug injecting for each Government Office Region are shown in Table S2.2 and Table S2.3.[4]

Table S2.1: National prevalence estimates and rates per 1,000 population aged from 15 to 64 with 95 per cent confidence intervals

Drug	Estimate	95% confidence interval	Rate	95% confidence interval
Problem	327,466	325,945 – 343,424	9.93	9.88 –10.41
Opiate	281,320	279,753 – 292,941	8.53	8.48 – 8.88
Crack	192,999	188,138 – 210,763	5.85	5.70 – 6.39
Injecting	137,141	133,118 – 149,144	4.16	4.04 – 4 .52

Table S2.2: Estimated number of problem drug (opiate and/or crack cocaine) users and opiate users by Government Office Region

Government Office Region	Problem drug users			Opiate users		
	Estimate	95% CI		Estimate	95% CI	
East of England	23,081	22,029	25,313	19,518	18,539	21,558
East Midlands	23,142	22,516	25,442	21,241	20,815	22,813
London	74,417	71,845	81,299	55,139	52,611	59,799
North East	15,853	14,987	17,683	13,729	13,099	15,657
North West	51,110	46,455	55,659	43,996	40,737	48,221
South East	33,854	32,495	38,306	28,727	27,267	31,648
South West	30,455	29,536	32,936	27,541	26,681	29,865
West Midlands	36,834	35,276	39,726	34,661	33,656	37,173
Yorkshire and the Humber	38,720	37,708	40,911	36,768	35,674	39,028
England	**327,466**	**325,945**	**343,424**	**281,320**	**279,753**	**292,941**

4 In the body of the report, data within tables are provided at the Government Office Region level. Full tables at the DAT area level are provided on the NTA website.

Table S2.3: Estimated number of crack cocaine users and drug injectors by Government Office Region

Government Office Region	Crack users			Injectors		
	Estimate	95% CI		Estimate	95% CI	
East of England	14,102	11,779	22,425	9,418	6,252	13,061
East Midlands	12,952	11,417	16,627	11,796	10,487	13,518
London	51,312	48,156	55,850	17,909	16,161	24,002
North East	7,376	5,636	10,899	8,959	7,638	10,597
North West	29,752	25,533	34,423	22,089	18,781	25,204
South East	19,261	16,097	26,554	13,778	12,044	17,802
South West	17,552	15,509	20,785	17,444	15,930	19,532
West Midlands	20,827	19,012	26,593	14,734	13,589	17,007
Yorkshire and the Humber	19,865	18,239	23,095	21,014	19,860	22,788
England	**192,999**	**188,138**	**210,763**	**137,141**	**133,118**	**149,144**

Table S2.4 and Table S2.5 present the prevalence rates per thousand of the population aged from 15 to 64, again by Government Office Region for problem drug use, opiate use, crack cocaine use and drug injecting.

Table S2.4: Estimated prevalence (rate per 1,000 population aged from 15 to 64) of problem drug (opiate and/or crack cocaine) use and opiate use by Government Office Region

Government Office Region	Problem drug use			Opiate use		
	Estimate	95% CI		Estimate	95% CI	
East of England	6.48	6.18	7.11	5.48	5.20	6.05
East Midlands	8.23	8.00	9.04	7.55	7.40	8.11
London	14.35	13.86	15.68	10.64	10.15	11.53
North East	9.50	8.98	10.59	8.22	7.85	9.38
North West	11.43	10.39	12.45	9.84	9.11	10.79
South East	6.40	6.14	7.24	5.43	5.15	5.98
South West	9.44	9.16	10.21	8.54	8.27	9.26
West Midlands	10.62	10.17	11.45	9.99	9.70	10.71
Yorkshire and the Humber	11.74	11.43	12.40	11.15	10.81	11.83
England	**9.93**	**9.88**	**10.41**	**8.53**	**8.48**	**8.88**

Table S2.5: Estimated prevalence (rate per 1,000 population aged 15 to 64) of crack cocaine use and drug injecting by Government Office Region

Government Office Region	Crack users			Injectors		
	Estimate	95% CI		Estimate	95% CI	
East of England	3.96	3.31	6.30	2.64	1.76	3.67
East Midlands	4.60	4.06	5.91	4.19	3.73	4.81
London	9.90	9.29	10.77	3.45	3.12	4.63
North East	4.42	3.38	6.53	5.37	4.57	6.35
North West	6.65	5.71	7.70	4.94	4.20	5.64
South East	3.64	3.04	5.02	2.60	2.28	3.36
South West	5.44	4.81	6.44	5.41	4.94	6.05
West Midlands	6.00	5.48	7.66	4.25	3.92	4.90
Yorkshire and the Humber	6.02	5.53	7.00	6.37	6.02	6.91
England	**5.85**	**5.70**	**6.39**	**4.16**	**4.04**	**4.52**

Thus, in total, there are an estimated 327,466 problem drug users (opiate and/or crack cocaine users) in England (95% CI 325,945 to 343,424); this corresponds to 9.93 per thousand of the population aged from 15 to 64 (95% CI 9.88 to 10.41). In terms of opiate users, there are an estimated 281,320 people (95% CI 279,753 to 292,941) in England who use these drugs (8.53 per thousand population aged from 15 to 64, 95% CI 8.48 to 8.88) whereas it is estimated that 192,999 people (95% CI 188,138 to 210,763) use crack cocaine (5.85 per thousand population aged from 15 to 64, 95% CI 5.70 to 6.39). It should be noted that the majority of people using crack cocaine are also using opiates and that crack cocaine may neither be their main drug of use nor indeed the drug that is causing them the most problems. Finally there are an estimated 137,141 opiate and/or crack cocaine users (95% 133,118 to 149,144) in England who inject drugs (4.16 per thousand of the population aged from 15 to 64, 95% CI 4.04 to 4.52).

In terms of regional differences, London is the Government Office Region with the largest prevalence of problem drug use at around 14 per thousand population aged from 15 to 64 compared with Yorkshire and the Humber at around 12 per thousand and the North West and the West Midlands at around 11 per thousand and with the lowest prevalence at around six per thousand in the South East and East of England. When comparing opiate use prevalence, both London and Yorkshire and the Humber have the highest prevalence at around 11 per thousand, followed by the North West and West Midlands at around ten per thousand and the South East and East of England again have the lowest prevalence at around five per thousand. Injecting drug use prevalence varies less between regions, ranging from around six per thousand in Yorkshire and the Humber to around three per thousand in London, East of England and the South East. London has a far higher estimated prevalence of crack cocaine use at just under ten per thousand population, in comparison to prevalence of between around four and seven per thousand in all other regions.

Discussion and conclusion

This study has demonstrated that it is possible to provide estimates of the prevalence of problem drug use (defined as the use of opiates and/or the use of crack cocaine) as well as the prevalence of opiate use, crack cocaine use and drug injecting at the local, Government Office Region and national level. The prevalence estimates do, however, need to be considered alongside their associated confidence intervals particularly when making comparisons across DAT areas or between different definitions of drug use within DAT areas. These estimates are the result of the first yearly sweep of a three-year project and while they are the best estimates that have ever been produced at a national level in England, there is further scope for methodological improvements which will be investigated in subsequent sweeps.

Background

Information about the number of people who use illicit drugs such as heroin, other opiates or cocaine is a key element of the evidence base used to formulate policy and inform service provision and provides a context in which to understand the population impact of interventions to reduce drug-related harm. To direct resources effectively, it is desirable to know about the prevalence of drug use at the local level. To determine the extent to which treatment may reduce harm to communities, it is necessary to know what proportion of the number of drug users in any given area is engaging with treatment. Direct enumeration of those engaged in a largely covert activity such as the use of heroin is not possible, and large, household surveys such as the British Crime Survey tend to underestimate numbers of those individuals whose drug use is the most problematic and whose lives are often the most chaotic. However, indirect techniques can be applied to provide estimates of drug use prevalence. The research covered in this chapter aims to use data sources that are available at the local and national level to provide estimates of the prevalence of problem drug use in all Drug Action Team areas in England and thus provide regional and national prevalence estimates.

Drug Action Team areas

Drug Action Teams (DATs) are the partnerships responsible for the delivering the Government's drug strategy at the local level. They ensure that the work of local agencies is brought together effectively and that cross-agency projects are co-ordinated successfully. DATs are supported by the Home Office teams in the nine Regional Government Offices and centrally by the Home Office Drugs Strategy Directorate. There are 149 DATs in England and, since April 2001, the DAT areas have been aligned with local authority boundaries. Figure 2.1 shows the boundaries of the DAT areas within Government Office Regions.

Table 2.1: Population estimates by Government Office Region, gender and age group (thousands), 2004 mid-year estimates

Government Office Region	Total (15-64)	Male	Female	15-24	25-34	35-64
East of England	3,562.3	1,733.0	1,829.3	650.1	694.4	2,217.8
East Midlands	2,813.1	1,351.2	1,461.9	554.1	527.4	1,731.6
London	5,184.2	2,345.0	2,839.2	989.6	1,454.7	2,739.9
North East	1,669.6	807.9	861.7	344.7	300.6	1,024.3
North West	4,470.7	2,147.4	2,323.3	905.9	853.9	2,710.9
South East	5,292.3	2,580.7	2,711.6	996.8	1,023.2	3,272.3
South West	3,225.8	1,550.4	1,675.4	608.4	580.0	2,037.4
West Midlands	3,469.7	1,691.1	1,778.6	697.0	671.9	2,100.8
Yorkshire and the Humber	3,298.7	1,586.3	1,712.4	686.2	626.0	1,986.5
England	**32,986.4**	**15,793.0**	**17,193.4**	**6,432.8**	**6,732.1**	**19,821.5**

Figure 2.1: Map of the 149 Drug Action Team areas (light outlining) and nine Government Office Regions (dark outlining) in England

The estimated size of the population[5] aged from15 to 64 in each DAT area (along with gender and age-group estimates) is provided within the regional reports which are available on the NTA website. This information is summarised at the Government Office Region level in Table 2.1.

5 2004 mid-year estimates from **http://www.statistics.gov.uk**, accessed November 2005

Case definitions

Any study that aims to estimate the prevalence of problem drug use needs to provide a description of the case definition employed for identifying problem drug use in the contributing data sources and thus the case definition of the resultant estimates. Clearly a pattern of drug use that constitutes a problem for one individual may not constitute a problem for another although opiates and crack cocaine are commonly considered to be the drugs that cause most harm to an individual and communities.

While measures of drug dependence have been devised, such as the International Classification of Disease (ICD) codes or the Diagnostic and Statistical Manual of Disorders (DSM) diagnostic criteria, these measures are not commonly used in sources of data on drug use and therefore, would be of little use in this type of prevalence estimation exercise. Rather, the study considers drug use measures that are readily available in data that can systematically be collated across the country. As such, the case definition of the prevalence estimates depends heavily on the case definitions used by the contributing sources. Moreover, the case definitions of the resultant prevalence estimates need to reflect case definitions that are common across all data sources. Thus, although a drug treatment service could provide information on clients whose cannabis use is so problematic that they are receiving structured treatment for their cannabis use, it is difficult to ascribe the level of someone's cannabis problem from possessions data supplied by the police. Similarly there may be large differences in the level of cocaine use by individuals identified in police data on cocaine possession compared with those identified in data on cocaine-related treatment demand. Non-problematic use of opiates and crack cocaine is comparatively rare so it is safe to assume that those convicted or cautioned for possessing those drugs are likely to be problematic users whereas the same could not be assumed for drugs such as cannabis and powder cocaine which are commonly used intermittently.

The study therefore employed the following as the case definition for problem drug use.

- Use of opiates and/or the use of crack cocaine.

It should be noted that the case definition focuses on the 'use' of opiates and/or crack cocaine rather than the 'misuse' of these drugs or addiction to either drug. The case definition does not include the use of cocaine in a powder form, the use of amphetamine, ecstasy or cannabis, or the injecting of drugs by people who do not use opiates or cocaine.

The study also provides separate estimates of the prevalence of opiate use, and of the prevalence of crack cocaine use. As many drug users use both drugs, but may only have been convicted of possession of one of these drugs, it is not possible to directly obtain estimates of the number of people who only use crack cocaine (i.e. use crack cocaine but do not use opiates) or only use opiates (without also using crack cocaine). It may be thought that subtracting the estimates of the number of people who use opiates from the problem drug use prevalence estimates may provide some information on the likely number of individuals who use crack cocaine but do not use opiates. However, care must be taken when making such a comparison, not least because the figures are not absolute numbers but estimates that should only be considered alongside their associated confidence intervals. Similarly, it was also not possible to directly estimate the number of problem drug users who use both opiates and crack cocaine, primarily due to difficulties in identifying such people from the available data sources.

In addition to the problem drug use (opiate and/or crack cocaine), opiate use and crack cocaine use estimates, the prevalence of drug injecting is also explored. For consistency between the problem drug use estimates and the drug injecting estimates, only the prevalence of injecting drug use by opiate and/or crack cocaine users has been estimated (i.e. they are a subset of the opiate and/or crack cocaine users).

All data quoted in this chapter refer to the financial year from 1 April 2004 to 31 March 2005. The age range employed within the study is from 15 to 64 and where the estimates have been stratified by age group, these are from 15 to 24, from 25 to 34 and from 35 to 64. To derive age from date of birth, the individual's age on 1 October 2004 (the midpoint in the financial year 2004/05) was calculated and those who were under the age of 15 or over the age of 64 were excluded. Individuals with missing data fields, such as gender, forename initial or surname initial were also excluded, as were individuals where it was not possible to assign DAT area of residence (or those that were resident outside England).

Research aims

The aims of the first sweep of the research study are to apply appropriate prevalence estimation methods to:

- establish estimates of the prevalence of problem drug use over three successive years for all 149 Drug Action Team areas within England; and
- establish the national prevalence of problem drug use in England over three successive years.

The study is being carried out over a three-year period and will provide prevalence estimates at the end of each of the three study 'sweeps'. This chapter describes the results from the first sweep of the project and provides estimates for the financial year 2004/2005.

The wider aims of the study are to examine changes in the prevalence of problem drug use over the three years of the study, to improve the methodologies used to estimate problem drug use prevalence and to build capacity at the national and local level for undertaking prevalence estimation. The first sweep of the study focuses primarily on obtaining reliable and timely prevalence estimates for a large number of DAT areas while leaving the capacity building element for later sweeps. Methodological improvements do, however, run as an important thread throughout all three sweeps of the study.

Prevalence estimation methods

This research applies two methods: the capture-recapture method and the multiple indicator method (also called the multivariate indicator method or MIM), to estimate the prevalence of problem drug use in England in 2004/05. These two methods appear to offer the most cost-effective and straightforward approach to establishing valid local and national prevalence estimates. The benefits of these methods are that: they do not rely on drug users' self-reported use of substances; it is possible to provide estimates of prevalence stratified by key characteristics such as age and gender; they use a standard set of procedures that are tried and tested and allow for replication; and they build upon existing routinely collected data. This section outlines the key issues when applying either method to problem drug use prevalence estimation.

The capture-recapture method has been used in a number of settings in the United Kingdom to estimate the prevalence of problem drug use at the local (e.g. DAT area) level and the national (e.g. Scotland) area level. The multiple indicator method has been less extensively used in the United Kingdom. Previously, the Home Office commissioned a series of feasibility studies that used the same two methods to estimates problem drug use prevalence in 2000/01 (Frischer *et al*., 2004; Hickman *et al*., 2004; Millar *et al*., 2004). The reports from those studies comprehensively describe the methods and further information can be found elsewhere, particularly within a series of reports prepared by the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA)[6].

In this study, the capture-recapture method has been used to estimate the prevalence of problem drug use in the majority of DAT areas in England. The multiple indicator method

6 **http://www.emcdda.eu.int?nnodeid=1372** (accessed 20/12/05)

provided local estimates in the remaining DAT areas. The national estimates were therefore derived as the sum of the 149 DAT area estimates.

This section provides a brief overview of the two methods whereas later sections describe the particular methodological issues encountered within this study.

Capture-recapture methods

Capture-recapture methods were first developed over a century ago to estimate the size of animal or fish populations. In its basic form the method involves capturing a sample of animals, marking and then releasing them. A second sample is then captured; the proportion of marked animals in this second sample is assumed to be equivalent to the proportion of animals in the population that were captured in the first sample, hence the population size can be estimated. For example, if 100 fish are caught, marked and released and a second sample of fish is caught, of which ten per cent are found to be marked, then the 100 fish in the first sample is equivalent to ten per cent of the population, hence the population size is 1,000.

When the method is applied to estimating the size of drug-using populations, the two samples are replaced with lists of individuals constructed from sources such as drug treatment services, police data, probation data and so on. The number of individuals in each data source is equivalent to the size of the animal samples in the example above and the number appearing in both sources is equivalent to the number of recaptured, marked, animals. Hence the size of the population can be estimated, as above. In drug misuse prevalence estimation, samples are typically drawn from health (e.g. client lists supplied by drug treatment agencies) and/or criminal justice (e.g. police records or probation data) sources.

There are several assumptions that must be met in applying capture-recapture methods. These have a bearing on the work reported here and are outlined below.

Assumption 1: The population is 'closed'
It is assumed that the population of drug users does not change during the period that is being studied. Specifically, it is assumed that drug users do not begin to use drugs or stop using drugs and that drug users do not move into or out of the area that is being studied within that time period. By taking a period of one year it may be true that the population can be assumed to be closed, however there may be areas which are affected by drug users moving in or out, i.e. due to seasonal work. Within the first sweep of the study it is indeed assumed that the population is closed; however, further analyses will be undertaken to examine whether movement between DAT areas (by individual drug users) within a calendar year would impact on the prevalence estimates.

Assumption 2: Drug users who are in more than one source are identified as such
To meet this assumption, the method of finding the overlap between lists must be accurate. In estimating drug misuse prevalence, individuals are usually identified by a code comprising initials, date of birth and gender (or an encrypted code based on those identifiers) to meet data protection and client confidentiality requirements. No single method of identifying overlaps is perfect; cross-referencing on full initials, gender and date of birth would not identify that William Smith and Bill Smith are the same person, or that Mary Jones is sometimes known by her married name of Mary Brown. Thus using such an approach could fail to identify some true matches. However, cross-referencing with a reduced set of identifiers can lead to false matching, i.e. counting more matches than are actually there, and thus impact on the validity of the prevalence estimates. In this study, two separate records were taken to be a match if both initials, the date of birth and gender were the same.

Assumption 3: Presence in one source does not influence presence in another
In other words, the fact that someone is in one data source does not increase (or decrease) the chances that they are in another one. If an individual's presence in one source affects the probability of his or her presence in the other, this will artificially inflate or deflate the observed overlap between samples. This assumption is often violated when looking at drug-using populations, for example those in probation data may be more likely to appear in prison data.

This problem can be reduced by examining three or more samples and by using log-linear modelling techniques and this is the approach taken in this study and other recent studies in the United Kingdom.

Assumption 4: All drug users are equally likely to appear in any data source

This assumption may not be met if there is 'heterogeneity' within data sources, leading to individuals having different probabilities of appearing in a particular source. For example, young drug users may be less likely than older drug users to have started treatment and therefore appear in a treatment data source. Stratifying samples according to known characteristics such as age and gender may reduce this problem and, in this study, age group and gender specific models were considered alongside the unstratified models in every DAT area.

Assumption 5: Data sources should be representative

In order to meet this assumption, data sources must have equal coverage of the area they serve and also be representative of gender, age group, ethnic group, type and severity of drug use and so on. That is not to say that, for example, a treatment service should have equal numbers of female and male clients, rather the probability that a female drug user in the community appears in a treatment data source is similar to that of a male drug user (or that drug users in a rural part of a county are as likely to access treatment as drug users living in a town or city).

Model selection

The capture-recapture method therefore fits statistical models to the overlap pattern between three or more sources of data on drug-using populations. Different statistical models are examined for each individual analysis to reflect different possible relationships between sources. For example, one model would assume that all sources are independent of each other (i.e. the chance of appearing in one source is not influenced by presence in another source) whereas a second model would assume, for example, that being in treatment would influence the chance that someone also turns up in a criminal justice source. When the capture-recapture analyses use four data sources there are over a hundred different models that could fit the overlap pattern; however, simpler models are preferred to more complex models. The study team therefore restricted the various analyses, in the first place, to the 22 more simple models (the model that assumes all sources are independent, six models that account for relationships between different pairs of sources and 15 models that account for relationships between sets of two different pairs of sources).

In most capture-recapture analyses, prevalence estimation results are taken to be the estimates derived from the best fitting models. Hook and Regal (1997) discuss the validity of methods for selecting between the estimates derived from each of the 22 different models (or indeed more complex models). They also consider approaches to combining different models together to give a 'weighted' estimate, such as using weights proposed by Schwarz (1978). Both approaches are based on values known as 'information criteria' that not only describe how well any given model fits the overlap data but also assist in comparing between different models to suggest a 'best fitting' model (and thus a best estimate). As referenced by Hook and Regal, one of the more commonly used information criteria is the Akaike Information Criteria (AIC) and this has been used to assist in model selection in this study.

Once estimates have been obtained, there are different approaches to deriving confidence intervals. The favoured approach in most studies, and the one that has been employed here, is that outlined by Cormack (1992) which results in more valid confidence intervals where the estimate almost always lies closer to the lower bound of the confidence interval.

Multiple indicator methods

The multiple indicator method can be used to generalise from areas for which local prevalence estimates have been established (for example by applying the capture-recapture method), to obtain estimates for areas without directly derived estimates. The areas that already have local prevalence estimates are known as 'anchor points'. The use of this method assumes that the prevalence of drug use is correlated with readily available data such as the published number of drug users in treatment, crime statistics or social statistics such as deprivation scores. This correlation or relationship can take the form of a regression model and standard regression methods can be used to identify the relationship and provide estimates for the areas that are not anchor points. Information on the indicators has to be available at the local level and relate to the same geographical areas as the prevalence estimates. Often this may not be the case as health and criminal justice areas may not be coterminous. The method also assumes that the relationship between prevalence estimates and indicators within the anchor point areas is the same as the relationship between the indicators and the areas that are not anchor points.

The prevalence estimates that are used as anchor points in a multiple indicator analysis will have an impact on the prevalence figures derived for other areas. These anchor points should be available for more than two of the areas (preferably far more than two areas) and must be valid and reliable as they determine the parameters of the regression model. By including a large number and range of anchor point areas in the analysis it can perhaps be assumed that they will be representative and that the estimates derived using the multiple indicator method will be more reliable than if a small number of anchor point areas had been used.

Data sources

This section outlines the data that have been used to estimate the prevalence of problem drug use in England. It outlines the data employed by the capture-recapture analyses and then goes on to consider the data used by the multiple indicator analyses.

Data used in the capture-recapture analyses

Four main sources of data on problem drug use, which were suitable for use in the capture-recapture analyses, were available at the national level.

- The National Drug Treatment Monitoring System (NDTMS).
- The National Offender Management Service Offender Assessment System (OASys).
- 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 from the Police National Computer (PNC).
- Counselling, Assessment, Referral, Advice and Throughcare services data for drug users in prison (CARAT).

All data were collated centrally and no primary data collection was involved in the study.

Use of National Health Service (NHS) data was subject to approval by the Scottish Multi-Centre Research Ethics Committee (reference MREC00/54). The use of the NDTMS data was subject to the approval of the National Treatment Agency for Substance Misuse, the use of CARAT and OASys data was subject to the agreement of the Home Office and the Chief Probation Officer for England and Wales and the use of PNC data was subject to the approval of the Home Office and the Association of Chief Police Officers (ACPO).

In addition, data from Hospital Episode Statistics (HES) were obtained from the Department of Health for inclusion in the capture-recapture analyses. As these data did not include initials, it was not possible (within this study sweep) to include that source as a fifth data source in the capture-recapture analyses. Aggregated HES data were, however, used within the multiple indicator analyses. As the data requested from the Department of Health could be considered as patient identifiable, additional approval was gained from the Security & Confidentiality Advisory Group of the Department of Health and the study obtained Section 60 support from the Patient Information Advisory Group[7] (reference PIAG 3-07(b)/2005).

7 http://www.advisorybodies.doh.gov.uk/piag/

One other potential source of data on problem drug use across England would have been data from the Drug Interventions Programme (DIP). A decision was taken not to include DIP data within the study as, over the time period being studied, consistent and comparable data at the individual level across the whole of England were not available.

National Drug Treatment Monitoring System

The National Drug Treatment Monitoring System is a treatment surveillance mechanism based on data collection by regional centres with a national dataset being co-ordinated by the National Treatment Agency for Substance Misuse (NTA) in collaboration with the National Drug Evidence Centre at the University of Manchester (NDEC). Public Health Observatories (PHOs) are responsible for gathering information about those problem drug users who contact agencies that provide treatment and care within tiers three or four of the National Commissioning Framework. In most cases, data are gathered direct from agencies' own electronic case management systems. Each month, PHOs provide a core set of data items about problem drug users in contact with agencies in their area to the NTA. These data are subject to preliminary validation by the NTA and are then provided to NDEC for further national validation and analyses. During the summer of 2005, NDTMS data were subject to supplementary verification by the NTA of the initials, date of birth and gender of the problem drug users that agencies reported during 2004/05.

NDTMS gathers information about up to three drugs that are used by each client. For the current study, NDEC selected from the 2004/05 NDTMS data a subset of records in which any use of heroin, methadone, other opiates and/or crack cocaine was mentioned. Records were then checked to ensure that details of initials, dates of birth and gender and DAT of residence were complete, and any records incomplete in this respect were discarded. Cases that were known to have not received treatment at tier three or four were also discarded. Dates of referral, triage, discharge, and treatment modality start and finish were checked to ensure that these conformed to a logical sequence and any records with an illogical sequence were discarded. Cases resident outside England were also discarded. Cases within the remaining subset were examined and those with a date(s) of triage, discharge, or modality start/finish that indicated contact with the reporting agency between 1 April 2004 and 31 March 2005 were extracted and retained for analysis.

Based on the remaining subset, NDEC constructed a sample that included one record for each problem drug user for each DAT area in which they were reported to be resident during the year. Individuals were isolated on the basis of their initials, date of birth and gender. Each record included details of the DAT area of residence and whether the individual had been reported to use heroin, methadone, other opiates and/or crack cocaine and injecting at any of the episodes at which they were reported to be resident in that DAT area. There were some individuals for whom DAT of residence was not reported at any treatment episode and these cases were also excluded from the analyses. Finally, details of initials, date of birth and gender were encrypted into an unrecognisable format, consistent with the encryption applied to other data sources in the study, before being supplied to the Centre for Drug Misuse Research at the University of Glasgow.

A problem with the NDTMS data has only recently been detected in that the data for some DAT areas, particularly DAT areas in the North West Government Office Region, have an excessively high level of missing data in the 'drugs used' fields. For some North West DAT areas, there were over 50 per cent of records that did not have a valid (for the purposes of this study) drug code. While the capture-recapture method is generally robust to incomplete data, a decision was taken to only consider undertaking a capture-recapture analysis for a North West DAT area if over 80 per cent of the records had a valid 'drugs used' code. This resulted in only four DATs in the North West receiving a valid capture-recapture estimate for problem drug use.

The National Offender Management Service Offender Assessment System

The Offender Assessment System (OASys) is a system developed jointly by the Prison and Probation Services. It is structured to help practitioners assess how likely an offender is to re-offend and assesses the risk of harm an offender poses to himself or herself and others. Within each OASys assessment there is a section about problem drug use which identifies the extent and type of problem drug use and its effects on the offender's life. This section takes the form of a grid where the levels of use of a range of drugs are recorded (and whether or not the drugs are injected). It was therefore possible to use OASys data to construct a data source for problem drug use, opiate use, crack cocaine use and also drug injecting. The Probation Service implemented OASys, initially as a paper-based system, across the 42 Probation Service areas of England and Wales during 2002/03 and completed conversion to an electronic IT-based system during 2003/04. Data for the financial year 2004/05 were used within the first sweep of this study. An extract of OASys that had any mention of any drug in the drug misuse section was supplied for the purposes of this study and this extract was stored in secure conditions within the Home Office. Members of the study team were given access to that extract to enable data cleaning and allocation of DAT area of residence prior to encryption. To allocate DAT area of residence, the postcode district of residence was considered. Postcode Address File (PAF) data were used to assign each postcode district (e.g. GL52 in Gloucestershire) to a single DAT area. If a postcode district was entirely within a DAT area (i.e. all of the population of the postcode district lived in the local authority area covered by the DAT) then the allocation was unequivocal. If, however, a postcode district straddled two (or more) DAT areas then assignment was made on the basis of where the largest proportion of the residents lived. Out of 20,798,136 addresses in England, 893,619 (4.3%) could potentially be assigned to the wrong DAT area using this approach to assigning DAT area of residence from postcode data.

As with the NDTMS data, those who were not within the age range employed by the study or not resident in England were excluded. In addition, those cases where it was not possible to assign DAT area of residence were also excluded. All offender-identifiable data were encrypted prior to the cleaned extract leaving the Home Office. The study team also had an opportunity to return to the Home Office to further clean the data and correct for any problems after an initial analysis of the data.

During collation of that year's data, it became apparent that there was a problem with the completeness of the data in one Probation Service area (Cheshire, covering the DATs of Cheshire, Halton and Warrington). Thus it has only been possible to use OASys data to inform prevalence estimation in 146 of the 149 DAT areas.

Although the OASys data should contain information on whether or not an individual drug user was a drug injector, initial analyses suggest that there may be a substantial amount of under-reporting of drug injecting within the OASys data in parts, or all, of the London Government Office Region. This has been borne in mind when obtaining drug injector prevalence estimates.

Misuse of Drugs Act 1971 Crimes from the Police National Computer

The Home Office also made available an extract of data concerning individuals convicted in connection with the Misuse of Drugs Act 1971 for either possession of or possession with intent to supply heroin, methadone or crack cocaine. This extract came from the Home Office's partial copy of the Police National Computer and the study team were given access to the extract and associated information under secure and supervised conditions within the Home Office. In a similar manner to the OASys data, DAT area of residence was assigned using postcode district of residence data and cases with insufficient information on the area of residence, those who lived outside England or those who did not meet the age criteria were excluded from the analyses. These data were encrypted prior to removal from the Home Office and the study team again had an opportunity to correct for any problems that arose on initial analyses of the data.

On initial inspection of the data, it appeared that there had been a problem with the data from one Police Force area (Suffolk covering the Suffolk DAT area). Thus it has only been possible to use the police data in 148 out of the 149 DAT areas. Furthermore, there was another Police Force area (Northumbria) where there were very few crack offenders identified from the PNC data. This could be because there were very few arrests for crack offences in the North East of England (possibly indicative of low levels of crack cocaine use) or it could be that crack cocaine offences are being recorded as cocaine offences within the PNC data for that area.

Counselling, Assessment, Referral, Advice and Throughcare

The CARAT service was established in 1999 as a universal service in every prison establishment across England and Wales. Prisoners can be assessed by a CARAT team, given advice about drug misuse and referred to appropriate drug services. Information on prisoners using the service was collected from April 2002 to March 2005. The submission of information was subject to the consent of the individual prisoner. CARAT data could potentially be collected at two points for each prisoner: firstly, after the initial CARAT assessment (usually soon after imprisonment); and secondly, when the case was closed (usually on the prisoner's release). The data included each prisoner's home area. Since May 2005, a modified form of the CARAT data collection has been part of the information system for the wider Drug Interventions Programme.

The study team were supplied an extract of the CARAT data from the RDS National Offender Management Service (NOMS) within the Home Office Research, Development and Statistics Directorate. These data relate to the financial year 2004/2005. As with the other data sources, cases that fell outside the required age range or were not residents of England were excluded from the analyses and only cases that identified opiate and/or crack cocaine use were used in the subsequent analyses. DAT area of residence was taken from the local authority of residence field supplied within each CARAT assessment completed by the CARAT worker. It is, however, possible that some cases have been allocated to the wrong DAT (for example by someone resident in Trafford, Greater Manchester stating that they lived in Manchester or, for example, by someone resident in Southend-on-Sea, Thurrock or even Barking and Dagenham stating that they lived in Essex). On initial analysis of the CARAT data there were no discernible patterns to suggest any systematic misallocation of DAT area of residence. The CARAT data were also encrypted prior to removal from the Home Office and the study team were able to seek additional clarification from the relevant senior scientific officer after the initial analyses of the data.

Data cleaning

Once data from the four contributing data sources had been collated, a further data cleaning exercise was carried out to further ensure that each record met the inclusion criteria for the study. Within each data source, the data were filtered to ensure that an individual only appeared once within a DAT area for the non-NDTMS data sources (this exercise had previously been undertaken on a wider NDTMS dataset within NDEC). To achieve this, 'hard' matching was employed. To expand on this, two records within the same DAT area were deemed to refer to the same individual if the gender, date of birth, forename initial and surname initial all matched. This matching procedure does not, however, account for different initials being used by individuals (i.e. difference in maiden name/married name or names such as Bill/William) but could occasionally incorrectly match different people with the same initials, gender and date of birth. Two other matching procedures were examined: 'soft' matching where forename initials are allowed to differ; and 'DoB' matching where only the date or birth (and gender) were used to match individuals within a DAT. The impact of only matching on dates of birth and gender was examined and it was concluded that it was not appropriate to match on such a limited set of identifiers as, within many DAT areas, there were many sets of records that would be assigned as matches despite the forename initials and/or surname initials being different. As with the NDTMS data, care was taken to include all relevant information within the resultant single record per DAT area (for example if someone was noted as a heroin injector in one OASys assessment then later in the year in another OASys assessment as using heroin and crack cocaine then a single record noting heroin use, crack cocaine use and drug injecting would be used within the analyses).

It is possible that the same individual appeared in data from different DAT areas within the same year. Care must therefore be taken in summing the 'known' number of drug users from any given source across DAT areas to provide Police Force or Government Office Region totals. This issue will, however, be less pertinent when prevalence estimates (and associated confidence intervals) are considered.

Data used in the multiple indicator analyses

There is a wide range of indicator data that may be correlated with drug use prevalence at the DAT area level that could be useful within a multiple indicator analysis. Three main types of indicator data were available to be used within this first sweep of the study: data that are currently in the public domain (e.g. published data on crime or income support claimants); data that are not currently in the public domain but have been provided to the study team (e.g. drug-related hospital admissions); and data that have been collected for use within the capture-recapture analyses (such at the NDTMS data). This subsection outlines the data sources that were considered within the MIM analyses (although not necessarily remaining in the best models).

Data sources in the public domain

Health data

There are a number of health-related data sources readily available at the DAT area level. The Office for National Statistics publishes a range of data[8] that can be used as indicators within a problem drug use prevalence multiple indicator analysis. One indicator has been included within the analyses.

- Number of people with a limiting long-term illness (2001 census data).

Crime data

The Home Office publishes data on crimes committed in England by type of crime.[9] The recorded crime statistics include the following categories.

- Burglary
- Criminal damage
- Drug offences
- Fraud and forgery
- Vehicle and other theft.
- Violence against the person
- All offences.

The inclusion of each of these categories of offences has been tried out within the MIM analyses. It should, however, be noted that the drug offences relate to all drugs, not just the drugs included in this study's definition of problem drug use.

Socioeconomic data

The Office for National Statistics also publishes a range of socioeconomic data that may be correlated (at the DAT area level) with the drug use prevalence. The following indicators have been tried within the multiple indicator analysis.

- Number of income support claimants
- Number of incapacity benefit claimants
- Number of jobseekers allowance claimants.

---

8 **http:/./www.statistics.gov.uk** (accessed 20/12/05)
9 **http://www.crimestatistics.org.uk** (accessed 20/12/05)

Population density

The published data on population density in 2001 for each DAT area (derived as total resident population divided by the size of the DAT area in km^2)[10] based on census data were also used as an indicator for the multiple indicator analyses.

Further specific information about all of these indicator data sources (that are in the public domain) is to be found in Appendix 1 of this report.

Data sources not in the public domain

Hospital Episode Statistics data

A request was made to the Department of Health to access patient-level data on hospital admissions related to the use of opiates or cocaine for inclusion within the capture-recapture analyses. Initial analyses found it was not possible to match the hospital data with the other four capture-recapture data sources as the hospital data did not include the patient's initials. These data could, however, when aggregated to the DAT area level, be used as an indicator within the multiple indicator analyses. Hospital episodes were classed as involving opiate use if they included a diagnosis of 'mental and behavioural disorders due to use of opioids'. Furthermore, any episode that referred to either opiate use or cocaine use was also classed as referring to drug injecting if it also had a diagnosis related to drug injecting such as endocarditis, abscesses or cellulitis.

Drug-related deaths

A request was made to the Office for National Statistics to obtain data on drug-related deaths by DAT area. Data were provided by DAT area for 2003; however, for confidentiality reasons, data on DATs where there were less than five drug-related deaths in 2003 were not supplied. Possible solutions to this problem would be to combine 2003 data with the previous years' data or to combine data from several DAT areas. Neither solution was considered satisfactory and therefore the 2003 data were tried within the multiple indicator analyses and where the number of deaths was fewer than five, an average value of two deaths was imputed.

Data sources used within the capture-recapture analyses

The aggregated NDTMS, OASys, PNC and CARAT data which were used in the capture-recapture analyses were all potential data sources for use within multiple indicator analyses. The NDTMS data were used in all multiple indicator analyses and the problem with the undercounting in the North West DATs was compensated for by dividing the aggregated numbers by DAT area by the proportion of records (within that DAT area) that had a valid drug code.

The OAsys, PNC and CARAT data indicators described above were also included in the multiple indicator analyses for problem drug use, opiate use, crack cocaine use and drug injecting. This also made it possible to check whether the data sources used with the capture-recapture analyses were being retained in the final multiple indicator regression model primarily because the capture-recapture estimates were highly correlated with the data they were derived from or whether those potential indicators were correlated with the true underlying prevalence of problem drug use.

For the areas where a problem had been identified within an individual data source (OASys data in Cheshire, Halton and Warrington and PNC data in Suffolk) the ratio between that particular data source and the NDTMS source for the rest of the relevant Government Office Region was used to impute the value to replace the known undercount value.

---

10 http:/.www.statistics.gov.uk (accessed 20/12/05)

Other potential data sources (not used)

Drug treatment data

The National Treatment Agency for Substance Misuse regularly publishes statistics for drug treatment activity in England derived from NDTMS. The most recent published documents containing drug treatment data that could have been used as possible indicators within this study were:

- provisional statistics for drug treatment activity in England 2004/05 from the National Drug Treatment Monitoring System;[11] and
- statistics from the National Drug Treatment Monitoring System, 1 April 2003 – 31 March 2004.[12]

The provisional statistics for 2004/05 provide a more up-to-date summary of drug treatment in England; however, they had not (at the time of the analyses) been broken down into the drugs used or injecting status. There were several reasons why these published sources of treatment data were not used. First, the data broken down by type of drug only refers to the main drug of use rather than including every drug used. This could give a substantial undercount of people using crack cocaine and the extent of the undercount could vary by DAT area. Secondly, in a small number of DAT areas, data are missing from a substantial proportion of records, particularly data related to the injecting status. Thirdly, the most up-to-date data available by drug type at the time of analysis were a year out of date and fourthly, although the 2004/05 data were more up to date, those figures would also include people only using drugs that are outside the remit of this study, e.g. cannabis.

Misuse of Drugs Act 1971 possession convictions

Although data on possession convictions in 2004 had been published by the Home Office, the information was only presented at the Police Force area level. A request was made to the Home Office to provide additional data on possessions by court areas (as information on the offender's area of residence is not included in the data collected). Although the areas covered by courts are not always coterminous with DAT areas it was possible to aggregate the court data to the DAT area in 105 of the 149 DAT areas of England. In the remaining areas it may have been possible to include this source within a specific MIM analysis that combined areas that straddled court boundaries (e.g. by combining Luton with Bedfordshire or combining Lancashire with the other two DAT areas of Blackburn with Darwen and Blackpool). This approach was not developed within this first sweep of the study; rather the PNC data were used instead. The potential benefits of using published data on possession offences may be considered at a later point within the study.

Two-sample capture-recapture estimate indicators

The capture-recapture method was primarily used with four sources (for problem drug use, opiate use and crack cocaine use) or three sources (for drug injecting); however, two-sample capture-recapture estimates can be easily derived for different combinations of data sources. Although not undertaken within this study, the use of two-sample capture-recapture estimates as indicators in a multiple indicator analysis could be explored in the future.

Capture-recapture analyses

In this section the specific application of the capture-recapture method in the context of this study is considered. In simple terms, the capture-recapture analysis involves testing a series of statistical formulae, or 'models', to find one that best matches, or 'fits' the pattern of overlap between data sources. This model is then used to calculate the number of problem drug

---

11 **http://www.nta.nhs.uk/programme/national/docs/200405_nos_in_treatment_by_DAT.xls** (accessed 20/12/05)

12 **http://www.nta.nhs.uk/programme/national/docs/NDTMS_200304_bulletin_Jul_05.**doc (accessed 20/12/05)

users who do not appear in any source. This estimate is then added to the total number of known problem drug users, to provide an overall estimate of prevalence.

The first stage of analysis involved testing how well a simple model, that assumed all samples were independent of each other, matched the observed overlap in the contingency table. Increasingly complex models, representing dependencies between single pairs of data samples ('one-way') and then two pairs of samples ('two-way') were then tested. The model that best matched the overlap was chosen using objective statistical criteria; more complex models were only chosen if they provided a better match (on comparing AIC values) than lower-level models. All capture-recapture analyses were carried out using the GLIM4 statistical package.

Analysis: prevalence of problem drug use (opiate and/or crack cocaine use) and opiate use

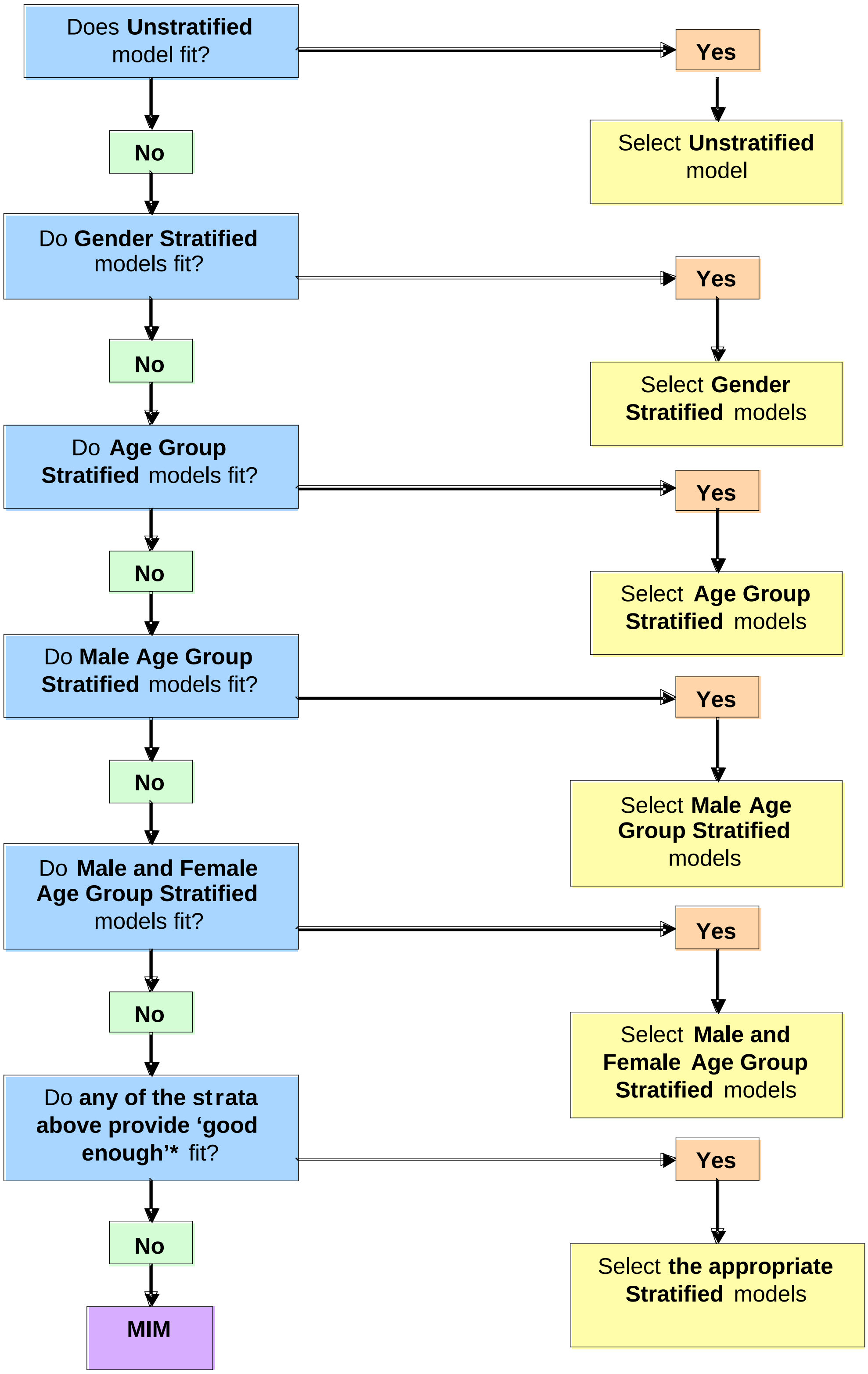
In most DAT areas, all four sources of data were available to estimate the prevalence of problem drug use and opiate use. Attempts were made to produce capture-recapture estimates in all 149 DAT areas but in the two smallest DAT areas there were too few data to carry out this analysis (City of London and Rutland). Furthermore, in the three DAT areas covered by the Cheshire Probation Service (Cheshire, Halton and Warrington) the three-sample method was tried (as OASys data were not available for those DAT areas). Due to difficulties in fitting valid models to the data in the three Cheshire DAT areas, compounded by the issue of missing data in the NDTMS data across the North West, multiple indicator estimates were used instead.

In the first stage of the analysis, the 22 simplest models were applied to the overlap data from each of the remaining 144 DAT areas in England. This was initially carried out on unstratified data, i.e. not splitting by gender or age group. This process was then repeated for the data stratified by age group (three strata) and by gender (two strata) giving five stratified estimates. At this stage the data were not stratified by both the age group and gender (e.g. young males, old females). Such an approach to stratification would have given another six stratified estimates.

Various methods were used to explore whether the model fitted to the unstratified data was a good fit (in particular if the AIC value was less than zero) and whether the resultant estimate was valid. This included checking whether the lowest deviance value indicated a good fit (a lower deviance value signifies a better fit of the model to the observed data), checking whether the estimate derived from applying the best model was similar to a weighted estimate (calculated as a weighted mean of the available 22 estimates) and whether the unstratified estimate was similar to the sum of the stratified estimate for both the age-stratified and gender-stratified model/estimates. In addition, it was considered whether each estimate was credible (i.e. not unfeasibly low or high in comparison with the known drug using population or underlying general population).

A flow diagram that summarises the analysis approach taken in each DAT area for problem drug use or opiate use is provided in Figure 2.2.

Figure 2.2: Flow diagram of the analysis strategy for problem drug use and opiate use



* One stratum does not fit but AIC<5 *and* is close to weighted estimate.

Thus to summarise, if the model fitted to the unstratified data did not offer a valid estimate, then either the summed gender-specific or age group-specific estimates were considered (with gender-specific estimates preferred if there was no discernable difference between the two approaches, again to ensure that the national confidence interval was not excessively wide). Under those circumstances, the best approach was often to stratify the males into three age groups but keep the female data unstratified. This was particularly important as, across the country, there were few data on female problem drug users over the age of 34. If that approach did not work, then the analyses were run on the six age and gender strata and those estimates were considered. If none of those unstratified estimates were deemed to be appropriate then any stratified analysis where the AIC value for one stratum was less than five was considered. If none of those approaches provided a valid estimate then a multiple indicator estimate was used instead.

Analysis: prevalence of crack cocaine use

To a certain extent the approach described above was carried out to estimate the prevalence of crack cocaine use at the DAT area level. However, due to the fewer data on crack cocaine use, only the unstratified analyses were carried out. The crack cocaine estimates were also compared to the problem drug use estimates to ensure consistency and where it was not possible to obtain a valid or feasible crack cocaine estimate using the capture-recapture method then a multiple indicator estimate was used instead.

Analysis: prevalence of drug injecting

As it was not possible to identify drug injectors from police data, only three sources of data were available to estimate the prevalence of drug injecting. It should also be noted that OASys data in many parts of London identified few drug injectors. A similar approach to that taken to obtain crack estimates was taken, i.e. only considering the unstratified estimate and comparing it to the problem drug use estimate for consistency. It should, however, be noted that as the capture-recapture analyses used to estimate the prevalence of drug injecting were three-sample analyses there were fewer models from which to select the prevalence estimate.

Multiple indicator analyses

In this section the specific application of the multiple indicator method within this sweep of the study is considered. All of the indicator data (apart from population density) and the anchor point data were converted to rates per 10,000 population aged from 15 to 64 prior to inclusion in the analyses.

Analysis: prevalence of problem drug use (opiate and/or crack cocaine use)

The capture-recapture analyses derived estimates of the prevalence of opiate use and the prevalence of problem drug use (opiate and/or crack cocaine use). Following the capture-recapture analyses, a set of anchor point DAT areas were constructed for use within the multiple indicator analyses. A DAT area was selected as an anchor point if both the opiate use and problem drug use capture-recapture analyses offered valid prevalence estimates. In total there were 117 DAT areas that were available to be used as anchor points in the multiple indicator analysis. However, there was one DAT area with a problem drug use capture-recapture estimate (Brighton and Hove) that was not used as an anchor point in the problem drug use multiple indicator analyses. This was because inclusion of that DAT area in the MIM analyses would have unduly influenced the analyses.[13]

13 As signified by the large value of the Cook's distance for that anchor point in the regression analysis.

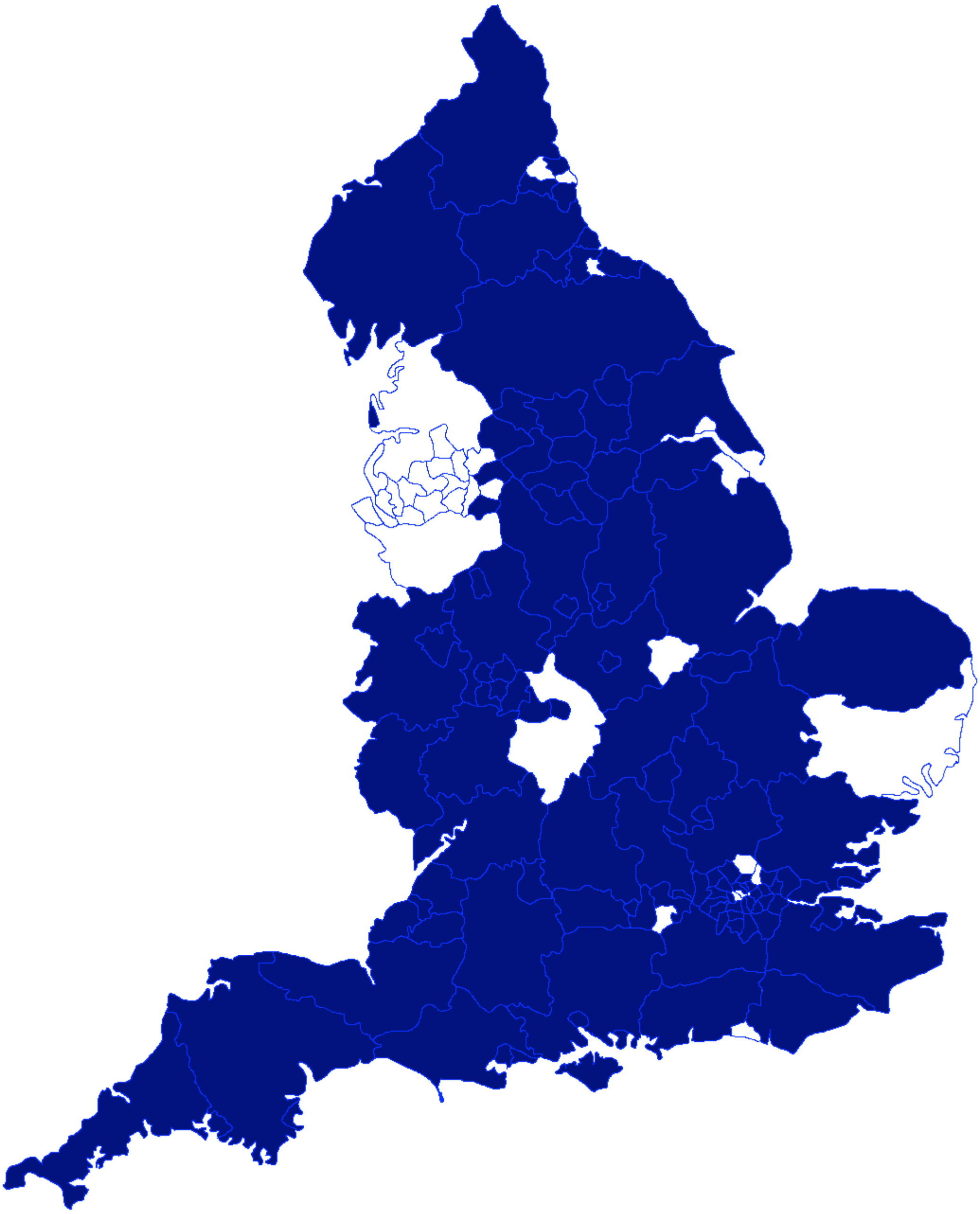
The number of DAT areas that were used as multiple indicator anchor points are summarised by Government Office Region in Table 2.2 and this information is presented, within Government Office Region, in the regional reports which are available on the NTA website.

Table 2.2: Summary of the number of DAT areas used as multiple indicator anchor points by Government Office Region

Government Office Region	Number of DAT Areas	Problem	Opiate	Crack	Injecting
East of England	10	10	10	6	4
East Midlands	9	8	8	8	6
London	33	27	21	15	16
North East	12	9	10	4	5
North West	22	4	4	3	3
South East	19	17	16	12	12
South West	15	15	15	11	12
West Midlands	14	13	14	11	13
Yorkshire and the Humber	15	13	13	12	12
England	**149**	**116**	**111**	**82**	**83**

The DATs that were used as anchor points in the problem drug use multiple indicator analyses are shown as the darker shaded areas in Figure 2.3 (map).

Figure 2.3: Map showing the problem drug use anchor point areas (darker shaded areas)



It should be again noted that, as clearly shown in Figure 2.3, there were data coverage issues in almost all of the DAT areas in the North West and also in Suffolk. None of these areas (nor the small DAT areas of City of London or Rutland) were used as anchor points and where there was known to be undercounting in any particular capture-recapture data source (e.g. NDTMS in the North West, OASys data in Cheshire, Halton or Warrington and PNC data in Suffolk), imputed or corrected values were used for the indicator data.

With well over 100 anchor points available for problem drug use or opiate use (and over 80 anchor points available for crack cocaine use or drug injecting), there was no need to use a technique known as principal component analyses that multiple indicator studies often use to ensure that the number of indicators is effectively less than the number of available anchor points (a prerequisite of the regression analysis); instead, the stepwise regression method in Minitab release 13.30 was used.

The stepwise regression approach considers all available indicators and only includes a particular indicator in the final regression model if it is significantly related to the available prevalence estimates. In contrast to the forward selection approach which starts with no indicators in the model and keeps including indicators until there are no more significant indicators, and the backward elimination approach which starts with all indicators in the model and removes non-significant ones until all remaining ones are significantly related to the available prevalence estimates, the stepwise regression approach alternates at each step between adding significant or deleting non-significant indicators[14] and can result in models that offer a good fit to the available data with a minimal number of indicators. The stepwise regression approach resulted in the following indicators remaining in the best regression model (in order of significance).

- NDTMS
- Population density
- CARAT
- Drug offences.

This model explained 90 per cent of the variance (i.e. provided a good fit to the available data) with the first indicator (NDTMS) explaining 69 per cent of the variance.

As described in the subsection on capture-recapture, the estimates for each age group and by gender group were obtained by applying the ratio of the weighted stratified estimates to the best estimate (no matter if it was derived from summing stratified estimates). For example, if the unstratified estimate suggested that there were 1,000 opiate users in a particular DAT area and a comparison of the two weighted estimates for males and females suggested that 70 per cent were males, then it was assumed that there were 700 male opiate users and 300 female opiate users. A similar approach was taken to get age-stratified estimates.

In the City of London DAT area it was not possible to produce a plausible estimate of problem drug use from the model. Further examination of the data for this area revealed that the estimate was being inflated by the drug offences data. This was due to the combined effects of the very small size of the area and the fact that the drug offence data refers to crimes detected within the DAT area rather than crimes committed by residents of the DAT area. Therefore, a different MIM model, using the only three indicators (population density, CARAT data and NDTMS data), was used in deriving a multiple indicator estimate for the City of London DAT area.

Analysis: prevalence of opiate use, crack cocaine use or drug injecting

The general approach outlined above for problem drug use was also taken to estimate the prevalence of opiate use or crack cocaine use and the prevalence of drug injecting. None of the available capture-recapture estimates was excluded from the multiple indicator analyses due to over-influencing the regression line. As with the problem drug use analyses, crime statistics indicators were not used within the multiple indicator analyses for the City of London DAT area. Although there were fewer anchor points for crack cocaine use or drug injecting, multiple indicator analyses appeared to be successful. The stepwise regression approach resulted in the following indicators remaining in the best regression models (in order of significance) for each definition.

Opiate use

- NDTMS
- Population density
- CARAT
- Drug-related deaths
- Drug offences

---

14 In these analyses α to enter and α to remove were both set to 0.15.

- Income Support claimants
- Burglaries

Crack cocaine use

- PNC
- NDTMS
- Drug offences
- CARAT

Drug injecting

- NDTMS
- CARAT
- Drug-related deaths
- Jobseeker's Allowance claimants

For the opiate use analyses, the indicators again explained 90 per cent of the variance and the crack cocaine and drug injecting analyses explained 84 per cent and 81 per cent respectively.

Confidence intervals

The favoured approach to deriving a 95 per cent confidence interval for a single capture-recapture estimate is the one proposed by Cormack (1992). This approach is preferred as it recognises that the confidence interval for these types of estimates should be asymmetric, or in other words, the higher limit of the confidence interval will be further away from the actual estimate than the lower limit would be. This is a consequence of using log-linear regression models instead of standard linear regression models to derive estimates within a capture-recapture analysis.

Obtaining a confidence interval for a local multiple indicator estimate is more straightforward; however, there are two different intervals that are offered when the regression model is fitted to the available anchor points/indicator data. One of these intervals, often called the confidence interval, is actually the confidence interval for the mean of the regression line. The other interval is known as the prediction interval, and would be the 95 per cent confidence interval for a new estimate derived from fitting the regression line to a set of indicator values for another area. Thus it should be the 95 per cent prediction interval that should be expressed when using multiple indicator methods to provide prevalence estimates for areas where capture-recapture estimates are not available (the non-anchor point areas). Unfortunately, the 95 per cent prediction interval is far wider than the 95 per cent confidence interval for the mean and can result in lower bounds that are lower than the known population or indeed negative lower bounds. When this has happened, the individual confidence interval has been truncated at the size of the known population, or in the case of North West DATs the number of drug users from the NDTMS data after adjusting for cases with missing data (if higher).

In order to get a national confidence interval, the various local confidence intervals must be combined. Unlike the national estimate which is simply the sum of the various local estimates from which it is constructed (i.e. the England estimate will be the sum of the nine Government Office Region estimates), the confidence intervals cannot be obtained by summing (for example) the lower limits of the local confidence intervals to get the lower limit of the national confidence interval. To do so would give an extremely wide and totally unrealistic confidence interval.

The favoured approach to combining confidence intervals, as carried out in the Home Office feasibility study that estimated the prevalence of problem drug use in Greater Manchester (Millar *et al*., 2004) and in similar studies in Scotland (Hay *et al*., 2005), is to use a computer-intensive set of techniques called 'bootstrap methods'. In this approach, all of the local confidence intervals are approximated by an assumed statistical distribution (the log-normal in the case of capture-recapture estimates and the normal in the case of multiple indicator estimates). Where there were negative lower bounds in the multiple indicator estimates they were kept as calculated as the bootstrap method could deal with that. In the small number of instances that the lower bound of the capture-recapture hidden population was zero then the value one was substituted to enable the log value to be calculated. A large number of different values from each of the distributions are randomly simulated (typically about 1,000 values) and then these distributions are summed. The confidence interval of the national estimate can then be approximated to the 2.5 per cent and 97.5 per cent percentile of that distribution.

Age group and gender specific estimates

The main estimate for each DAT area, referring to the prevalence in the 15 to 64 age range for males and females combined, was derived using two different methods (capture-recapture and multiple indicator) and where the capture-recapture method was used, different approaches to stratification were employed. In some of the areas where the estimates were based on capture-recapture estimates, some were derived as the sum of the gender specific estimates, some were derived as the sum of the age group specific estimates and others were either the non-stratified estimates or combinations of other stratified estimates. To obtain comparable age group or gender specific estimates for all DAT areas, for PDU (opiate and/or crack cocaine use) and 'opiate use' particularly, including those where the multiple indicator method was used, the following approach was taken.

Apart from the City of London and Rutland DAT areas (where there were too few data to run any capture-recapture analyses) or the Cheshire, Halton and Warrington DAT areas (where OASys data were not available) a weighted capture-recapture estimate was derived for each stratum (males, females, 15 to 24, 25 to 34 and 35 to 64). A weighted capture-recapture estimate essentially captures all of the information in the 22 individual models fitted to the available data in each of the individual (stratified) capture-recapture analyses and can offer a good estimate, even when individual point estimates markedly vary across the 22 models or when there is difficulty fitting a model. The proportion of problem drug users in each stratum is then estimated using the weighted estimates. For example, the proportion of female drug users in the Barking and Dagenham DAT area is calculated as the female weighted capture-recapture estimate divided by the sum of the female weighted capture-recapture estimate and the male weighted capture-recapture estimate for that area.

This is a relatively new approach which was used in order to provide comparable age group or gender-specific estimates even for areas where the capture-recapture method did not offer useful estimates. It is an approach that needs further consideration and development and therefore the gender and age-group estimates should be viewed with caution. This will be addressed further in future sweeps of the study particularly in terms of obtaining appropriate confidence intervals for the stratified estimates.

Results

In this section the prevalence estimates for problem drug use (opiate and/or crack cocaine use), opiate use, crack cocaine use and drug injecting are presented, aggregated to the national and Government Office Region level. The estimates by individual DAT area are given in the Regional reports which are available on the NTA website.

Prevalence estimates

In total there were 117 areas where the capture-recapture analyses offered valid estimates of the prevalence of problem drug use. In those areas the prevalence of problem drug use was provided by the capture-recapture estimate whereas in the remaining 32 areas the multiple indicator estimates were used. There were 111 areas that had capture-recapture estimates for the opiate use estimate. In terms of crack cocaine use, 82 areas had capture-recapture estimates, and 83 areas had a capture-recapture estimate for drug injecting. The decision to use a capture-recapture estimate instead of a multiple indicator method was always taken on the basis of the validity of the capture-recapture estimate, both in terms of how well the capture-recapture model fitted the available data and how feasible the estimate was compared to the known population and the estimates for other drugs.

There are an estimated 327,466 problem drug users (defined as 'opiate and/or crack cocaine users'),in England, (95% CI 325,945 to 343,424). This corresponds to 9.93 per thousand population aged from 15 to 64 (95% CI 9.88-10.41). In terms of opiate users, there are an estimated 281,320 people in England who use those drugs (8.53 per thousand population aged from 15 to 64) whereas it is estimated that 192,999 people use crack cocaine (5.85 per thousand population aged from 15 to 64). It should be noted that the majority of people using crack cocaine are also using opiates and that crack cocaine may neither be their main drug of use or indeed the drug that is causing them the most problems. Finally, there are an estimated 137,141 opiate and/or crack cocaine users in England who inject drugs (4.16 per thousand of the population aged from 15 to 64).

Table 2.3 summarises the national prevalence estimates along with their associated confidence interval.

Table 2.3: National prevalence estimates and rates per thousand aged from 15 to 64 with 95 per cent confidence intervals

Drug	Estimate	95% confidence interval	Rate	95% confidence interval
Problem	327,466	325,945 – 343,424	9.93	9.88 –10.41
Opiate	281,320	279,753 – 292,941	8.53	8.48 – 8.88
Crack	192,999	188,138 – 210,763	5.85	5.70 – 6.39
Injecting	137,141	133,118 – 149,144	4.16	4.04 – 4 .52

Table 2.4 presents the prevalence estimates by Government Office Region for problem drug use, opiate use, crack cocaine use and drug injecting. Table 2.5 presents the prevalence estimates per thousand of the population aged from 15 to 64. These estimates are presented together with their 95 per cent confidence intervals in Tables 2.6 to 2.9.

Table 2.4: Estimated number of problem drug (opiate and/or crack cocaine) users, opiate users, crack cocaine users and drug injectors by Government Office Region

Government Office Region	Problem	Opiate	Crack	Injectors
East of England	23,081	19,518	14,102	9,418
East Midlands	23,142	21,241	12,952	11,796
London	74,417	55,139	51,312	17,909
North East	15,853	13,729	7,376	8,959
North West	51,110	43,996	29,752	22,089
South East	33,854	28,727	19,261	13,778
South West	30,455	27,541	17,552	17,444
West Midlands	36,834	34,661	20,827	14,734
Yorkshire and the Humber	38,720	36,768	19,865	21,014
England	**327,466**	**281,320**	**192,999**	**137,141**

Table 2.5: Estimated prevalence of problem drug (opiate and/or crack cocaine) use, opiate use, crack cocaine use and drug injecting by Government Office Region (per thousand population aged from 15 to 64)

Government Office Region	Problem	Opiate	Crack	Injecting
East of England	6.48	5.48	3.96	2.64
East Midlands	8.23	7.55	4.60	4.19
London	14.35	10.64	9.90	3.45
North East	9.50	8.22	4.42	5.37
North West	11.43	9.84	6.65	4.94
South East	6.40	5.43	3.64	2.60
South West	9.44	8.54	5.44	5.41
West Midlands	10.62	9.99	6.00	4.25
Yorkshire and the Humber	11.74	11.15	6.02	6.37
England	**9.93**	**8.53**	**5.85**	**4.16**

In terms of regional differences, London is the Government Office Region with the largest prevalence of problem drug use at around 14 per thousand population aged from 15 to 64 compared with Yorkshire and the Humber at around 12 per thousand and the North West and the West Midlands at around 11 per thousand. The South East and the East of England have the lowest prevalence of problem drug use at around six per thousand. When considering opiate use prevalence, both London and Yorkshire and the Humber have around 11 per thousand and the North West and West Midlands have around ten per thousand. The lowest prevalence of opiate use is in the South East and East of England at around five per thousand. London has the highest estimated prevalence of crack cocaine use at just under ten per thousand population compared to a prevalence of between four and seven per thousand in the other regions. Injecting drug use prevalence ranges from around six per thousand in Yorkshire and the Humber to around three per thousand in the South East, East of England and London.

Individual DAT area estimates are presented the regional reports which are available on the NTA website. These show variations in prevalence of problem drug use across the country and also variations within Government Office Regions.

Confidence intervals

Confidence intervals for each DAT level estimate have been produced and these are presented in the relevant tables in the regional reports which are available on the NTA website. Tables 2.6 and 2.7 present the estimates and confidence intervals at the Government Office Region level for problem drug use, opiate use, crack cocaine use and drug injecting. Tables 2.8 and 2.9 present the information as estimated prevalence rates per thousand population aged from 15 to 64.

Table 2.6: Prevalence estimates and 95 per cent confidence intervals (problem drug (opiate and/or crack cocaine) use and opiate use), by Government Office Region

Government Office Region	Problem drug users			Opiate users		
	Estimate	95% CI		Estimate	95% CI	
East of England	23,081	22,029	25,313	19,518	18,539	21,558
East Midlands	23,142	22,516	25,442	21,241	20,815	22,813
London	74,417	71,845	81,299	55,139	52,611	59,799
North East	15,853	14,987	17,683	13,729	13,099	15,657
North West	51,110	46,455	55,659	43,996	40,737	48,221
South East	33,854	32,495	38,306	28,727	27,267	31,648
South West	30,455	29,536	32,936	27,541	26,681	29,865
West Midlands	36,834	35,276	39,726	34,661	33,656	37,173
Yorkshire and the Humber	38,720	37,708	40,911	36,768	35,674	39,028
England	**327,466**	**325,945**	**343,424**	**281,320**	**279,753**	**292,941**

Table 2.7: Prevalence estimates and 95 per cent confidence intervals (crack cocaine use and drug injecting), by Government Office Region

Government Office Region	Crack users			Injectors		
	Estimate	95% CI		Estimate	95% CI	
East of England	14,102	11,779	22,425	9,418	6,252	13,061
East Midlands	12,952	11,417	16,627	11,796	10,487	13,518
London	51,312	48,156	55,850	17,909	16,161	24,002
North East	7,376	5,636	10,899	8,959	7,638	10,597
North West	29,752	25,533	34,423	22,089	18,781	25,204
South East	19,261	16,097	26,554	13,778	12,044	17,802
South West	17,552	15,509	20,785	17,444	15,930	19,532
West Midlands	20,827	19,012	26,593	14,734	13,589	17,007
Yorkshire and the Humber	19,865	18,239	23,095	21,014	19,860	22,788
England	**192,999**	**188,138**	**210,763**	**137,141**	**133,118**	**149,144**

Table 2.8: Estimated prevalence (rate per 1,000 population aged from 15 to 64) of problem drug (opiate and/or crack cocaine) use and opiate use by Government Office Region including 95 per cent confidence intervals

Government Office Region	**Problem drug use**			**Opiate use**		
	Estimate	**95% CI**		**Estimate**	**95% CI**	
East of England	6.48	6.18	7.11	5.48	5.20	6.05
East Midlands	8.23	8.00	9.04	7.55	7.40	8.11
London	14.35	13.86	15.68	10.64	10.15	11.53
North East	9.50	8.98	10.59	8.22	7.85	9.38
North West	11.43	10.39	12.45	9.84	9.11	10.79
South East	6.40	6.14	7.24	5.43	5.15	5.98
South West	9.44	9.16	10.21	8.54	8.27	9.26
West Midlands	10.62	10.17	11.45	9.99	9.70	10.71
Yorkshire and the Humber	11.74	11.43	12.40	11.15	10.81	11.83
England	**9.93**	**9.88**	**10.41**	**8.53**	**8.48**	**8.88**

Table 2.9: Estimated prevalence (rate per 1,000 population aged 15 to 64) of crack cocaine use and drug injecting by Government Office Region including 95 per cent confidence intervals

Government Office Region	**Crack users**			**Injectors**		
	Estimate	**95% CI**		**Estimate**	**95% CI**	
East of England	3.96	3.31	6.30	2.64	1.76	3.67
East Midlands	4.60	4.06	5.91	4.19	3.73	4.81
London	9.90	9.29	10.77	3.45	3.12	4.63
North East	4.42	3.38	6.53	5.37	4.57	6.35
North West	6.65	5.71	7.70	4.94	4.20	5.64
South East	3.64	3.04	5.02	2.60	2.28	3.36
South West	5.44	4.81	6.44	5.41	4.94	6.05
West Midlands	6.00	5.48	7.66	4.25	3.92	4.90
Yorkshire and the Humber	6.02	5.53	7.00	6.37	6.02	6.91
England	**5.85**	**5.70**	**6.39**	**4.16**	**4.04**	**4.52**

Stratified prevalence estimates

This section provides estimates for problem drug use and opiate use broken down by gender and by age-group. Although, as previously described, the method for producing these stratified estimates is relatively new and is still being developed, it is possible to make broad comparisons by age and gender at the regional level. It should be noted that there was insufficient information to obtain age or gender-stratified estimates for crack cocaine use or drug injecting.

Tables 2.10 and 2.11 present the estimates for the Government Office Regions for both problem drug use and opiate use broken down by gender. Again, DAT level information is presented in the regional reports which are available on the NTA website.

Table 2.10: Estimated gender breakdown for problem drug (opiate and/or crack cocaine) use by Government Office Region

Government Office Region	Female	%	Male	%
East of England	7,222	31.29	15,859	68.71
East Midlands	5,548	23.97	17,594	76.03
London	17,359	23.33	57,059	76.67
North East	4,247	26.79	11,606	73.21
North West	13,152	25.73	37,958	74.27
South East	9,185	27.13	24,669	72.87
South West	8,133	26.70	22,322	73.30
West Midlands	8,046	21.84	28,788	78.16
Yorkshire and the Humber	10,224	26.40	28,496	73.60
England	**83,116**	**25.38**	**244,351**	**74.62**

Table 2.11: Estimated gender breakdown for opiate use by Government Office Region

Government Office Region	Female	%	Male	%
East of England	5,878	30.12	13,640	69.88
East Midlands	5,263	24.78	15,978	75.22
London	13,588	24.63	41,581	75.37
North East	3,678	26.79	10,051	73.21
North West	12,009	27.30	31,987	72.70
South East	7,926	27.59	20,801	72.41
South West	7,266	26.38	20,275	73.62
West Midlands	7,443	21.47	27,218	78.53
Yorkshire and the Humber	9,661	26.28	27,107	73.72
England	**72,712**	**25.84**	**208,638**	**74.16**

As shown in Tables 2.10 and 2.11, approximately a quarter of problem drug users or opiate users in England are female.

The breakdown by age group was also examined, using the three age groups of those aged from 15 to 24, those aged from 25 to 34 and those aged from 35 to 64. These breakdowns are given in Tables 2.12 and 2.13.

Table 2.12: Estimated age group breakdown for problem drug (opiate and/or crack cocaine) use by Government Office Region

Government Office Region	15 to 24 years	%	25 to 34 years	%	35 to 64 years	%
East of England	5,143	22.28	10,284	44.56	7,652	33.15
East Midlands	6,959	30.07	10,458	45.19	5,726	24.74
London	14,068	18.91	25,509	34.28	34,837	46.82
North East	4,979	31.41	7,237	45.65	3,637	22.94
North West	8,264	16.17	22,552	44.12	20,290	39.70
South East	7,235	21.37	14,714	43.46	11,902	35.16
South West	6,112	20.07	14,325	47.04	10,020	32.90
West Midlands	10,747	29.18	15,753	42.77	10,332	28.05
Yorkshire and the Humber	9,284	23.98	19,444	50.22	9,993	25.81
England	**72,838**	**22.23**	**140,365**	**42.84**	**114,459**	**34.93**

Table 2.13: Estimated age group breakdown for opiate use by Government Office Region

Government Office Region	15 to 24 years	%	25 to 34 years	%	35 to 64 years	%
East of England	4,038	20.69	8,804	45.11	6,676	34.20
East Midlands	6,460	30.41	9,849	46.37	4,931	23.21
London	8,927	16.19	18,727	33.96	27,487	49.85
North East	4,105	29.90	6,526	47.53	3,100	22.58
North West	6,414	14.58	20,052	45.58	17,531	39.85
South East	5,416	18.85	13,035	45.38	10,280	35.79
South West	5,018	18.22	13,036	47.33	9,489	34.45
West Midlands	10,699	30.87	15,373	44.35	8,590	24.78
Yorkshire and the Humber	8,506	23.13	18,602	50.59	9,656	26.26
England	**59,583**	**21.18**	**124,004**	**44.08**	**97,740**	**34.74**

From Tables 2.12 and 2.13, the younger age group accounts for around 20 per cent of the estimated population. Despite being the largest age group in terms of baseline population, the 35-64 age group accounts for approximately 35 per cent of the total estimate. It is clear from Tables 2.12 and 2.13 that most problem drug users in England are aged between 25 and 34.

There are some marked regional differences in terms of the proportions of problem drug users who are in each of the age groups. In London and the North West the percentage of problem drug users in the 15 to 24 age group is lower than 20 per cent but in East Midlands, West Midlands and the North East around 30 per cent are in the youngest age group. There is a similar pattern for opiate use. In London around 47 per cent of problem drug users are in the 35 to 64 age group in contrast to the North East where only 23 per cent are in this age group.

The gender and age-group specific estimates can also be considered in terms of prevalence rates as presented in Tables 2.14 to 2.17.

Table 2.14: Problem drug (opiate and/or crack cocaine) use prevalence rates per thousand population aged from 15 to 64, by gender and Government Office Region

Government Office Region	**Female**	**Male**
East of England	3.95	9.15
East Midlands	3.80	13.02
London	6.11	24.33
North East	4.93	14.37
North West	5.66	17.68
South East	3.39	9.56
South West	4.85	14.40
West Midlands	4.52	17.02
Yorkshire and the Humber	5.97	17.96
England	**4.84**	**15.47**

Table 2.15: Opiate use prevalence rates per thousand population aged from 15 to 64, by gender and Government Office Region

Government Office Region	**Female**	**Male**
East of England	3.21	7.87
East Midlands	3.60	11.83
London	4.79	17.73
North East	4.27	12.44
North West	5.17	14.90
South East	2.92	8.06
South West	4.34	13.08
West Midlands	4.18	16.09
Yorkshire and the Humber	5.64	17.09
England	**4.23**	**13.21**

Table 2.16: Problem drug (opiate and/or crack cocaine) use prevalence rates per thousand population by age group and Government Office Region

Government Office Region	**15 to 24 years**	**25 to 34 years**	**35 to 64 years**
East of England	7.91	14.81	3.45
East Midlands	12.56	19.83	3.31
London	14.22	17.54	12.71
North East	14.44	24.08	3.55
North West	9.12	26.41	7.48
South East	7.26	14.38	3.64
South West	10.05	24.70	4.92
West Midlands	15.42	23.45	4.92
Yorkshire and the Humber	13.53	31.06	5.03
England	11.32	20.85	5.77

Table 2.17: Opiate use prevalence rates per thousand population by age group and Government Office Region

Government Office Region	**15 to 24 years**	**25 to 34 years**	**35 to 64 years**
East of England	6.21	12.68	3.01
East Midlands	11.66	18.67	2.85
London	9.02	12.87	10.03
North East	11.91	21.71	3.03
North West	7.08	23.48	6.47
South East	5.43	12.74	3.14
South West	8.25	22.48	4.66
West Midlands	15.35	22.88	4.09
Yorkshire and the Humber	12.40	29.72	4.86
England	9.27	18.42	4.93

Ethnicity

The study also aimed to investigate the potential for estimating the prevalence of problem drug use by ethnic group. There are many important considerations when attempting to use prevalence estimation methods to estimate ethnic minority drug use. First of all, the likelihood of obtaining a valid prevalence estimate increases when the size of the known population increases. Thus it may be easier to obtain estimates in larger DAT areas (which are not necessarily DAT areas with a larger proportion of their minority ethnic communities using drugs). Thus the results of a study such as this may appear misleading if it is assumed that those areas that did not receive a prevalence estimate are the ones that do not have any, or very few, ethnic minority drug users.

Secondly, the completeness of the available data depends heavily on how comprehensively and accurately each source monitors ethnicity (and the approach to monitoring ethnicity). In the NDTMS data, the OASys data and the CARAT data it is thought that the ethnic group of each individual is that offered by that individual (however, some research has suggested that some of those undertaking ethnic monitoring are uncomfortable asking someone their ethnic origin (Fountain *et al*., 2003)). It is thought that the PNC data contains the ethnic group assigned by police officer. Little is known about variations in the completeness of ethnic monitoring across the country and differences in this completeness serve to make comparisons across DAT areas difficult.

Thirdly, the ethnic grouping used by each contributing data source differs and can be unclear. From the different classifications used, there are two ethnic groups that appear to be common across groups – Asian and Black. It is thought that Asian more accurately refers to those from the South Asian[15] ethnic group as there is also a Chinese ethnic group category, although this cannot be verified for all sources. Thus in this study two ethnic groups have been considered – Asian and Black.

It was found that the confidence intervals attached to the ethnic minority estimates were often particularly wide and therefore they were of limited use for inferring the proportion of problem drug users within any particular DAT area that are from an ethnic minority group. Estimates by ethnic group have not been presented in this chapter as further work is needed on this and a report addressing the issues described will be produced at a later stage in the study.

Comparing estimates

The case definitions employed within this study include any use of opiates or crack cocaine and not necessarily frequent or primary use of either of those drugs. The case definitions for each estimate are such that it is possible to express the estimated number of opiate users, crack cocaine users or drug injectors at each area level as an estimated proportion or percentage of the number of problem drug users in that area. In other words, since both crack cocaine and opiate use is part of the problem drug use case definition it is appropriate to look at the percentage of problem drug users that are using crack cocaine or opiates. As the drug injecting case definition was restricted to those problem drug users (defined as opiate and/or crack cocaine users) who inject drugs, it is also possible to estimate the proportion of problem drug users who inject (although consideration of such a proportion should recognise the different propensity of opiate users or crack cocaine users to inject drugs).

While it is straightforward to derive an estimate for the percentage of problem drug users who use crack cocaine, use opiates or inject, it is far less straightforward to derive an appropriate confidence interval. This is due to the fact that the figures needed to derive such a percentage (e.g. the number of problem drug users and the number of crack cocaine users) are both estimates and as such both have associated confidence intervals. Thus bootstrap methods need to be employed to construct appropriate upper and lower bounds for the confidence intervals.

While it is possible to derive an appropriate confidence interval for any of the percentages, such as those presented below, another issue related to the fact that both the numerator and denominator of the percentage are estimates is that the confidence intervals can be quite wide. Indeed it is possible that the upper bound of any percentage is greater than 100 per cent or even that the lower bound is lower than zero per cent Thus the upper or lower bounds of any confidence interval for a percentage needs to be truncated at zero per cent or 100 per cent. This is not an issue for the national or Government Office Region estimated percentages, but can often be an issue for the DAT area estimates.

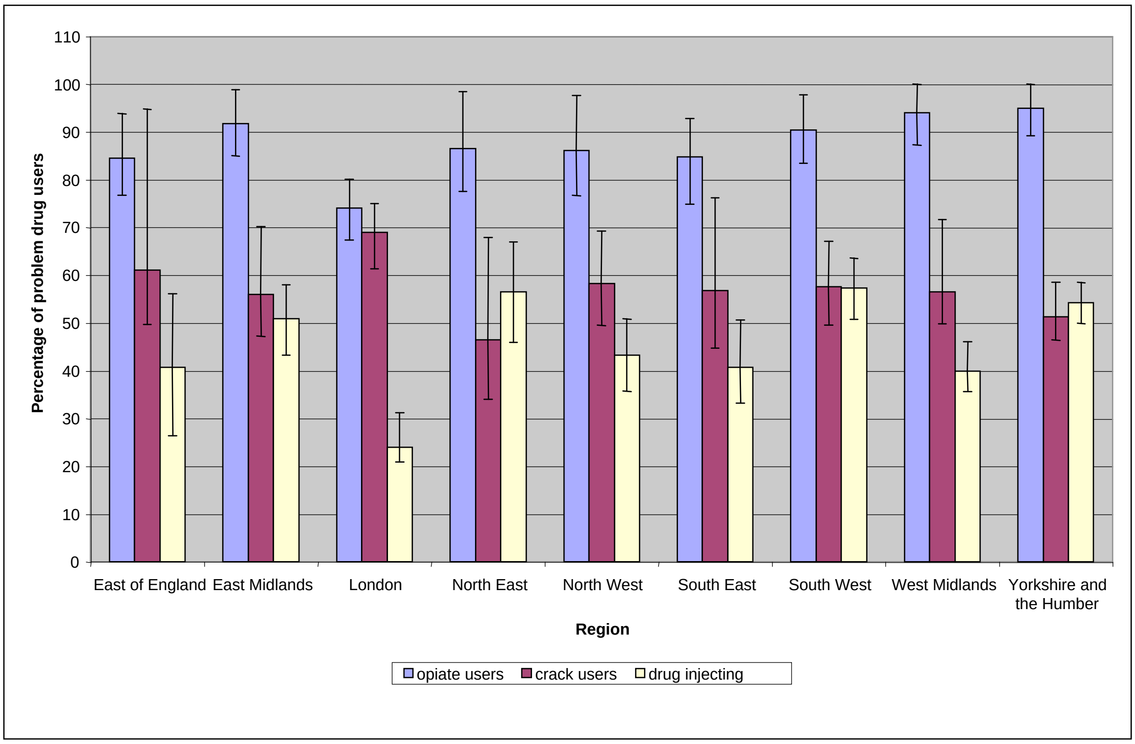
Figure 2.4 presents the percentage of problem drug (opiate and/or crack cocaine) users in each Government Office Region that are estimated to be opiate users, crack users and drug injectors.

Figure 2.4 shows that the percentage of problem drug users who are estimated to be opiate users is quite high, with London being the Government Office Region with the lowest percentage of problem drug users who use opiates at 74.09 per cent (95% CI 67.38% –

15 Bangladeshi, Indian, Pakistani or Sri Lankan.

80.11%). London has the highest percentage of problem drug users estimated to be using crack cocaine at 68.95 per cent (95% CI 61.40% – 75.00%). This contrasts with the North East, which has the lowest estimated percentage of percentage of problem drug users, using crack cocaine at 46.53 per cent (95% CI 34.12% – 67.87%). In terms of drug injecting, London is the Government Office Region with the lowest percentage of problem drug users who are estimated to be drug injectors at 24.07 per cent (95% CI 20.93% – 31.28%). This may be due to the comparatively high level of crack cocaine use in that region.

Figure 2.4: Graph to show estimated percentage of problem drug users who are opiate users, crack cocaine users and injecting drug users, (with associated confidence intervals) by Government Office Region



Discussion and conclusions

This study is the first that has systematically applied the capture-recapture method to such a large number of DAT areas in England. Not only have capture-recapture estimates of the prevalence of problem drug use been obtained for 117 out of the 149 DAT areas, but opiate use prevalence estimates have been obtained in 111 DAT areas. Even for crack cocaine use or drug injecting, there are over 80 capture-recapture estimates each. Thus approximately 80 per cent of DATs in England can base their policy and planning decisions on capture-recapture estimates that have been specifically derived from data relating to their area. Unfortunately, there was an issue with the drug treatment data quality in the North West Government Office Region; outside of that area over 90 per cent of DATs will get a problem drug use capture-recapture estimate.

Due to the large number of capture-recapture estimates derived during the study, there was a large number and wide range of DAT areas which could be used as anchor points for the MIM estimates. Therefore, it is likely that the multiple indicator estimates accurately reflect the true problem drug use prevalence in those DATs that did not receive capture-recapture estimates and thus the national prevalence will be the most accurate ever produced for England.

It should, however, be recognised that neither method can be considered as an 'off the shelf' tool that can be straightforwardly applied to available data to offer prevalence estimates.

Although both approaches, in isolation, can be considered as fairly standard statistical methodologies, the entire process from the initial approach to gaining approval to collate data through to obtaining estimates and confidence intervals is a complex system involving a large number of decisions at various stages. Some of the decisions have been described in previous sections; however, some of the more important methodological issues are discussed here in more detail.

Methodological issues

Matching

The first main methodological issue faced when carrying out a capture-recapture study is how to match within data sources to avoid double counting and to match across data sources to identify the overlap pattern on which the statistical models are based. The desire for as much personal information as possible needs to be balanced against the very valid need to protect an individual's personal data. Thus capture-recapture studies have typically used people's initials, dates of birth, gender and postcode area of residence to match. Matching can be complicated further when data are encrypted, as was a requirement for this study. If, however, the encryption process follows clear rules, then it is still possible to consider different matching algorithms.

In this study a match was taken to be when both initials, along with the dates of birth and gender were all the same. This 'hard' matching would therefore be insensitive to people changing surnames (e.g. married women) or names such as Bill/William. If, however, a reduced set of indicators were used to match then it would increase the risk that false matches occur that could artificially deflate the size of the resultant estimates. In order to explore the possibility of using hospital episode data within a capture-recapture analysis, matching only on dates of birth and gender was examined, but it was found that this substantially decreased the size of the estimates, sometimes even halving them. While using a 'soft' matching process where the forename initials could be different also had the effect of decreasing the prevalence estimates; the decrease was not as pronounced and appeared to be within the confidence intervals of the local estimates (although probably not within the confidence interval of the national estimates). The main reason that 'hard' matching was used in this study is that: it appears to be the matching algorithm applied in previous capture-recapture studies in England; it is similar to the approach taken within the NDTMS data; and it was felt that that basing the first in a series of three annual estimates on hard matching would mean that trend data would be less influenced by any improvements in data collection throughout the three-year period. Nonetheless, it is possible that the prevalence estimates derived after using hard matching are over-estimates. Compared to soft matching, hard matching reduces the overlap between data sources and, in general, increases the prevalence estimates. However, it is then possible that a different model would provide the best fit to the overlap pattern and this different model may offer a lower estimate.

Any possible inflation of prevalence estimates due to an over-strict matching algorithm does need to be considered along with other possible violations of the capture-recapture assumptions, some of which could result in under-estimation. It would therefore be wrong to focus only on the matching assumption in isolation and conclude that the prevalence estimates are overestimates. The effect of using hard matching, as opposed to matching algorithms that allow some identifiers to differ, will be investigated as part of the next phase of the study.

Stratification/model selection

Other issues arise when considering how to stratify the local estimates: should estimates be stratified as far as possible within a DAT area to give as much information as possible, or does the resulting comparatively larger confidence interval make over-stratified data worthless? Also, when summing local estimates to get regional or national estimates, there are benefits in keeping the number of stratified estimates low in order that the larger area estimates do not have overlarge confidence intervals.

Questions should, however, be asked about the comparability of DAT areas where some capture-recapture estimates refer to large county areas (such as Kent with a population aged from 15 to 64 of 870,200) when others refer to relatively small urban unitary authority areas (such as Hartlepool with a population aged from 15 to 64 of 57,600) In this, the most extreme comparison, the population of the smaller DAT area is less than seven per cent of the population of the larger one. There may have been merit in stratifying large areas into geographical sub-areas to see if geographical heterogeneity was present. There did appear to be an issue in the analyses that it was harder to fit models to the data for large areas, and it may not just be due to geographical issues. It could be that capture-recapture methods (or traditional methods for gauging the goodness of fit of capture-recapture models) are struggling to cope with the large numbers seen in some stratified analyses. This issue will be explored in greater detail as part of the study.

Confidence intervals

The approach to producing confidence intervals taken in this study is similar to that taken in other recent studies funded by the Home Office and by the Scottish Executive, but the issue as to whether using bootstrap approximations to the underlying distributions is appropriate has still to be fully considered. While it would clearly be wrong and totally worthless just to sum the lower or upper bounds, there is a sense that rigidly applying the bootstrap method will result in a national confidence interval that masks the very real issue that there are various degrees of uncertainty introduced at each stage of the process and that individual capture-recapture or multiple indicator confidence intervals do not adequately grasp this uncertainty. One obvious example is that the capture-recapture confidence interval is strictly only valid if the model does, indeed, point to the true prevalence in that area, but methods that account for this flaw (such as weighted estimates/confidence intervals) have yet to find favour in drug misuse prevalence estimation. Another example, which is perhaps more important for this study, is that the multiple indicator confidence intervals ignore the uncertainty in the underlying capture-recapture (i.e. the anchor point) confidence intervals. Again, techniques to address this could be developed, such as using weighted regression where the weights are developed to reflect how well the capture-recapture estimate fits the available data.

Multiple indicator methods

The multiple indicator method has been developed within the area of drug misuse prevalence estimation typically to provide estimates from a relatively small number of anchor points. For example, the previous Home Office multiple indicator study, which was groundbreaking at the time, only had 13 anchor points to estimate problem drug use prevalence in 149 areas. This study had a far larger number of anchor point estimates available and therefore a technique such as principal component analysis was not needed. Instead the stepwise regression approach was used and, in general, this appeared successful in that the final regression models included a relatively small number of indicators. There is, however, scope for exploring other regression techniques such as weighted regression which could include some measure of how well the capture-recapture anchor point estimates fitted the available data.

Other issues

The capture-recapture method and the multiple indicator method are still seen as separate approaches to estimate prevalence; however, there should be scope for them to inform each other more than just the capture-recapture estimates slotting into the multiple indicator analyses. To a certain extent this has been done in this study by briefly comparing the multiple indicator estimates with the capture-recapture estimates and if there were any glaring discrepancies these were further examined. However, more formal approaches could, and perhaps should, be taken. For example, a formal re-examination of the capture-recapture estimates in the areas that were deemed to be outliers in the multiple indicator analysis to see if there is perhaps another estimate that is more similar (but still a good enough fit) to that obtained from the multiple indicator approach. This, however, does end up becoming a repetitive, iterative process in that once capture-recapture estimate changes, the multiple indicator analyses would have to be entirely rerun, something that is relatively time-consuming and which could then potentially produce new outliers on each run. Similarly, there may be merit in specifically focusing more on the DAT areas at the extremes of the indicator spectrum as these have more impact on the regression line (and thus the prevalence estimates) and improving a capture-recapture estimate for such an area. The study has looked at four different case definitions but, to a certain extent, has treated them as distinct analyses. Given the case definitions, the number of drug injectors (or indeed the number of opiate users or crack cocaine users) in any DAT area has to be lower than the number of problem drug users in that area, yet there has been no attempt to recognise this within the application of the two methodologies used in the first sweep of the study. Related to this, there may also be scope to further improve both methods by the use of Bayesian techniques, for example by using the multiple indicator estimates as prior weights in the capture-recapture recapture analyses,

The first sweep of the study has been successful in producing a large number of capture-recapture estimates for four different case definitions of problem drug use and thus the most comprehensive and accurate assessment of opiate and/or crack cocaine use, opiate use, crack cocaine use and drug injecting that there has even been. Prior to the next sweep of the study, and in tandem with the very important strand of the study that aims to engage with DATs in interpreting and using the prevalence estimates, there is a rare opportunity to undertake more methodological analyses and further improve the methods.

3. The economic and social costs of Class A drug use in England and Wales, 2003/04

Lorna Gordon[16], Louise Tinsley[16], Christine Godfrey[17], Steve Parrott[17]

This chapter provides an update to the estimates of the costs of Class A drug use published in 2002, in Home Office Research Study (HORS) 249 (Godfrey et *al.*, 2002). Estimates for 2003/04 have now been produced using more recently available data and an improved methodology.

Due to methodological and data improvements, the results for 2003/04 are not comparable to those for 2000.

Box 3.1: Key points

- Based on the methodological and data updates, the economic and social costs of Class A drug use are estimated to be around £15.4 billion in 2003/04. This equates to £44,231 per year per problematic drug user. The associated confidence range is between £15.3 billion and £16.1 billion.
- Problematic Class A drug use accounts for most of the total costs (99%, or £15.3 billion). Viewed from a different perspective, drug-related crime is the domain that accounts for the largest proportion of cost (90%, or £13.9 billion).
- Since the original estimates were first published in 2000, a number of methodological and data improvements have been made to estimate the economic and social costs of Class A drug use. Updated estimates for 2003/04 are based on this improved methodology and more recently available data.
- Methodological and data improvements relate to the inclusion of additional health harms, criminal justice system costs of drug-related crime, victim costs of drug-related crime and the estimates of the prevalence of problematic drug use. Given these changes, total cost estimates in 2000 and 2003/04 are not comparable.

Background

The estimates of economic and social costs of Class A drug use provide a measure of the total costs to society of Class A drug use. The methodology identifies Class A drug users by type – young recreational, older regular and problematic – and prevalence estimates for each are derived. The consequences of Class A drug use are identified as drug-related crime,

[16] Home Office, 2 Marsham Street, 3rd Floor, Seacole Building, SW1P 4DF
[17] Department of Health Sciences, Area 2, 1st Floor, Seebohm Rowntree Building, University of York, Heslington, York, UK YO10 5DD
Address for correspondence: Lorna Gordon, Economics and Resource Analysis Group, Home Office: Telephone: 020 7035 3382;
Fax: 0870 336 9137/8; E-mail: Lorna.Gordon9@Homeoffice.gsi.gov.uk

health service use, drug-related deaths and cost of social care, and these are applied to each type of user. These consequences each have a unit cost which is used to provide an estimate for the overall costs for each consequence for each type of drug user. These overall costs are then summed to provide an estimate of total economic and social cost of Class A drug use. Further details of the methodology can be found in HORS 249 (Godfrey *et al.*, 2002).

The original costs were estimated to be around £12 billion in the year 2000, within a wide confidence range of £10 billion to £18 billion. Problematic drug use accounted for 99 per cent (or £11.96 billion) of costs. Further, drug-related crime accounted for 88 per cent (or £10.56 billion) of total costs.

The methodology has now been improved, and more recent data have been used to provide an updated estimate of the total economic and social cost due to Class A drugs use in 2003/04. In particular, these methodological changes are: inclusion of a broader range of drug-related health costs; improvements to the methodology used to estimate the criminal justice system cost component; improvements to the methodology used to estimate the victim cost component; and incorporation of the most recent estimates of the prevalence of problem drug use.[18] Together, these improvements have provided more robust estimates. This does, however, mean that estimates for 2003/04 are not comparable to those in 2000.

Results for 2003/04

As in HORS 249, the cost domains are drug-related crime, health service use, drug-related deaths and social care. The total economic and social cost of Class A drug use in 2003/04 is estimated to be around £15.4 billion. This equates to £44,231 per year per problematic drug user. Based on the lower and upper range of estimates of the prevalence of drug use, the range around this estimate is between £15.3 billion and £16.3 billion. Estimates are presented for different groups (problematic drug users, young recreational and older regular users) in Tables 3.1 and 3.2. The different cost categories for the groups are consistent with the methodology in HORS 249.

Problematic drug users account for 99 per cent (or £15.3 billion) of total costs. Drug-related crime costs account for 90 per cent of costs associated with problematic drug use. Young recreational and older regular users account for less than one per cent (£52 million and £9 million respectively) of total costs.

[18] As reported later in this chapter.

Table 3.1: Economic and social costs of Class A drug use

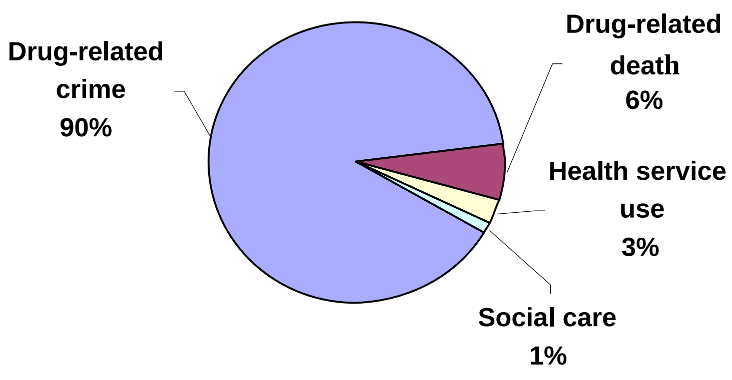
	£m	% of total cost
Drug-related crime		
Fraud	£4,866	32%
Burglary	£4,070	26%
Robbery	£2,467	16%
Shoplifting	£1,917	12%
Drug arrests	£535	3%
Health costs		
Inpatient care	£198	1.2%
Inpatient mental health	£88	0.6%
A&E	£81	0.5%
Community mental health	£61	0.4%
Primary care – GP visits	£32	0.2%
Neonatal effects	£3	0.1%
Infectious diseases	£25	0.1%
Drug-related deaths	£923	6.0%
Social care	£69	0.4%
Total	**£15,337m**	**99%**

Table 3.2: Economic and social costs of Class A drugs use associated with young recreational and older regular use, 2003/04

	Young recreational users (£)	Older regular users (£)
Drug-related deaths	£43,304,580	-
Drug possession	£6,770,885	£6,770,885
Toxicity and overdose	£4,539	£252,636
Mental health admissions	£1,805,000	£2,338,140
Total	**£51,885,004**	**£9,361,661**

Total costs broken down by domain are presented in Figure 3.1. Drug-related crime accounts for 90 per cent (or £13.9 billion) of total costs. Drug-related deaths account for six per cent (or £923 million), health service use three per cent (or £488 million) and social care 0.4 per cent (or £69 million).

Figure 3.1: Estimated total cost of Class A drug use by type of cost, 2003/04



As in HORS 249, the total cost estimates include two components of the cost of crime: expenditure by the criminal justice system in dealing with the crimes committed and the cost consequences for the victims of crime. These costs are presented in Table 3.3, which illustrates that the victim costs of crime constitute the majority of the costs of drug-related crime.

Table 3.3: Summary of estimates of criminal justice costs and victim costs of crime in 2003/04

	CJS costs (£mn)	*Victim costs (£mn)*
Fraud	£877	£3,989
Burglary	£1,419	£2,651
Robbery	£822	£1,585
Shoplifting	£383	£1,533
Drug arrests	£535	n/a

Methodological and data update

HORS 249 identified some areas for improvement in the methodology. Methodological and data improvements have now been made in four key areas.

Infectious diseases

HORS 249 identified the need for inclusion of the costs associated with infectious diseases relating to intravenous drug use in the total cost estimate. However, data to estimate these costs were limited at that time and it was therefore not possible to include these costs. Data have now become available, which has enabled the inclusion of the costs of new incident cases of HIV, Hepatitis B and Hepatitis C in the most recent estimates. The number of new cases for each disease is combined with an estimate of the present value of lifetime treatment costs for each disease. The present value of the lifetime treatment costs associated with infectious diseases are now estimated to be £23 million for HIV, £608,475 for Hepatitis C and £580,568 for Hepatitis B.

CJS cost component

Costing drug-related crime involves considering separately costs borne by the criminal justice system and those borne by the victims of crime. In the original methodology, estimation of these cost components were based largely on the National Treatment Outcome Research Study (NTORS). NTORS provides data on both the direct contact of problematic drug users with the criminal justice system and self-reports on the number of offences committed. For the CJS component, costs were based on reported contacts in order that actual public finance expenditure, as it occurred, could be estimated. However, this CJS contact did not necessarily relate to offences and drug use in the current period. Work was therefore undertaken to estimate the number of different types of offences committed by problematic drug users in different treatment states (Godfrey *et al*., 2004). These numbers of offences are costed using the CJS component of the updated cost of crime estimates that have since been published by the Home Office (Dubourg *et al*., 2005).

Victim cost component

Estimation of the victim cost component was based on the pattern of offences self-reported by NTORS clients in the original methodology, and the victim cost component based on the original Home Office costs of crime (Brand and Price, 2000). It was not possible at that time to break these self- reported offences down into types. Therefore, updates would only be possible based on general inflation. With the improved methodology the number of different types of offences committed by problematic drug users is based on Godfrey *et al.,* 2004 and the volumes are therefore now consistent with the CJS costing approach. The cost component is based on the Home Office updated cost of crime estimates.

Prevalence of problematic drug use

The main data source development has been the use of improved estimates of the prevalence of problematic drug use in England and Wales. The estimates used in the original costing model were the best estimates available at that time. These were based on the treatment demographic and treatment coverage methods and produced three alternative estimates of the number of problematic drug users in England and Wales. These were 281,125 (treatment demographic method), 337,350 (treatment coverage method) and 506,025 (treatment coverage method). The Home Office has since commissioned research to provide estimates of the number of problematic drug users in England, based on more robust indirect estimation methods.[19] The first sweeps of estimates (for 2004/05) are now available and these have been incorporated into the revised costing model. This report estimates there are around 327,466 problematic drug users (defined as opiate and/or crack users) in England, within a 95 per cent confidence range of between 325,945 and 343,424.

[19] As described in Chapter 2 of this report.

4. Estimating the size of the UK illicit drug market

Stephen Pudney, Celia Badillo, Mark Bryan, Jon Burton, Gabriella Conti, Maria Iacovou from the Institute for Social and Economic Research, University of Essex[20]

Summary

Results

This chapter describes a study estimating the size of the UK market for six categories of illicit drugs for the reference year 2003/4. The drug categories are: cannabis, amphetamines, ecstasy, powder cocaine, crack cocaine and heroin. Figure S4.1 shows that total market size is estimated as £4.6bn in England and Wales and £5.3bn in the UK as a whole: roughly 33 and 41 per cent of the size of the tobacco and alcohol markets respectively. There are major uncertainties in the measurement process, so the margins of error for these estimates are wide: £3.5bn – £5.8bn for England and Wales, and £4.0bn – £6.6bn for the UK. In terms of total street value, crack and heroin account for the largest shares of expenditure (respectively 28% and 23%); cannabis and powder cocaine make up 20 and 18 per cent of expenditure, while amphetamines and ecstasy have market shares of only six and five per cent. The picture is quite different in volume terms; for instance, the powder cocaine market is estimated to be 14 per cent bigger than the crack market in terms of street quantities, but 34 per cent smaller in value terms.

S4.1: Baseline estimates of market size for England and Wales and the UK for 2003/4

	England and Wales			UK		
	Aggregate street quantity (tonnes)	**Aggregate pure quantity (tonnes)**	**Aggregate expenditure (£million)**	**Aggregate street quantity (tonnes)**	**Aggregate pure quantity (tonnes)**	**Aggregate expenditure (£million)**
Cannabis	360.33 *±135.81*	360.33 *±135.81*	900.8 *±372.4*	412.41 *±155.44*	412.41 *±155.44*	1031.0 *±432.5*
Ampheta-mines	32.68 *±17.33*	3.60 *±2.31*	277.8 *±72.9*	36.70 *±19.46*	4.04 *±2.60*	312.0 *±81.9*
Ecstasy (millions of tabs)	52.79 *±23.84*	13.72 *±8.14*	237.5 *±76.2*	59.52 *±26.88*	15.47 *±9.18*	267.8 *±85.9*
Powder cocaine	15.7 *±12.17*	7.85 *±6.16*	863.4 *±237.1*	17.70 *±13.72*	8.85 *±6.94*	973.3 *±267.3*
Crack	13.79 *±11.76*	8.96 *±7.71*	1309.8 *±348.9*	15.58 *±13.29*	10.13 *±8.71*	1480.4 *±394.29*
Heroin	17.60 *±13.14*	7.04 *±5.32*	1055.9 *±199.2*	20.11 *±15.02*	8.04 *±6.13*	1206.7 *±227.65*
Total market value (£bn)			**4.645** *±1.154*			**5.271** *±1.310*

[20] Address for correspondence: Steve Pudney, ISER, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ. Email: spudney@essex.ac.uk.

Methodology and approach

The analysis uses a survey-based demand-side approach, building on previous attempts to estimate the size of the illicit drugs market, in particular the NERA study (Bramley-Harker, 2001). Improvements have been made through the use of better data; the inclusion of juveniles; various statistical innovations, including the treatment non-response and under-reporting in surveys; and the calculation of margins of error for the estimates. The authors of this chapter are explicit about the *a priori* assumptions which have to be made to deal with limitations in the data, and the quoted error margins include the uncertainty associated with these assumptions. Finally, the market structure estimates are checked against supply-side data on drug seizures.

Because of the incomplete coverage of available surveys, the authors have combined three main sources of evidence to obtain a figure for total consumption. The population is split into three groups, each of which is directly accessible by survey.

- *Juveniles* (aged 10 to 16), covered by the Schools Survey (SS) 2003.
- *Non-arrestee adults*, covered by the Offending, Crime and Justice Survey (OCJS) 2003.
- *Adult arrestees*, covered by the Arrestee Survey (AS) 2003-04.

The choice of surveys was based on their geographical coverage, reference periods and question content. Due to a lack of UK-wide evidence, the core analysis relates to England and Wales, and the estimates are then scaled up to UK level, using the relative sizes of juvenile, arrestee and non-arrestee populations. Survey data are used to derive average frequencies of drug use in each sub-population. To minimise non-response bias, the authors impute frequency of use for non-respondents in the OCJS and AS by matching them to respondents with similar personal and area characteristics.

The available surveys provide good information about the frequency of drug use, but only incomplete and unreliable data on drugs prices and quantities per episode of use, and no information on purity. Based on external data sources and previous case studies, the authors make explicit assumptions about prices, quantities per episode and purity, as well as the likely level of uncertainty in these assumptions. Using the price, quantity and purity assumptions, the frequency estimates are converted to consumption volumes and expenditures.

The authors have investigated the possible impact of under-reporting by respondents to the Arrestee Survey, by comparing self-reported drug consumption with the results of bio-assay test results. The authors found this to have a moderate effect, raising the estimate of total market size by seven per cent, from £5.271bn to £5.639bn. It should be borne in mind that no comparable adjustment has been made to the estimates for non-arrestee population groups and it would consequently be possible to make further upward adjustments in the total market size estimate.

Comparison with previous estimates

Bramley-Harker (2001) estimated the size of the illicit drugs market in 1998 to be £5.5bn in England and Wales and £6.6bn in the UK. The figures are substantially higher than the estimates of £4.6bn and £5.3bn in this study. Part of the gap can be attributed to sustained falls in drugs prices since 1998. The remainder reflects differences in methodology and sources of evidence. Overall, the major structural differences between these estimates and Bramley-Harker are the lower aggregate volumes and values for cannabis (66% and 85% of the earlier estimates, respectively) and heroin (53% and 65%); and the higher volume and value of the cocaine market (400% and 270%).

This study also estimates a smaller gap (14% compared to 19% in Bramley-Harker) between the size of the England and Wales market and the UK market. This difference stems from the use of age-specific population differences, rather than a single population scaling factor, and explicit consideration of the lower arrest rate in Northern Ireland.

Implications for future areas of work

The authors think there are a number of implications that have arisen from this research which if acted upon, could lead to significant improvements in the accuracy and reliability of market size estimates.

- Consider harmonisation of the question format and questionnaire content across surveys (OCJS, SS and AS), so that these sources can more readily be combined.
- Reconsider the design of survey questions on drug use to improve recall accuracy and give more direct evidence on consumption.
- Improve the system of price measurement, which is not adequate at present. The availability of a dataset giving information on amounts, prices, purities and detailed locations at the level of the individual deal or seizure would be a valuable research resource.
- Improve the data resources available in Scotland and Northern Ireland to give comparability with evidence from England and Wales, particularly in terms of arrest statistics and arrestee surveys.
- Review the collection of police arrest statistics to assess the feasibility of measuring the number of individuals experiencing arrest and the volume of non-notifiable arrests.
- Measures of aggregate market size can be updated at regular intervals to monitor changes over time. It would be inadvisable, however, to repeat the analysis more than once a year, given the large margins of error in the estimates. If the market size estimates are repeated annually, the resulting time series should only be interpreted as a broad indicator of long-term trends in drug use, not as a measure of precise year-on-year changes.

Objectives

The authors' aim in this study is to develop and implement a suitable methodology for estimating the size of the market, based on the best sources of survey evidence currently available. The research has a number of specific objectives.

- To produce an estimate of the size of the UK market for six categories of illicit drugs for the reference year 2003/4, using a survey-based demand-side approach to estimation. The six drug categories are: cannabis, amphetamines, ecstasy, powder cocaine, crack cocaine and heroin.
- To make clear the *a priori* assumptions underlying this estimate and indicate the degree of uncertainty associated with them.
- To compare the estimates with the estimates for 1998 previously produced by Bramley-Harker (2001) and identify the major sources of difference.
- To indicate the range of uncertainty associated with the estimates. This uncertainty stems from two main sources: the sampling error associated with survey estimation; and the possibility of error in the *a priori* assumptions.
- To identify weaknesses in the measurement process and make recommendations for future improvements.

Previous measurement attempts

It is beyond the scope of this report to survey the huge international research literature relevant to the measurement of drug use. However, it is worthwhile to give a brief review of previous comparable work for the UK and some indication of work that has been done for the USA and other countries with data resources comparable to those available for the UK.

Estimates for the UK

There have been two major attempts at measuring the size of the UK illicit drugs market, by research teams from the Office for National Statistics (ONS) and National Economic Research Associates (NERA). Their findings are reported in Groom *et al.* (1998) and Bramley-Harker (2001) respectively.

The aim of the ONS study was to examine the feasibility of including estimates of illegal activity in the UK National Accounts. A method was proposed, largely for illustrative purposes, to estimate the value of the illicit drugs market. The estimation approach used both supply measures (for example, administrative data on drug seizures) and demand measures (national and targeted surveys of individuals). The two sets of estimates were then "balanced" to reconcile them with the national accounting identity equating consumption with overall production. The outcome of the balancing process was a range of estimates corresponding to different assumptions about the proportion of drugs seized by police. Total consumers' expenditure on illicit drugs was estimated to be between £4.3bn (20% seizure rate) and £9.9bn (5% seizure rate) at 1996 prices. The range of estimates is not a statistical confidence interval and does not take account of the other sources of uncertainty in the data.

On the supply side, the estimates were based on seizures, prices and purity data. On the demand side, the study split the population of drug users into three: regular recreational users, occasional recreational users and problem users. The numbers of recreational users were estimated from the British Crime Survey (BCS), while information on problem users (assumed to be under-represented in the BCS) was taken from the Department of Health Drug Misuse Database. Estimation of total expenditure then required data on prices and average quantities used, or expenditure per user. Some average expenditure data were available from a previous small-scale survey, but for recreational users quantities had to be assumed. The authors recognised that their estimates depended crucially on these and the other assumptions made, and on the imperfect coverage of the data (for example, the omission of the under 16s). They concluded that their methods were not yet sufficiently developed to justify inclusion of measures of illegal activity in the National Accounts.

The NERA study can be seen as the natural successor to the ONS analysis. The demand-side approach was developed to produce an estimate of the total value of the UK illicit drugs market of £6.6bn. In value terms, heroin had the largest share of the market (35%), followed by crack (28%) and cannabis (24%). Amphetamines, cocaine and ecstasy were each estimated to account for only four to five per cent of the market. Like its predecessor, the NERA study made a distinction between regular and occasional users, and it also considered drug use by prisoners. The 1998/9 New English and Welsh Arrestee Drug Abuse Monitoring programme survey (NEW-ADAM) was used to estimate consumption by regular users. Unfortunately, NEW-ADAM covered only a small number of disproportionately urban police arrest suites, and so it is likely that the estimates of drug use were subject to biases. The British Crime Survey (BCS) was used to estimate the number of occasional users, in combination with information about frequency of use from the Youth Lifestyles Survey (YLS). Based on self-reported frequency of use, regular users were identified and excluded from the BCS sample, allowing estimation of the number of annual episodes of use among the remaining occasional users. Expenditure for each drug was imputed from total expenditure figures and usage frequencies reported in NEW-ADAM. As in the earlier ONS study, it was recognised that the estimates were sensitive to the various assumptions made, but no explicit margin of error was quoted. A number of recommendations for improvement were made in a comment by Pudney (2001) and the present study can be seen as an attempt to implement these and other improvements, where possible.

Estimates for the USA

The USA forms the world's largest market for illegal drugs (Reuter, 2000), and has the most well-developed data resources and research literature in the area.

General-population individual-level data on drug use are available through the National Household Survey on Drug Abuse (NHSDA)[21] (see OAS, 2004). There are particularly rich data from surveys focusing specifically on the young, including the Monitoring the Future survey (MtF), the National Longitudinal Survey of Youth (NLSY), the National Parents' Resource Institute for Drug Education survey (PRIDE) and the Youth Risk Behavior Survey (YRBS). More targeted sources focusing on high-prevalence groups include the survey of Substance Abuse among Probationers and Inmates, conducted by the Bureau of Justice Statistics, and state- and federal-level surveys of inmates in correctional facilities. The Arrestee Drug Abuse Monitoring (ADAM), established by the National Institute of Justice in 1987 was the core of the international ADAM programme.

The USA also has extensive data on drug prices and purity, through the "System to Retrieve Information on Drug Evidence" (STRIDE), compiled by the Drug Enforcement Administration (DEA). This provides data on laboratory analyses of street-level drug purchases and of drugs removed from the marketplace where DEA participated in the seizure(s). STRIDE also provides chemical analyses to determine the geographic origins of drugs, and is the source for data on long-term trends in the prices and purities of drug exhibits.

US experience underlines the uncertainty inherent in measurement of market size. Reuter (2005), in his review of strategies for estimating the size of the global drugs market, notes that even in the US, which currently has the most complete and sophisticated data on drug use and prices, expenditure estimates could range between $40 billion and $100 billion.

The US DEA produces estimates of drug production by other nations. These reports, published annually in the *International Narcotics Control Strategy Report* are based on reports of growing areas, crop per acre and refining yield of the raw product in source countries. Reuter (1996) enumerates several problems with these estimates, including inadequate methodology, unresolved discrepancies between different series of data, and generally implausible figures. However, he expresses a much greater degree of confidence in the estimates produced by the US Office of National Drug Control Policy, published as the series *What America's Users Spend on Illegal Drugs*. The latest in the series covers the years 1988-2000 (ONDCP 2001) and draws data from a range of sources, including those listed above, to estimate total aggregate expenditures on drugs.

Table 4.1 reproduces the series of estimated aggregate expenditures for a range of drug categories. As an example of the way the ONDCP estimate is built up, Table 4.2 gives more detail for marijuana, which shows a decline in market value of 14 per cent over the period. This estimated fall in market size is the result of a 28 per cent fall in price, outweighing the 20 per cent increase in prevalence. The fall in price is shared by the other drugs included in the ONDCP analysis and is apparent in the UK also.

[21] Now renamed the National Survey on Drug Use and Health.

Table 4.1: Estimated total expenditures on illlicit drugs, USA 1994-2000 ($ billions)

	1994	1995	1996	1997	1998	1999	2000
Cocaine	42.8	40.0	39.2	34.7	34.9	35.6	35.3
Heroin	13.2	13.2	12.8	11.4	11.1	10.1	10.0
Methamphetamine	7.6	9.2	10.1	9.3	8.0	5.8	5.4
Marijuana	12.2	10.2	9.5	10.5	10.8	10.6	10.5
Other drugs	2.6	2.7	2.7	2.5	2.3	2.6	2.4
Total	78.4	75.2	74.3	68.4	67.2	64.6	63.7

Columns may not add due to rounding error. Estimates for 2000 are projections.
Prices in 2000 dollar equivalents
Source: Taken from *What America's Users Spend on Illegal Drugs 2001*, Table 11.

Table 4.2: Calculation of total marijuana consumption, USA 1994-2000

	1994	1995	1996	1997	1998	1999	2000
Number of users (millions)	10.1	9.8	10.1	11.1	11.0	11.9	12.1
Joints used per month	18.7	18.7	18.7	18.7	18.7	18.7	18.7
Weight of a joint (ounces)	0.0136	0.0136	0.0136	0.0136	0.0136	0.0136	0.0136
Price per ounce, 1/3 ounce purchase	$397	$340	$309	$311	$322	$292	$284
Total real expenditure	$12.2	$10.2	$9.5	$10.5	$10.8	$10.6	$10.5
Tonnes	874	848	874	960	952	1028	1047

Expenditures are in $ billion dollars, in 2000 dollar equivalents. Estimates for 2000 are projections.
Sources: Table taken from *What America's Users Spend on Illegal Drugs 2001*, Table 9. Figures in the table are based on information from NHSDA and the STRIDE database.

The ONDCP estimates for marijuana are quite straightforward: a relatively large proportion of marijuana users are to be found in the general population as sampled by the NHSDA survey, and because the use of marijuana is more widespread than other drugs, and the penalties for use less severe, there is less of a problem with under-reporting. For other drugs, the estimation process is more complicated. For example, the heaviest users of heroin are likely to be missed from a household-based survey, and thus, calculations incorporate other sources of data, including arrestee surveys and emergency-room admissions. Rates of methamphetamine use vary widely between areas, so instead of using household-level data, information on drug treatment admissions was used.

Several US studies exist on markets for individual drugs, although most are concerned with modelling behaviour rather than measuring aggregate market size. This research nevertheless has important implications for aggregate measurement. For example, Caulkins and Pacula (2002) examine the nature of marijuana markets and acquisition using data from the 2001 NHSDA. They estimate that in the US there are around 400 million retail marijuana purchases annually, and that the average purchase size is small, of the order of 6-7 joints' worth. They also find that a minority of users is responsible for a majority of purchases. This has implications for the use of purchase data in the measurement of market size: if a large proportion of marijuana purchases are sold on to friends, there may be double counting in the data.

Estimates for other countries

There are relatively few systematic attempts at demand-side estimation of the aggregate size of the domestic drug market for countries other than the US. Outside the UK, Australia and New Zealand have particularly well-developed measurement systems.

Estimates by Wilkins *et al.* (2005) of the size of the market for cannabis in New Zealand are based on self-reported purchases recorded in the 2001 New Zealand Drug Survey (NZDS). The NZDS is a large-scale telephone survey of the 13 to 45 age group, with a high response rate. The sample was assumed to be nationally representative in terms of cannabis purchases. Almost half of the sample had tried cannabis and almost a fifth had used it in the last year. A high proportion (84%) of users received at least some of their cannabis for free, suggesting a large informal market of people who share their cannabis, once purchased.

Using data on the weight of cannabis each sample member had purchased in the last year, multiplied by the price paid, a figure was calculated for each sample member's annual expenditure. The proportion of the sample who had bought cannabis in the last year was multiplied by the total population aged 13 to 45 to estimate the number of cannabis buyers in the population. This was multiplied by the sample average dollar amount spent by cannabis buyers to give a national total. Confidence intervals were calculated to take account of sampling variation. The total NZ$ amount spent on cannabis in 2001 was estimated to have been just under $190 million ($131-$249 million). The authors note the limitation that the data relates only to those aged 13 to 45, but suggest that purchases made by people outside this age group are relatively low. The sample is restricted to those members of the population who reside in households and have a telephone line, so this may under-estimate the size of the cannabis-buying population. Non-response and misreporting are further potential sources of bias. The authors point out a danger of double-counting since some cannabis may be purchased to sell on to others and so the same quantity of cannabis may be reported twice. This would have the effect of over-estimating the value of the market.

There is little work on cocaine, crack or heroin in New Zealand, although the market is believed to be small. There has, however, been a recent study (Wilkins *et al.*, 2004) of Amphetamine Type Stimulants (ATS) which cover synthetic illicit drugs such as methamphetamine, ecstasy and crystal methamphetamine. This study also uses the 2001 NZDS, a national survey of drug treatment workers and in-depth qualitative interviews with heavy methamphetamine users. It also gives a comparison with Australian data on ATS use. From the 2001 NZDS it was found that six per cent of the sample (of 13- to 45-year-olds) had used an ATS in the previous year. Compared to Australia, New Zealand had a lower rate of ecstasy use.

The size of the ATS market was estimated in much the same way as for cannabis expenditure, with some differences: multiple imputation (based on quantity, age by gender and amount typically paid) was used for cases of non-response or where they had received their drugs for free. The resulting estimate was adjusted by 20 per cent to allow for general under-reporting of drug use. This assumed under-reporting rate was reached through a comparison of findings from the NZDS and figures from two New Zealand birth cohort surveys, which suggested that under-estimation of drug use in the former ranged from 10 to 30 per cent. The amphetamine trade was estimated to be worth NZ$123 million (US$40-242 million) annually with users spending, on average, NZ$1000 a year. The ecstasy market was estimated to be NZ$46 million (US$30-62 million) a year with users spending an average of NZ$600 annually.

Australia has a wide range of drug-use surveys, including an arrestee survey (DUMA), which is similar to other ADAM surveys. General population surveys include the National Survey of Mental Health and Wellbeing of Adults and the large-scale National Drug Strategy Household Survey (NDSHS), covering the over-14 population using a mixed-mode survey method. There is also an Australian School Students Alcohol and Drugs Survey (ASSADS), covering the 12 to 17 age group and asking about prevalence with respect to various recall periods, but not about quantities or expenditures. The Australian Illicit Drug Reporting System (IDRS), conducted annually in capital cities, monitors the price, purity and availability of the main illicit drugs. This is collected through interviews with injecting drug users (implying respondent-initiated sampling), experts and professionals in the drug field and through existing data collected by other agencies. The IDRS reports average prices and purity of drugs, rather than total expenditure.

The authors are aware of no systematic attempt to combine these sources of information into a single estimate of market size for Australia.

Methodology

The analytical approach used here is designed to address several methodological concerns.

- In the present state of knowledge, supply-side estimates are unlikely to be reliable. This is because the trade in illicit drugs and the organisation of supply within the UK are very difficult to observe in a systematic way.
- No single source of survey data will be adequate for measuring total demand, because of incomplete coverage and non-response.
- Survey response is potentially an important issue, since the characteristics of survey non-respondents may differ systematically from those of respondents.
- The available survey data may be subject to measurement error. In particular, there is reason to believe that some survey respondents may under-report their use of drugs.
- No estimate of market size will be completely reliable, so indicators of reliability should be offered to users of the estimates.

Unlike much of the earlier work in this area (for example, Rhodes *et al.*, 1997, and Bramley-Harker, 2001), the authors do not make a formal distinction between "regular" and "occasional" users, since these terms are ambiguous and they define consumer groups which cannot be surveyed directly. The analysis is based on a partition of the population into three groups, which are more directly accessible by survey.

Juveniles, aged 10 to 16, observed through a school-based survey.

Non-arrestee adults, defined as people aged 17 to 65 who do not experience arrest within the reference year; they are observed through a survey of the household-resident population.

Arrestee adults, defined as those aged over 16 who experience arrest within the reference year; they are observed through a survey of arrestees in police custody.

Children aged under ten and non-arrestees aged over 65 are assumed to have zero drug consumption. Each of the three groups is partitioned further, according to their response behaviour in the relevant surveys; thus, in principle, the authors allow non-respondents to be distinguished from respondents. The explicit treatment of non-response means that it is possible to explore the consequences for measurement of making different assumptions about the behaviour of non-respondents.

A major distinction between the authors' approach and the earlier NERA analysis is that this study does not assume that all regular drug users are necessarily subject to arrest and thus representable solely by evidence from the arrestee population. Moreover, unlike NERA, this study does not "scale up" evidence from the arrestee population to represent the hypothetical wider set of potentially arrestable drug users, since the source of evidence on the general population appears to represent adequately those who have experienced arrest but are not in the population of current arrestees.

The full approach to measurement is set out schematically in Figure 4.1. The process has several stages.

Stage 1 Imputation of frequencies of drug-use for survey non-respondents. This is done by fitting a statistical ordered-response model to data on respondents and then using it to predict the probabilities of each category of consumption-frequency for all non-respondents. The explanatory variables in these models are variables which can be observed for both respondents and non-respondents. For the household-population survey, these variables are mainly descriptors of the local area; for the arrestee survey, they include demographic characteristics and circumstances of arrest. Non-response is a relatively minor problem for the survey of juveniles and the authors make no adjustment in that case.

Stage 2 Adjustment of the estimated distribution of frequency-of-use to deal with under-reporting by survey respondents. This is difficult, owing to the lack of external checks on the validity of responses. The authors did not use adjustments for under-reporting in the main estimates but did report the results of an experimental adjustment procedure, which makes use of the biological tests made on saliva samples from arrestee survey respondents.

Stage 3 Use of assumptions about average drug prices, quantity consumed per episode of use and purity to convert frequency-of-use to predicted annual consumption and expenditure rates. This process is a simple matter of multiplying by assumed average price, purity, and rate of use per day to convert predicted frequency-of-use probabilities into expected consumption. However, these conversion factors are based on the rather thin evidence currently available and they introduce the largest element of uncertainty into the estimation process.

Stage 4 Construction of population aggregates for England and Wales using external information on the sizes of the adult non-arrestee, juvenile and arrestee sub-populations. The authors first estimated the size of the arrestee population, allowing for the fact that many individuals appear multiple times in official arrest statistics. They then combined estimated consumption figures for the juvenile, adult non-arrestee and arrestee populations, using their relative sizes as weights.

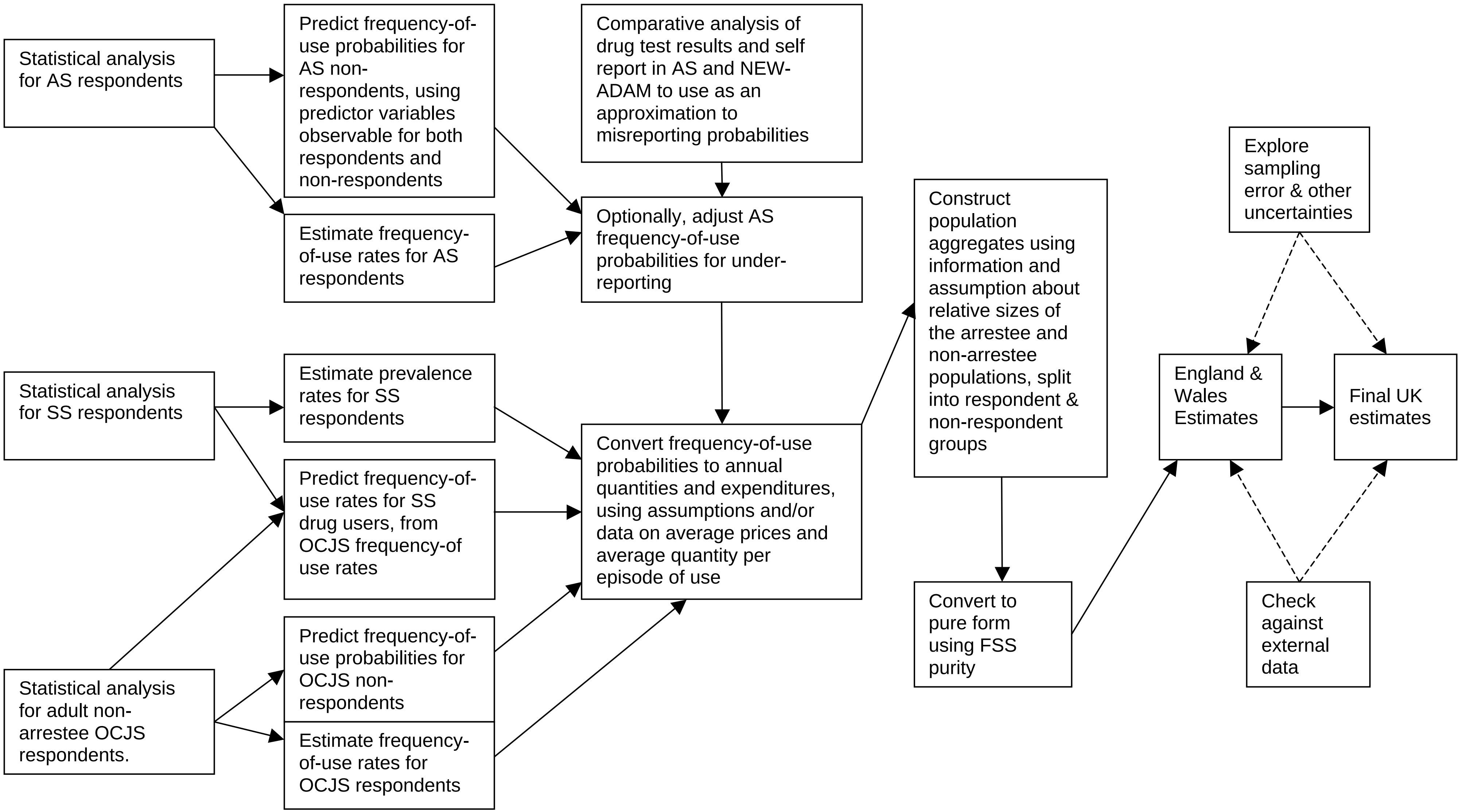
Stage 5 Extension of the estimates from an England and Wales to a UK basis. For this, official estimates of the size of the Scotland, Northern Ireland and England and Wales populations, together with information on the (proportionately smaller) arrestee population in Northern Ireland were used. The assumption underlying the procedure is that juveniles, arrestees and non-arrestees have essentially the same behaviour patterns in all nations of the UK and that rates of drug use by drug users are also uniform.

Stage 6 Comparison with external indicators of the structure of the drugs market. The authors compared the estimated market structure with the composition of drug seizures, to give a necessarily rough assessment of the concordance of the demand-side estimates with (arguably) the most reliable supply-side indicator.

Stage 7 Calculation of indicators of the reliability of measured aggregates. The authors used an *ad hoc* method of combining two sources of uncertainty into a single rough indicator of the likely scale of error in their estimates. For each of the parameters fixed by external assumption (average street prices, purities and consumption amounts per use-day), the estimation process was simulated by drawing a large number of alternative values from an *a priori* range of uncertainty (assuming a uniform distribution). For each of these replications, the authors estimated total market size, computing confidence intervals (by means of a bootstrap method) for the intermediate quantities which are constructed from survey data. The total variation produced by this simulation process then gives an indication of the range of uncertainty in the final results.

More details of the various stages of this process are given in Appendices 3, 7 and 8.

Figure 4.1: Sizing the UK illicit drugs market methodology



A review of data sources on drug use

There now exist new data sources which were not available when the ONS and NERA studies (Groom *et al.*, 1998; Bramley-Harker, 2001) were carried out and which offer scope for improvement. This section aims to review the data resources currently available including those which were available to the NERA team as well as those which have become available more recently.

For this project it was necessary to know the proportion of the population using different drugs, how often they use drugs and the amount they spend on those drugs. The questions that the authors were mainly interested in cover:

- the prevalence of drug-use;
- the frequency of drug-use;
- the quantities of drugs used; and
- expenditure on drugs.

The main surveys are briefly summarised below, with more technical details presented in appendix Table A4.1.

Surveys of the general population

The Offending, Crime and Justice Survey (OCJS)

The Offending, Crime and Justice Survey was commissioned by the Home Office with the objective of providing a base for measuring prevalence of offending behaviour and drug use in the general population. Specifically, the survey covers the general public aged between 10 and 65, living in private households in England and Wales. As detailed in appendix Table A4.1, the core sample of the 2003 survey consists of 10,085 respondents, interviewed in the period January-July.

The survey is mainly carried out with computer-assisted personal interviewing (CAPI), but a computer-assisted self-administered questionnaire (CASI) is used for sensitive areas like experience of arrest and drug use. Relevant questions ask which drugs have ever been taken, which have been taken in the past 12 months or four weeks, the frequency of drug use, the time since each drug was last used, the total amount spent on drugs and the effects of drugs.

Information on OCJS non-respondents is limited and relates mainly to the neighbourhood. It is summarised in appendix Table A4.2, which demonstrates the strong divergence between the characteristics of respondent and non-respondent households.

The British Crime Survey

The British Crime Survey (BCS) is a long-established Home Office survey, conducted on an annual basis since 2001. The context of the BCS differs from the OCJS, as it is focused more on victimisation than illegal activities. The BCS target population comprises adults aged 16 and over in private households in England and Wales. The 2002-03 survey yielded a core sample of 37,395 interviews over the period April 2002 to March 2003, representing a response rate of 74 per cent (Table A4.1). In 2002/03, the questions about drugs, asked in a self-completion CASI module, asked: whether the respondent had ever taken any of a particular drug; whether it was in the last 12 months; or in the last month; and whether the respondent had injected any drugs. Questions were also asked about the age each drug was first used, the ease in obtaining each drug and the frequency of drug use in the last 12 months. Critically, for these purposes, the BCS does not contain questions which would allow arrestees and non-arrestees to be distinguished in the sample.

The Schools Survey

The Schools Surveys (SS) are carried out on behalf of the Department of Health by the National Centre for Social Research (NatCen) and the National Foundation for Educational Research (NFER). The SS uses a self-completion questionnaire and smoking diary, given to pupils in "exam conditions" at school, under the supervision of an interviewer. It is believed that the collection of data in this setting encourages honest reporting.

There were annual cross-sectional waves from 1998 to 2003. In 1998 and 2000 the sample covered both England and Scotland, but in 1999-2003 coverage was restricted to England only. The target population is pupils in secondary school years 7 to 11 (age 11 to 15); however, special schools and hospital schools are excluded from the sample. The sampling is carried out over multiple stages: first, the schools are sampled, then pupils are selected within each school. In 2003, the overall response rate was 65 per cent, yielding a sample of 10,390 pupils. The drugs questions ask whether respondents had heard of particular drugs, had been offered any of them, had they ever tried them, their age at first use and the timing of their last use. There were also questions about the nature of the first and most recent use, frequency of drug use and the behaviour of friends and siblings. Unlike the other surveys considered here, no survey weights are produced for the SS, owing to the low variations in response across types of individual and schools.

The Youth Lifestyles Survey

The Youth Lifestyles Survey (YLS) was based on the 1998 BCS sample and can be viewed as a youth booster sample to the BCS: respondents in households with at least one person aged 12 to 30 were asked if these young people could be interviewed for the YLS. Focused enumeration (asking about adjacent addresses) was also used to build a further sample. In total, the YLS contains 4,848 12- to 30-year-olds living in private households in England and Wales. This represents an overall response rate of 69 per cent, somewhat lower than is typical in the BCS and OCJS because the YLS sample had already experienced a stage where non-response was possible. Survey weights are provided to correct for non-response (and the over-sampling of large urban areas).

The survey was conducted using CAPI, with a self-completed (CASI) element which included the questions about drugs. Respondents were asked if they had ever taken drugs, if they had done so in the last 12 months, their frequency of drug use in that period, their average monthly expenditure on drugs, their age of initiation into drug use and the context in which they last used drugs.

Surveys of arrestees

The Arrestee Survey

The Arrestee Survey (AS) is a survey of people who are in police custody on suspicion of a criminal offence. In the October 2003-September 2004 period to which the data relate, 60 police custody suites were sampled, after stratification by arrestee inflow. In each of the sampled suites, a set of six-hour interviewer shifts was allocated by random assignment. During each shift, the interviewer attempted to contact every eligible arrestee who was in custody for any part of the shift. Once consent was achieved, the interview was conducted using (mainly) computer-assisted self-interviewing (CASI) and a saliva sample taken and then tested for cocaine and opiate residues.

Non-response is a major issue in the AS. In the sample period used, of all eligible arrestees recorded by interviewers as present during some part of the shift, only 28 per cent of cases (7,549 respondents) yielded an interview and a further 11 per cent refused to participate (a refusal rate of 28%). The remaining 61 per cent of arrestees could not be interviewed for a range of reasons, including lack of time (of either the interviewer or custody staff) and the poor physical or mental state of the arrestee. However, an unusual feature of the AS is that it gives quite detailed information on non-respondents, including basic demographic characteristics (age, sex and ethnicity), the time and date of entry into custody and the suspected offence that gave rise to the arrest. The analysis exploits this information on noninterviewed arrestees, which is summarised in appendix Table A4.3. There are significant differences between respondents and non-respondents, particularly in terms of timing of entry into custody, the reason for arrest and age.

NEW-ADAM

The New English and Welsh Arrestee Drug Abuse Monitoring programme (NEW-ADAM) was linked to the international Arrestee Drug Abuse Monitoring (I-ADAM) Program. The fieldwork for the UK version took place in three phases: July 1999-April 2000, May 2000-April 2001 and May 2001-April 2002. Interviewers were present continuously in custody suites for a one-week period and eligible adult arrestees were interviewed as they became available. Of all those approached for interview, around 85 per cent agreed (resulting in 3,091 interviews in the first two years and 1,554 in the third).

The NEW-ADAM questionnaire varied over custody suites, so that not all questions are asked of all respondents. All interviewed arrestees were asked whether they had ever taken any of a list of drugs, if they had taken each in the last 12 months, the number of days in the last 30 they had used each drug and the age at which they first took each drug. Respondents were also asked whether they had ever injected drugs, if they had injected in the last 12 months, in the last 30 days and the last three days. Arrestees were also asked for the amount spent on drugs in the past seven days, and whether this amount was usual. Questions also covered previous contact with the Criminal Justice System, and frequency of arrest during the past five years and the past 12 months.

Summary and choice of data sources

The authors' aim in this project is to combine data sources to estimate the size of the drugs market. This implies that there is a need to have surveys covering a common reference period, where possible. They also need to cover the same geographic area or at least cover the minimum area necessary. The surveys should be of high quality, with clearly stated sampling designs.

The ideal would be to cover the whole of the UK. However, most of the available surveys cover only part of the UK. Those discussed above typically cover England and Wales. Limitations of coverage mean that other surveys such as the Health Education Population Survey 2001-2003 (Scotland), the Scottish Crime Survey, the Continuous Household Survey (Northern Ireland), Reaching and Keeping Tweenagers (various parishes) and Drug Pathways into Young Adulthood (North West England) are of limited use in the main project design.

The arrestee data used comes from the Arrestee Survey, which is available only from 2003/4. Matching general population surveys must also relate approximately to this reference period. Surveys with an earlier fieldwork and reference periods will be of limited use in the main project. This rules out the Mental Health of Children and Adolescents in Great Britain survey (1999), the National Child Development Study and 1970 British Cohort Study Follow-ups (1999-2000), Psychiatric Morbidity Among Adults Living in Private Households (2000), Youth Lifestyles Survey (1998) and the ONS Survey of Psychiatric Morbidity Among Prisoners in England and Wales (1997).

On grounds of timing, the general population surveys that are most suitable for this project are the BCS (2003) and the OCJS (2003). The BCS has the advantage of continuity with past analyses but this is not a major benefit, since the sources underlying the other parts of the analysis have not previously been available. There are several reasons to prefer the OCJS to the BCS for the purposes of this study. Firstly, the absence of questions on arrest in the previous year makes the BCS hard to incorporate in the analytical framework. Secondly, the more coherent design of the OCJS may give it better response properties than the BCS, where the drug use questions are added on to what is, primarily, a victimisation survey. This view is consistent with the significantly higher rate of prevalence apparent in the OCJS sample, suggesting that under-reporting may be a lesser problem than in the BCS. Thirdly, the presence of a re-interview sub-sample in the OCJS should permit more precise estimates of change over time, when this analysis is repeated in the future. For these reasons, the authors have replied on the OCJS as their main data source for the adult non-arrestee population.

For juveniles, they have used the SS (2003) to avoid the under-reporting that they believed is produced by the interviewing of young people in a home environment, where adults may be present. The higher prevalence rates found in the SS relative to the OCJS are consistent with this view. This chapter now goes into more detail on these differences between surveys.

Data reliability

The impact of survey design on drug self-reports

In the sensitive area of drug use, survey responses may be particularly vulnerable to aspects of survey design such as sample selection, the phrasing of the questions, interview conditions, and the perceived confidentiality of the answers. If survey data are to produce credible measurements of drug use, researchers need to be confident that the answers are not just an artefact of the particular survey chosen. And if surveys do differ, they need to know by how much and, ideally, why. A necessary first step in exploiting survey data is therefore to assess the practical effect of differences in design amongst commonly used surveys.

In this section the authors have compared the data sources which they used in this study – the OCJS, SS (both 2003) and AS (2003-04) – with each other where appropriate, and with two other similar surveys, the BCS (2002/3) and NEW-ADAM (1999-2002). In a fuller study (Badillo *et al.*, 2005), the authors have also compared the 1998 releases of the BCS and SS with the YLS, and the authors refer briefly to these results in this section. To facilitate comparisons, they have focused on two simple consumption measures of the six drugs considered: whether respondents reported ever taking each drug, and whether they had done so in the last year. The questions required to construct these measures appear in all the surveys and also have the advantage that they are generally more straightforward and unambiguous than questions about quantities purchased and consumed. The first measure is a useful benchmark for past drug use, whilst the second measure is a better reflection of "current" drug use and should also be less affected by recall error. This approach has been to compare subgroups of respondents who are as similar as possible across the different surveys (specifically, the same groups surveyed in the same year in the same geographical area). The data have also been weighted where possible to adjust for differing sample composition.

The 2003 OCJS and 2002/3 BCS

This subsection begins with a comparison of the 2003 OCJS, with another general population survey, the 2002/3 BCS.[22] Both cover private household in England and Wales, but in contrast to the BCS, the OCJS is less concerned with experience as a victim of crime and more with offending behaviour by the respondent. Both surveys use CASI and have similar responses rate (about 75%). However, question wording and routing differ somewhat. Whilst the question asking about ever taking drugs is similar in the two surveys, the BCS then asks if the respondent has taken that drug in the last 12 months whilst the OCJS splits the time period up into the last four weeks and then the last 12 months (not including the last four weeks).

As shown in appendix Table A5.1, the OCJS elicits a higher proportion of drug users than the BCS. For example, the weighted estimates show that 34.4 per cent of respondents in the OCJS reported ever having taken cannabis, compared to 30.6 per cent in the BCS. This is an absolute gap of 3.8 percentage points and a proportionate gap of 13 per cent. There are similar proportionate gaps for the other drugs, with, in some cases, larger proportionate gaps when the measured prevalence is low (for example, 0.2% reported taking crack in the last

[22] The OCJS took place in the first half of 2003 whilst the BCS was conducted on a monthly basis throughout the 2002 Financial Year and so, whilst the two surveys have an overlapping period of time, they are not concurrent.

year in the BCS, compared to 0.3% in the OCJS). However, these differences are not dramatic and, for three drugs (amphetamines, crack and heroin), they are not statistically significant. In Badillo *et al.* (2005) the authors compare 16- to 24-year-olds from the BCS and OCJS youth boosts. The OCJS, once again, suggests higher prevalence and a higher level of recent drug use than the BCS. The authors find similar proportionate differences, and, as for the full sample, no significant evidence of a difference for amphetamines, crack and heroin.

In Badillo *et al*. (2005) the authors also analyse the sub-sample of interviewees who responded to both the BCS and the YLS. With the caveat that respondents were on average eight months older in the YLS than they had been in the BCS, the expectation would be to see very similar responses from these identical samples. In fact, the results show differences of similar magnitude to those above. This suggests that the gaps between the surveys are likely to be due to survey and interview factors rather than the sampling of respondents.

The 2003 SS and 2003 OCJS

The SS sample of young people (aged 11 to 15) in England with English 11- to 15-year-olds from the OCJS are now compared. The main difference between the surveys was the interview mode: a paper-based self-completion survey conducted under “exam conditions” in the SS, compared to CASI self-completion in the OCJS.

The SS elicits a much higher proportion of respondents who claim to have ever used drugs in their life (see appendix Table A5.2). Whilst almost eight per cent of those questioned in the OCJS say that they have tried cannabis, 15.3 per cent give the same response in the SS. This is the largest absolute difference. The biggest proportionate difference is for heroin: whilst just 0.1 per cent of the OCJS sample in England aged between 11 and 15 say that they have tried heroin before, the proportion in the SS is 1.1 per cent. There is a similar pattern to reports of drug use in the past 12 months, for example the proportion who have used cannabis in the last 12 months in the OCJS (6.2%) is less than half of that elicited in the SS (12.5%).

It is hard to say for certain which set of figures is accurate but, given that the OCJS is conducted in the student’s home whilst the SS is conducted in a more anonymous classroom setting, it seems plausible that the SS context encourages respondents to be more honest in their answers. Some evidence consistent with this was found when the analysis was repeated using only those respondents in the OCJS who completed the questionnaire with no other household member present in the room. The proportion reporting some use of cannabis increased to 10.6 per cent and the proportion reporting amphetamine use increased to 0.7 per cent.[23] This narrows the gap between the OCJS and the SS, but the estimates are still significantly different. An alternative (and, in the authors’ view, generally less plausible) interpretation of the SS-OCJS difference is that proximity to friends and peers in the classroom in the SS may lead to ‘bragging’, despite the fact that other pupils cannot see the answers of the respondent. This seems most likely in the case of heroin, where the prevalence figure is higher than one might expect.

NEW-ADAM and the Arrestee Survey

Finally, the Arrestee Survey (2003-04) is compared with its predecessor, NEW-ADAM (1999-2002). Because they cover different years and were designed differently, the estimates are expected to differ more than in the previous comparisons. The results (see appendix Table A5.3) confirm that there are some quite large differences between NEW-ADAM and the AS. The ‘ever taken’ prevalence of amphetamines is 13 percentage points lower in the AS than in NEW-ADAM and the proportion having taken cannabis in the last year is ten percentage points lower. These differences may be due in part to changes in demand (there is strong evidence of a downward trend in amphetamine use) and in police arrest practices. Other estimates are very similar across the two surveys: for example, the proportion of respondents ever having used crack is about 36 per cent in both. To investigate whether the observed differences may stem from the different geographical coverage and time periods of the

[23] For the less common drugs, there are too few observations to make reliable inferences in the restricted OCJS sample.

surveys, the authors restricted NEW-ADAM data to its most recent year before the AS, and only used police suites common to both NEW-ADAM and the AS. However, the discrepancies between the two surveys are not obviously smaller in the restricted comparison. A caveat is that the estimates in the smaller samples are less precisely estimated, and that the surveys are still separated by two years, during which drug consumption could have changed.

Overall, larger proportionate differences are found between the two surveys of arrestees than between the general population surveys (OCJS and BCS). Much of the difference is likely to be caused by the very different survey designs and survey periods.

Under-reporting of drug use: self-reports and biological tests

Not everyone, when asked about drug use, will give honest or accurate answers. All the surveys covered in this chapter are confidential, and respondents are assured that their answers will be used only for scientific purposes, and that they will not be personally identifiable from the survey data. However, it is understandable that interviewees will interpret these assurances in different ways, and this will affect the extent to which they are inclined to divulge information about illegal activities to interviewers.

It is difficult to assess the extent of under-reporting, but one line of approach to the problem is offered by the NEW-ADAM and the Arrestee Surveys, where, in addition to the questions on drug use, consenting respondents are also given biological tests for drug residues. In the NEW-ADAM survey, a urine sample is taken, which is tested for traces of amphetamines; crack and powder cocaine (the test does not distinguish between these two forms of cocaine); opiates, and cannabis. The NEW-ADAM urine tests are expected to be sensitive to consumption of the relevant drug within the previous three days on average, although this may vary between individuals and with the intensity and other conditions of drug use. In the AS, a saliva sample is tested for opiates and crack/powder cocaine (again, not distinguishing between these two). The saliva test is generally regarded as having a two-day detection window on average.

The proportion of interviewees testing positive for drugs, but who report no use of the drug during the appropriate window of time, is rather high. For example, in the AS, 40 per cent of those testing positive for cocaine deny having used either crack or powder cocaine within the past 48 hours, and 29 per cent of those testing positive for opiates deny having used heroin within 48 hours.

However, some people testing positive for drugs *do* admit using the drug recently, but outside the appropriate window of time. In the AS, the proportion denying use of the relevant drugs within the past four weeks falls to 23 per cent of those testing positive for cocaine and 21 per cent of those testing positive for heroin. Only 16 per cent of those testing positive for either drug deny using the drug within the past year. Table 4.3 shows, for each drug, the percentage of respondents in the AS and NEW-ADAM with a positive drug test, who deny using the drug within the past year. The proportion of misreporting ranges from six per cent for cannabis, up to 26 per cent for amphetamines.

Table 4.3: Percentage of interviewees with positive drug test results who deny using the drug in question during the previous year

Survey	NEW-ADAM	AS (unweighted)	AS (weighted)
Opiates	11.4 *7.8*	16.1 *22.4*	20.5 *24.2*
Cocaine	15.3 *16.1*	15.0 *21.9*	17.2 *24.3*
Amphetamines	26.3 *30.6*	-	-
Cannabis	6.0 *5.4*	-	-

Figures in italics refer to the same analysis excluding interviewees who said they had recently taken prescription or over-the-counter medicines.

A “quick and dirty” approach to correcting for the under-reporting of drug use would be to inflate reported rates of drug use from other surveys by factors based on the figures in Table 4.3. For example, it could be assumed that the rate of under-reporting of cocaine use in the general population is 15.3 per cent, the same as in the arrestee population, and construct an adjustment factor of 1/(1-0.153) = 1.18, based on this figure, which calls for an 18 per cent inflation of recorded consumption rates. More sophisticated adjustment methods would be preferable. In the results presented below, the effects of adjusting AS estimates for under-reporting are explored but no analogous adjustment to estimates based on the OCJS and SS is made. These calculations should be seen as experimental rather than firm estimates.

Prices and quantities

The prices paid and the quantities consumed by drug users are not directly observed in the available survey data. Instead, external evidence must be used as the basis for the core assumptions on prices and quantity per episode of use. The difficulties in using survey-based unit values as price measures are first considered.

Price measurement: unit values from survey data

Most general population surveys (like the OCJS or BCS) do not give enough information to allow prices to be estimated. However, arrestee surveys (such as NEW-ADAM and the Arrestee Survey) give information on transacted quantities and the associated expenditure, which allows a unit value to be calculated for each transaction. The authors have concluded that these unit values are very unreliable, with enormous dispersion and a mean or median that differs substantially from other available price indicators. One cause of this measurement problem is that survey respondents may have difficulty in distinguishing between the different units in which drugs are purchased or in relating those units to the options offered by standard survey questions. In the Arrestee Survey, respondents were given the option of answering in terms of ounces or grams and this appears to have caused significant problems. Heroin is often purchased in £10 or £20 “bags” and crack is bought by the “rock”, both of which are variable in terms of quantity. Powder cocaine is generally sold, at least nominally, in grams so there may be less difficulty than for heroin and crack. However, there are also large outliers in responses to expenditure questions, so it should not be assumed that it is only confusion over physical units that is responsible for measurement error.[24]

Table 4.4 and Figures A5.1-A5.3 show the distributions of unit values from the Arrestee Survey sample for Oct 2003-Sep 2004. (Before analysing the data, the sample has been

[24] A revised questionnaire has been introduced in Wave 2 of the AS in an effort to overcome these problems. It is likely that AS price and quantity responses will become more stable in the future.

trimmed by dropping cases with a reported transaction of less than a tenth of a gram, or more than 10 grams.) On this evidence, allowing for sampling error, heroin, crack and cocaine have mean prices per gram in the range of £30-£50, with crack commanding the highest prices. These prices are noticeably out of line with evidence from other sources.

Table 4.4: Unit value distributions from AS data
(£ per street gm.; weighted for survey design and non-response; sample trimmed to give reported quantity between 0.10 and 10 grams; standard errors in parentheses)

	Heroin	**Powder cocaine**	**Crack**
Mean	£38.36 (£1.41)	£40.66 (£0.87)	£45.00 (£1.26)
n	1384	1209	1378

NCIS price data

Data on price are collected by the National Criminal Intelligence Service, using information generated during the course of police operations. This is not published routinely by the Home Office,[25] but serves as a nationwide price index for UK law enforcement. There are regional figures available, often given in the form of a range of prices. These price ranges are not easy to interpret (they are not interpretable as confidence intervals, for example) and there is no fully reliable way of converting them into point estimates. Tri-annual time series have been constructed for the UK by using the mid-points of price ranges as point estimates for geographical areas, which are then combined to form a national average. These series are given in Table 4.5 for the period Dec 2000-Dec 2004. They suggest a substantial fall of around 20 per cent in most drug prices over this four-year period, with a particularly striking price fall for ecstasy.

The NCIS price data suggest that unit values from the Arrestee Survey greatly under-estimate drug prices, particularly for crack, where the AS mean unit value is less than half the NCIS figure.

Table 4.5: NCIS price series for 2000-2004

	Cocaine (gm)	**Crack (gm)**	**Heroin (gm)**	**Ecstasy (dose)**	**Amphetamine (gm)**	**Herbal Cannabis (oz)**	**Cannabis Resin (oz)**	**Skunk (oz)**
Dec 2000	£65	£115	£70	£9	£9	£82	£85	£145
Dec 2001	£60	£105	£63	£7	£9	£80	£77	£140
Apr 2002	£60	£105	£60	£8	£9	£79	£73	£132
Aug 2002	£58	£110	£61	£6	£9	£87	£79	£133
Dec 2002	£56	£100	£61	£6	£8	£83	£81	£134
Apr 2003	£56	£95	£61	£6	£9	£82	£76	£127
Aug 2003	£55	£95	£61	£5	£9	£74	£66	£121
Dec 2003	£55	£95	£62	£5	£9	£72	£66	£114
Apr 2004	£52	£95	£59	£4	£8	£73	£66	£118
Aug 2004	£53	£95	£56	£4	£9	£71	£62	£121
Dec 2004	£51	£90	£55	£4	£8	£72	£61	£121

[25] Average prices based on NCIS data are published by the UN, however.

Other UK price data

There are many other sources of information relating to drug prices but very little is as systematic or has such wide coverage as the NCIS data. Many internet sources quote "typical" street prices but there is wide variation in the figures quoted and it is not possible in most cases to assess the statistical reliability or the geographical specificity of the figures.

Widely-quoted non-government sources include the annual survey of the readers of Mixmag magazine, which gives rather lower mean prices than NCIS. Price data are also published by the Independent Drug Monitoring Unit (IDMU) Ltd, which runs a survey of drug use, conducted at music festivals and similar events (and, recently, as a web-based survey). Average prices for IDMU data are estimated and published (Atha and Blanchard, 1997). Like NCIS data, neither the Mixmag nor the IDMU survey follows standard survey design principles, so all of these price estimates should be treated with caution. Nevertheless, the absolute and relative prices of different drugs quoted on the IDMU website are broadly consistent with the NCIS national average figures.[26]

Purity

The Forensic Science Service (FSS) regularly tests seized substances for the presence and purity of illicit drugs. This is done separately for seizures made by HM Revenue and Customs (HMRC, formerly HM Customs and Excise) and the police. Figures for police seizures, which are likely to be a better reflection of "street" purities have been used here. Table 4.6 reports the most recently-published averages for cocaine, crack, heroin and amphetamines (Ahmad and Mwenda, 2004).[27] Purity figures for ecstasy are not routinely published, nor is the THC content of cannabis samples.

Table 4.6: Drug purity
(Forensic Science Service estimates for police seizures, published in Murray and Fiti, 2004)

	Cocaine	**Crack**	**Heroin**	**Amphetamines**
2002Q1	64%	74%	40%	14%
2002Q2	63%	74%	34%	15%
2002Q3	58%	69%	40%	14%
2002Q4	59%	68%	45%	14%
Mean 2002	61.0%	71.3%	39.8%	14.3%
Quarterly std dev 2002	2.9%	3.2%	4.5%	0.5%
Mean (1997-2002)	56%	76%	42%	12%
Quarterly std dev (1997-2002)	5.5%	7.3%	5.9%	4.4%

The distribution of purity levels across seized samples is not generally published. However, distributions for police seizures made in 2004 have been provided by FSS and are reproduced in Table 4.7.[28] These distributions are highly dispersed, with coefficients of variation in the range 0.3 to 1.0.

[26] IDMU price data are not reproduced here for copyright reasons.
[27] These figures are unweighted averages over all police seizures; quantity-weighted averages are higher but might be misleading, since some large seizures would be cut further before their final sale.
[28] These figures are reproduced from the restricted publication *FIB Drugs Update*, issue 29.

Table 4.7: Distributions of drug purity, Jan-Sep 2004 (Forensic Science Service estimates for police seizures)

Purity	Amphetamines	Cocaine	Crack	Heroin
0-9%	0.69	0.04	0.00	0.01
10-19%	0.24	0.13	0.01	0.05
20-29%	0.03	0.18	0.04	0.16
30-39%	0.01	0.19	0.09	0.27
40-49%	0.01	0.11	0.13	0.33
50-59%	0.01	0.09	0.13	0.15
60-69%	0.01	0.07	0.12	0.02
70-79%	0.00	0.12	0.16	0.00
80-89%	0.00	0.06	0.23	0.00
90-100%	0.00	0.00	0.07	0.00

Drug quantities per episode of use

One of the major problems involved in any attempt to estimate turnover in illicit drugs markets is that direct survey estimates of amounts consumed or purchased in a specific reference period are lacking. Most surveys ask about ‘usual’ frequency of use or typical frequency of use within the last year, where the allowable responses correspond to the number of days in which the drug is consumed. There is no direct question in the OCJS or AS on the amount of money spent or the physical quantity purchased or used in a given period. This means that the main indicator of demand is a measure of frequency, rather than amount, of use. To convert this into quantity terms, a figure for the average quantity consumed per episode of use is required, where an episode is defined as a day in which the drug is consumed.

Evidence from the Arrestee Survey

The only available wide-coverage source for such an estimate is the Arrestee Survey, which asks directly about the quantity consumed on the last occasion the drug was used. Respondents can choose to answer this question in ounce or gram terms, but the question is only asked for cocaine, crack and heroin. Inspection of the distribution of responses (see Figures A5.4-A5.6 in appendix 5) suggest that there are cases of confusion over units and probably also confusion between amounts purchased and amounts actually used. This confusion leads to large outliers, which can seriously distort sample estimates. If only the most extreme outliers (above 10 gm) are trimmed, the estimated means are implausibly high: 1.3, 1.5 and 1.3 gm per day for heroin, cocaine and crack respectively. At an average purity of around 40, 50 and 65 per cent, these would imply very high consumption rates per episode of 520mg, 750mg and 845mg respectively.

There is some evidence from Table A5.1 in the appendix that intensive users (five or more days a week) have higher rates of use per episode, at least for heroin and crack, where less frequent use is associated with a daily consumption level about 70 per cent of that for users in the highest frequency category. However, the AS evidence summarised in Table A5.1 is clearly not reliable as the sole basis for the authors’ assumptions about rates of consumption, so further fragmentary evidence available from other UK and foreign sources has been considered.

Use-quantities: comparison with evidence from other sources

Cannabis

The Forensic Science Service (FSS) has estimated the mean cannabis content of seized “spliffs” as roughly 200mg (Humphries and Joyce, 1982). This figure broadly corresponds with more recent evidence published by IDMU on the basis of a user survey (Atha and Blanchard, 1997). In a day of cannabis use, the average number of spliffs smoked by high frequency (at least three times per week) users has been estimated as 7.7 by Atha and Blanchard (1997).

Recent Australian general-population survey data (Table 4.8) suggests a figure of 5.5 for users who had consumed cannabis in the previous week (not all of whom would be high-frequency users).[29] A reasonable working assumption is that intensive users consume on average 1.2 gm per day of use, with a margin of error of ±0.3gm (equivalent to roughly 4-8 spliffs per use-day). Both IDMU and Australian estimates suggest that non-intensive users have a per episode consumption rate 40 to 50 per cent of that for intensive users. Accordingly, a figure of 0.5 ± 0.4gm for those with a consumption frequency of less than three days per week has been used.

Amphetamines

The authors found little UK evidence relating directly to the average quantity per use-day for amphetamines. However, the Australian evidence (Table 4.9) suggesting a figure of around 1 gram (=10 points) is broadly consistent with anecdotal evidence, and is the central estimate that is used in the analysis.

Ecstasy

The authors were able to find very little systematic evidence on ecstasy consumption per use-day. The Australian evidence in Table 4.9 suggests an average consumption per episode of 1.75 tablets for those who have used ecstasy in the preceding week. For intensive users, an average rate of 2.0 ± 0.25 tablets per use-day seems reasonable, with a lower rate of 1.5 ± 0.25 for less intensive users.

Cocaine

There are many publicly-quoted estimates of typical consumption quantities for cocaine. For example, the Drugscope website[30] suggests that "a gram of cocaine can make between 10 and 20 lines for snorting, depending on its strength, which can last two people anything from a couple of hours to a whole night". This suggests a typical quantity per use-day per person of around 0.8gm for an intensive user. Information from NCIS suggests that one gram per use-day might be a reasonable estimate, giving the user around six lines.[31]

The Australian evidence (Table 4.9) suggests about five lines per use-day for people who have used cocaine within the last week. This group of recent users will include some infrequent users, so an average of 0.8 gm per use-day may be a reasonable estimate for the subset of intensive users. As the baseline assumption, the authors have used the figure 0.8 ± 0.2 for intensive users and 0.5 ± 0.2 for others.

Crack

Little systematic evidence was found for typical rates of crack consumption per use-day. The AS evidence, unreliable as it is, suggests a level only slightly lower than that for powder cocaine. Accordingly, the authors assume 0.7 ± 0.2 gm and 0.4 ± 0.2gm for intensive and non-intensive users respectively. Note that the higher average purity of crack implies a higher true level of consumption for crack users (530mg rather than 450mg under the assumptions for intensive users).

Heroin

The Australian evidence from Table 4.8 suggests that the average usage rate of heroin users is a little over two "hits" per use-day. At an average of 0.3 street grams per hit[32] and typical purity in Australia of around 30 per cent, this would suggest an average level of equivalent

[29] Although the Australian NDSHS survey is an interesting example of questionnaire design, which may avoid some of the apparent confusion over units, it still suffers from the problems of small cell sample sizes and non-response that afflict other comparable surveys elsewhere. It should also be borne in mind that the Australian drug market is nearly as distant geographically as it is possible to get from the UK market.

[30] http://www.drugscope.org.uk/st_about.asp?file=\wip\11\1\1\cocaine_crack.html

[31] Communication from NCIS.

[32] Value suggested by Toni Makkai on the basis of the figures published in *Illicit Drug Data Report 2002-3* (Australian Crime Commission, 2004).

pure consumption of roughly 200mg. per use-day. The Australian equivalent pure rate is thus less than half that implied by self-reported quantities in the AS. Other evidence comes from the Oslo needle-exchange programme (Bretteville-Jensen, 1999), where a sample of registered addicts had an observed consumption rate of 3.7 injections per use-day at an average of 0.143 (street) grams per injection, implying just over 0.5 grams per use-day. Purity was not observed but is believed to be similar to that in the UK on average, so average equivalent consumption of pure heroin would have been around 200mg per use-day. The Swiss heroin maintenance experiment that began in the mid-1990s (MacCoun and Reuter, 2001 page 288-296) gives another indication of usage rates. Under this scheme, addicts had free access to safe heroin supplies at essentially zero marginal cost, resulting in an average daily consumption rate of approximately 500-600 pure mg. This is a plausible estimate of the typical heroin satiation level. Evidence from Britain on the heroin maintenance programme that operated before the 1967 Dangerous Drugs Act showed a daily rate of consumption (not necessarily per use-day) of around 60mg of prescribed pure heroin (Hartnoll *et al.*, 1980). However, doctors tended to put pressure on addicts to keep the prescribed amounts low (Edwards, 1969) and addicts also had access to the illicit market, both as buyers and sellers.

This range of evidence shows very clearly that, even when the sample is heavily trimmed, the AS is not a reliable basis for estimating consumption rates per use-day. The authors' working assumption for average heroin consumption per use-day is 0.48 ± 0.1 and 0.34 ± 0.1 street grams for intensive and non-intensive users respectively. At typical purity rates, this is equivalent to pure consumption rates of 160-240mg. and 100-185mg. respectively and is broadly representative of the range of figures to be found in the literature.

Table 4.8: Average quantity used by user type[a] at typical episode, Australia 2001

	$N^{weighted}$	Mean	Median	Min / Max
Cannabis (cones, bongs, joints)				
Used in last year	764851	2.04	1	0 / 25
Used in last month	319871	2.20	1	0 / 15
Used in last week	840795	5.54	3	0 / 90
Total (any in last year)	*1925517*	*3.60*	*2*	*0 / 90*
Amphetamines (points)				
Used in last year	264797	9.41	10	0 / 400
Used in last month	124853	11.40	10	1 / 400
Used in last week	79915	8.05	4	0 / 200
Total (any in last year)	*469565*	*9.71*	*10*	*0 / 400*
Ecstasy (tablets, pills)				
Used in last year	266131	1.32	1	1 / 11
Used in last month	96046	1.67	1	1 / 7
Used in last week	69534	1.75	1	1 / 8
Total (any in last year)	*431711*	*1.47*	*1*	*1 / 11*
Heroin (hits)				
Used in last year	18061	2.24	2	0 / 10
Used in last month	4478	1.45	1	1 / 4
Used in last week	2862	2.17	2	1 / 5
Total (any in last year)	*25402*	*2.09*	*2*	*0 / 10*
Cocaine POWDER (hits)[b]				
Used in last year	118259	2.57	2	1 / 15
Used in last month	33841	4.13	3	1 / 20
Used in last week	16387	5.51	5	1 / 15
Total (any in last year)	*168487*	*3.17*	*2*	*1 / 20*
Cocaine CRACK (hits)[c]				
Used in last year	3213	4.38	2	1 / 8
Used in last month	1137	4.00	4	4 / 4
Used in last week	–	–	–	–
Total (any in last year)	*4350*	*4.28*	*4*	*1 / 8*

Notes:
Based on answers to the question "*On a day you use [DRUG], on average how many [QUANTITY] do you*

normally have?"
[w] = Data Weighted to WTS8.
[a] = User type is defined by period of last use. Respondents are counted one only.
[b] = Estimates are for persons reporting POWDER cocaine use only.
[c] = Estimates are for persons reporting CRACK cocaine use only.
Source: Adapted by the Australian Institute of Criminology from the National Drug Strategy Household Survey 2001 [computer file], Australian Institute of Health and Welfare

Assumptions

Assumptions need to be made about four things: the prices paid by drug users; the average amount consumed on a day in which consumption occurs; the average purity of each drug; and the average annual number of days' consumption for each category of consumption frequency. In addition to these mean values, assumptions also need to be made about their degree of variation between individuals. Tables 4.10 and 4.11 set out these assumptions. Given the thin evidence available, they necessarily contain a large element of guesswork. The assumptions about average prices are based loosely on the NCIS national averages for the period April 2003-April 2004 (Table 4.5). That source gives little basis for an assumption about the margin of error associated with the NCIS average price. Evidence from elsewhere (for example the analysis by Reuter and Caulkins, 2004, of the US STRIDE data) suggests a wide range of price variation between individual deals, so this suggests a large element of sampling error in the NCIS average, quite apart from other, more systematic, error. In Table 4.9 the authors therefore specify rather wide error margins.

Mean quantities per day of use are based on the range of fragmentary evidence summarised in the previous section. Mean purity is based very loosely on figures summarised in Table 4.6 above. The ranges of variation given in Table 4.9 below are to be interpreted as analogous to 90 per cent confidence intervals. However, they are arbitrary and are used only to give a rough indication of the likely range of estimation uncertainty.

Table 4.9: Assumed values for mean price, quantity and purity, with assumed margins of error

Drug	Mean price [a]	Mean quantity per day of use for intensive users [a,b]	Mean purity [c,d]
Cannabis	£2.50 ± £0.75	1.2 ± 0.4 (intensive) 0.55 ± 0.4 (non-intensive)	-
Amphetamines	£8.50 ± £2.50	1.0 ± 0.2 (all users)	11% ± 4%
Ecstasy	£4.50 ± £1.50	2.0 ± 0.5 (intensive) 1.5 ± 0.5 (non-intensive)	26% ± 10% (65±25mg per tab)
Cocaine	£55.00 ± £10.00	0.8 ± 0.2 (intensive) 0.55 ± 0.2 (non-intensive)	50 ± 6%
Crack	£95.00 ± £15.00	0.7 ± 0.2 (intensive) 0.4 ± 0.2 (non-intensive)	65 ± 7%
Heroin	£60.00 ± £10.00	0.48 ± 0.1 (intensive) 0.34 ± 0.1 (non-intensive)	40% ± 6%

Notes: (a) all quantities in grams, except for ecstasy in tabs; (b) intensive users defined as using at least three times a week; (c) there is no purity factor for cannabis, since it is not usual to measure cannabis quantities in terms of the active constituent (THC); (d) estimate of the mean MDMA content of ecstasy tabs supplied by NCIS

A further important issue is the impact of treatment on consumption rates. Addicts often spend a large (and recently increasing) amount of time in treatment, so this is a potentially important factor. Table 4.10 shows that the AS gives little evidence of a difference in consumption levels per use-day associated with participation in a drug treatment programme.[33] However, there are differences in reported frequency of use, which are significantly lower for arrestees in treatment. This finding suggests that use of reported frequency as a predictor will automatically account for the impact that treatment has in suppressing drug use, so no adjustment is made for the impact of treatment on the rate of use per use-day.

Table 4.10 Distribution of self-reported quantities consumed from AS data
(weighted for survey design and non-response; sample trimmed to give reported quantity between 0.1 and 2 grams; standard errors in parentheses)

Mean consumption per use-day by treatment status	**Heroin**	**Powder cocaine**	**Crack**
currently in treatment	0.95 (0.04)	1.34 (0.36)	0.97 (0.11)
not currently in treatment	0.82 (0.04)	1.11 (0.03)	0.84 (0.03)
Proportion of users consuming at least three times a week	**Heroin**	**Powder cocaine**	**Crack**
currently in treatment	0.09 (0.01)	0.11 (0.01)	0.11 (0.01)
not currently in treatment	0.48 (0.02)	0.28 (0.09)	0.32 (0.07)

A further set of assumptions is needed to translate the verbal descriptions of frequency of use, which are embedded in the OCJS and AS survey questions, into numerical ranges for the number of days' consumption per year. There is an important issue about the interpretation of reported frequencies of use. Survey recall over a period as long as a year and the assessment of "usual" frequencies of consumption are both known to be potentially unreliable. In general specific questions about actual behaviour in a relatively short, recent reference period are to be preferred. This has not been investigated very fully for surveys of drug use. However, Johnson *et al.* (1985, Tables 3-2 and 3-3) give some comparisons between claimed expenditures over the past year and monitored expenditures over a shorter reference period for a group of US heroin users. They found that reported annual consumption was seriously overstated: by a factor of two on average, with the effect being particularly strong for irregular users. Daily users were found to overstate their rate of expenditure by around 47 per cent. The discrepancies were considerably reduced when reported usual daily expenditure rates were compared with reported expenditures per use-day. In their enquiry, heroin-using subjects appeared to be liable to "forget about days without heroin use and, perhaps, days with low heroin use". Thus, the reported AS and OCJS frequencies may be biased upwards. In the absence of relevant UK evidence on this source of bias, reported frequencies are treated as essentially accurate for the baseline assumptions, which are given in Table 4.11. The potential bias arising from this should be an urgent issue for future research.

[33] A regression of quantity on frequency of use and treatment status gives an insignificant treatment coefficient in every case.

Table 4.11: Numerical interpretation of frequency-of-use questions (baseline assumptions)

OCJS: frequency in last year		AS: usual frequency	
Description	*Assumed range days per year*	*Description*	*Assumed range days per year*
Once only	0 – 2	Less often than once a year	0 - 1
More than once, less than once every couple of months	2 – 4	A few times a year	1 – 9
Once every couple of months	4.5 – 9	One or two days a month	9 – 30
Once a month	9 – 21	One or two days a week	30 – 130
Two or three times a month	21-54	Three or four days a week	130 – 234
Once or twice a week	54 – 156	Five or more days a week	234 - 320
Most days	156 – 320		

Estimating the size of the arrestee population

Surprisingly little is known about the population of people who experience arrest in any given period, as opposed to the population of arrest events. The Police National Computer (PNC) system does hold a range of details on an individual basis. It is used operationally by the police and a PNC-based dataset is also used for research purposes by the Home Office. However, the latter only covers arrests resulting in convictions, cautions and warnings or reprimands. Other practical difficulties, such as the possibility of multiple identities for some individuals, make the PNC a problematic source of information for the purposes of this study.

The Home Office publishes an annual analysis of arrests recorded by the police in England and Wales (see Murray and Fiti, 2004), but there is no comparable analysis of the group of people experiencing those arrests. Arrestees are frequently arrested more than once a year, so the count of arrests greatly exceeds the underlying number of arrestees. A method of estimating the size of the arrestee population, using data from multiple sources must therefore be developed and applied.

Data sources

The approach used here combines population data published by the Office for National Statistics (ONS), police arrest data and two sources of survey data: the Crime and Justice Survey and the Arrestee Survey, both of which are described in the section on data sources above.

ONS population estimates

For reasons of comparability with survey data, this study distinguishes between the household and non-household populations and between juveniles (aged 10 to 16) and adults (aged 17 to 65). It is assumed that there are essentially no arrests among the over-65s.[34] There is no published version of official population estimates broken down in this way, so some approximation is required. For mid-2003, official estimates of the whole England and Wales population are available broken down by age (ONS, 2005). 'Experimental' estimates of the household component of these populations are also published (ONS, 2004). Unfortunately, the age breakdown used for these estimates does not coincide with the 10- to 16- and 17+ groups that are relevant for arrest data. The procedure here is to combine the population estimates and the household population estimates and to make some assumptions about the relationship between the two.

[34] Considerably fewer than one per cent of the arrestees recorded by Arrestee Survey interviewers were over 65. Given the wide margins of error in other aspects of the authors' estimates, the decision to ignore the over-65s is innocuous.

To estimate the required totals for the 10- to 16- and 17- to 65-year-old populations, a number of assumptions must be made. The first of these is that the resident population is equal to the household population plus the non-household population. ONS provides figures for the total population and household population and so by subtracting one from the other it is assumed that the non-household population is left. From the single-year age-specific population estimates (ONS, 2005); the totals for various age groups, as presented in the first column of Table 4.12 are known.

Table 4.12: Population estimates

<table>
<tr><th>Age group</th><th>Resident population</th><th>Household population</th><th>Non-household population</th></tr>
<tr><td>0-9 years</td><td>6.243</td><td rowspan="3">11.654</td><td rowspan="3">0.083</td></tr>
<tr><td>10-16 years</td><td>4.819</td></tr>
<tr><td>17 years</td><td>0.675</td></tr>
<tr><td>18-59 years</td><td>29.997</td><td>29.516</td><td>0.484</td></tr>
<tr><td>60-65 years</td><td>3.100</td><td rowspan="2">4.915</td><td rowspan="2">0.033</td></tr>
<tr><td>66-69 years</td><td>1.847</td></tr>
<tr><td>All ages</td><td>52.794</td><td>51.829</td><td>0.965</td></tr>
<tr><td colspan="2">10-16 years</td><td>4.768</td><td>0.051</td></tr>
<tr><td colspan="2">17-65</td><td>33.265</td><td>0.506</td></tr>
</table>

Note: figures in millions, rounded; those in italics are authors' estimates; others are ONS published figures

The second assumption is that the proportion of 10- to 16-year-olds who are household-resident is the same as the proportion for 10- to 17-year-olds. The 2001 Census tells us that 1.06 per cent of the 10 to 17 age group were living in communal establishments and it is assumed that these are coincident with the non-household population. Thus the estimated number of non-household 10- to 16-year-olds is 1.06 per cent of the 4.819m 10- to 16-year-olds in the whole population. This gives figures of 51,275 and 4.768m for the non-household and household populations aged 10 to 16, respectively.

The third assumption is that the proportion of those aged 60 to 65 who are resident in households is the same as the proportion of those aged 60 to 69 (99.34%). From the residential population it is known that there are 3.100m aged 60 to 65 in the residential population, giving an estimate of 3.079m for the household component. Adding to this the ONS estimate of the 18 to 59 household population and the estimate of the 17-year-old household population, gives a figure of 33.265m for the 17 to 65 household population, which in turn implies 0.506m in the non-household 17 to 65 population.

Police arrest data

Arrest is only normally permissible for suspects over the age of ten years. The published arrest figures (Murray and Fiti, 2004) only cover arrests for 'notifiable' offences. These include all indictable and triable-either-way offences and a few closely associated summary offences (see Dodd, *et al.*, 2004, Appendix 2).[35] The main categories of arrest events excluded from this definition are a range of minor summary offences and detention in custody as a place of safety for reasons of intoxication or mental disturbance. There are no figures available for the number of non-notifiable arrests. The published figures relate to England and Wales, for the reference period May 2003 to April 2004. Table 4.13 gives these separately for juveniles (aged from 10 to 16) and adults (aged 17 and over).

[35] Indictable-only offences, such as murder, manslaughter, rape and robbery must be heard at a Crown Court. Either-way offences can be dealt with either by the magistrates or before a judge and jury at the Crown Court, which can impose heavier penalties. Summary offences are less serious cases, such as motoring offences and minor assaults, where the defendant is not entitled to trial by jury.

The Crime and Justice Survey

OCJS respondents were asked if they had ever been arrested by the police (with response categories "no", "yes – once" and "yes – more than once") and if so, how long ago (with the shortest time frame being "within the last year"). Thus the authors were able to divide OCJS respondents into those who experienced arrest within a one-year reference period and those who did not. One potential difficulty is that respondents may have a broad definition of arrest in mind when answering the questions on arrest. There is a possibility that some respondents will describe as arrests events involving temporary detention not involving an explicit suspected offence, or arrest for non-notifiable offences such as motoring offences or minor assaults. A bias in the opposite direction might be induced by non-response, since groups with high offending rates are widely believed to be under-represented in general population surveys like the OCJS.

The Arrestee Survey

AS interviewees are asked for the number of arrests they have experienced in the previous year; before analysis, the authors add one to this count to include the current episode. The analysis is carried out using weighted data on the 7,549 respondents[36] for the year October 2003 to September 2004. The survey weights are designed to adjust for non-proportional survey design, differential non-response and the tendency towards over-sampling of individuals with a high arrest rate. As in the OCJS, there is a strong possibility that AS respondents use a wider definition than the police, when recording their arrest histories, leading to over-estimation of their arrest rates.

From these four sources of data, a set of 11 basic items of information relevant to the measurement problem can be derived. These are set out in Table 4.13.

Table 4.13: Available estimates relevant to measurement problem
(England and Wales; standard errors in parentheses)

Symbol	Definition	Estimate (std. error)
$\hat{N}_1$	Household population aged 10-16 (millions: source CSO + authors' estimates)	4.768 (-)
$\hat{N}_2$	Household population aged 17-65 (millions: source CSO + authors' estimates)	33.265 (-)
$\hat{N}_3$	Non-household population aged 10-16 (millions: source CSO + authors' estimates)	0.051 (-)
$\hat{N}_4$	Non-household population aged 17-65 (millions: source CSO + authors' estimates)	0.506 (-)
$\hat{M}_{13}$	Juvenile arrests (millions: source Home Office)	0.313 (-)
$\hat{M}_{24}$	Adult arrests (millions: source Home Office)	1.017 (-)
$\hat{Q}$	Proportion of non-h/h arrestees (source: AS Oct 2003-Sep 2004)	0.106 (0.007
$\hat{\mu}_2$	Mean arrest rate per h/h arrestee (source: AS Oct 2003-Sep 2004)	3.379 (0.133)
$\hat{\mu}_4$	Mean arrest rate per non-h/h arrestee (source: AS Oct 2003-Sep 2004)	5.741 (0.422)
$\hat{P}_1$	Arrest incidence for h/h juveniles (source: OCJS 2003)	0.0191 (0.0030)
$\hat{P}_2$	Arrest incidence for h/h adults (source: OCJS 2003)	0.0151 (0.0016)

[36] Of whom 6,296 responded to the relevant question.

Method

The authors' method is set out in technical detail in Appendix 6. It distinguishes between four population groups: household residents aged from 10 to 16; household residents aged over 16; non-household residents aged from 10 to 16; and non-household residents aged over 16. This division was chosen because one of the sources (the AS) does not cover the under-17 population and because there is evidence of a significant difference between the arrest rates in the household and non-household populations.

The authors' method has to overcome two major difficulties. One is that the 11 items of available information set out in Table 4.13 are not sufficient to determine the arrestee numbers in each of the four population groups. The second difficulty is that the average arrest incidence in the OCJS sample and the annual frequency of arrest in the AS sample are not consistent with the relatively small total number of arrests recorded by the police, because the police use a narrower definition of ("notifiable") arrest when recording arrests than do individuals when responding to the OCJS and AS. To overcome these difficulties, two further assumptions are made: that OCJS and AS arrest rates are both biased upwards (relative to the police definition) in the same proportion; and that the incidence (but not necessarily annual frequency) of arrest for juveniles is approximately the same in the household and non-household populations. These assumptions are sufficient to allow the construction of a single estimate from the available data.

To give an indication of the degree of reliability of this estimate, the authors use a simulation method to explore the potential impact of: survey sampling error; the measurement error in police arrest statistics and ONS population estimates; and the approximation errors in the *a priori* assumptions. This simulation method allows the authors to construct a margin of error analogous to a conventional confidence interval.

Results

The results of implementing this method are given in Table 4.14. The estimates imply that police reports of notifiable arrest numbers are around 71 per cent of the total number of arrests as perceived and reported by survey respondents. Using the narrow police definition of arrest, the estimated total number of adult arrestees is 396 thousand, implying a mean rate per arrestee of just under 2.9 arrests per year. This implies that around 1.2 per cent of the 17 to 65 England and Wales population experiences police-recorded arrest in a given year. However, a broader definition of arrest is appropriate here, since the Arrestee Survey is not restricted to notifiable offences. On the broader definition, arrest numbers are estimated as 560 thousand adults. Thus the estimate of the adult arrest rate (as perceived by survey respondents) is $P(A) \approx 0.0166$, with a confidence interval of ± 0.0019.

These 95 per cent confidence intervals are quite wide and it turns out that that survey sampling error is the major source of uncertainty in the calculation. However, it should be noted that the systematic error induced by response bias and the authors' additional assumptions is difficult to capture in a confidence interval and may be larger than they have assumed.

Table 4.14: Estimates of the arrestee population

Arrestee population	**Estimate** (thousands)	**95% confidence interval** (all sources of error)
Narrow definition of arrest (as recorded by police)		
Juveniles aged 10-16	64.9	± 21.0
Adults aged 17-65	396.2	± 46.2
P(A) = Arrest proportion for adults aged 17-65	1.173%	± 0.14%
TOTAL	**461.1**	**± 44.9**
Broad definition of arrest (as perceived by survey respondents)		
Juveniles aged 10-16	91.8	± 29.7
Adults aged 17-65	560.2	± 65.3
P(A) = Arrest proportion for adults aged 17-65	1.659%	± 0.19%
TOTAL	**652.0**	**± 63.5**

Extension to the UK level

There are few data sources for Scotland and Northern Ireland comparable to those used by the authors to construct the estimates for England and Wales. Although there are general-population surveys giving information on drug use in Scotland and Northern Ireland, there is currently no analogue of the England and Wales Arrestee Survey, so the authors' multi-survey methodology is not applicable outside England and Wales.

There are other data problems. Although population figures are available for Scotland and Northern Ireland, the authors have been unable to find data on the aggregate number of arrests in Scotland. However, for Northern Ireland, figures are available for the number of arrests under the Police & Criminal Evidence (NI) Order. In 2003/04, this total was 27,221. Expressing this as a proportion of the NI 17 to 65 population gives a ratio of 0.0183. The comparable ratio for England and Wales is 0.0344, suggesting that the arrest rate in Northern Ireland is only half that in England and Wales.

In the absence of adequate data for Scotland and Northern Ireland, the authors use a simple method of extending their estimates to a UK basis. The method, which is detailed in Appendix 7, takes account of differences in the age structures of the Scottish and Northern Irish populations relative to England and Wales. The authors assume (on the basis of the low volume of arrests in Northern Ireland) that the arrest rate there is just over half that in England and Wales (0.0183/0.0344 = 53.2%), whilst the arrest rate in Scotland is identical to that in England and Wales. They further assume that mean drug consumption rates per head for adult non-arrestees, juveniles and arrestees are the same in each country of the UK.[37]

The outcome of this necessarily crude method is a UK estimate of total market size roughly 13.5 per cent larger than that for England and Wales, whereas the UK population is 15.0 per cent larger.

[37] The authors did not have access to drug use survey data for Northern Ireland and Scotland for the relevant period. Past survey evidence from Northern Ireland, in particular, suggests a lower prevalence rate than in England and Wales, so there is likely to be a consequent (small) degree of over-estimation in the report's UK figures.

Baseline estimates

The result of applying the method outlined above is summarised in Table 4.15. The authors' baseline estimate for the size of the UK illicit drugs market is £5.3bn. There are a number of points to be made about these estimates.

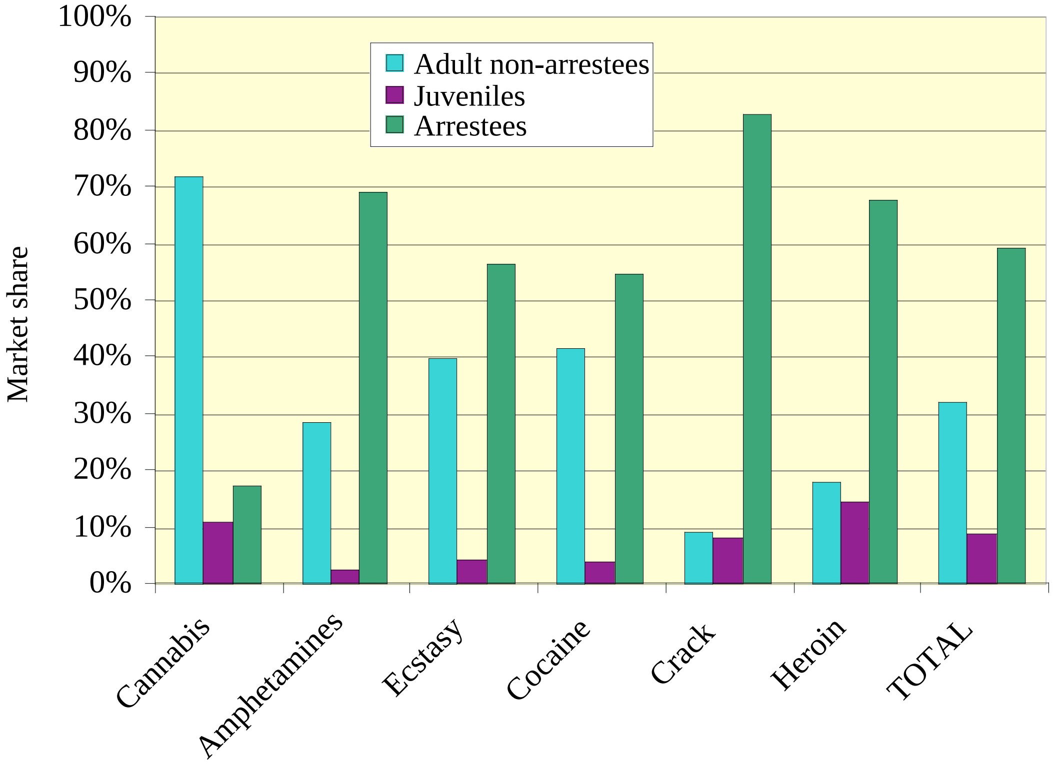
- No adjustment has been made for under-reporting by survey respondents. If made, such an adjustment would increase the estimates, with a larger impact on “hard” than “soft” drugs.
- The Schools Survey has been used as the sole basis for estimating the drug use of juveniles. This gives a substantially larger result than would use of OCJS data. The authors believe that the SS interview context is more conducive to accurate responses than that of the OCJS and that the higher prevalence rates observed in SS data are more plausible than the low rates suggested by the OCJS. A different view would lead to a reduction in the market size estimates. Note that the authors have not attempted to adjust the OCJS estimates for adults to bring them in line with SS data.
- A literal interpretation of responses to survey questions on frequency of use has been used. It could be argued that these responses are biased towards over-statement of consumption, so an *ad hoc* adjustment for this would have the effect of reducing estimated market size.
- The authors' method involves the assumption that, after dealing appropriately with survey non-response, potential offenders and drug users who happen not to be arrested in the reference year are adequately represented in the analysis of general-population survey data (the OCJS). If this is not so then the estimates will tend to be biased downwards. The authors believe that this is most likely to be the case for cocaine.
- The error margins quoted in Table 4.15 are a very rough attempt to quantify the large range of uncertainty surrounding the estimates. They should not be regarded as precise statements of reliability.

Table 4.15: Baseline estimates of market size for England and Wales and the UK for 2003/4

	England and Wales			UK		
	Aggregate street quantity (tonnes)	**Aggregate pure quantity (tonnes)**	**Aggregate expenditure (£million)**	**Aggregate street quantity (tonnes)**	**Aggregate pure quantity (tonnes)**	**Aggregate expenditure (£million)**
Cannabis	360.33 *±135.81*	360.33 *±135.81*	900.8 *±372.4*	412.41 *±155.44*	412.41 *±155.44*	1031.0 *±432.5*
Ampheta-mines	32.68 *±17.33*	3.60 *±2.31*	277.8 *±72.9*	36.70 *±19.46*	4.04 *±2.60*	312.0 *±81.9*
Ecstasy (millions of tabs)	52.79 *±23.84*	13.72 *±8.14*	237.5 *±76.2*	59.52 *±26.88*	15.47 *±9.18*	267.8 *±85.9*
Powder cocaine	15.7 *±12.17*	7.85 *±6.16*	863.4 *±237.1*	17.70 *±13.72*	8.85 *±6.94*	973.3 *±267.3*
Crack	13.79 *±11.76*	8.96 *±7.71*	1309.8 *±348.9*	15.58 *±13.29*	10.13 *±8.71*	1480.4 *±394.29*
Heroin	17.60 *±13.14*	7.04 *±5.32*	1055.9 *±199.2*	20.11 *±15.02*	8.04 *±6.13*	1206.7 *±227.65*
Total market value (£bn)			**4.645** ***±1.154***			**5.271** ***±1.310***

The estimated composition of the illicit drugs market by population is summarised in Figure 4.2. For the market as a whole, the authors estimate that 59 per cent of the market (by expenditure) is attributable to adult arrestees, 32 per cent to other adults and nine per cent to juveniles. Cannabis, cocaine and ecstasy have the highest penetration of the general adult population, whilst crack, heroin and amphetamines are dominant in the arrestee population. There are several important things to note, however, when interpreting Figure 4.2. Firstly, that there are wide margins of error for expenditure shares by drug within population group, particularly for juveniles. Secondly, the picture would be quite different in quantity, rather than expenditure, terms. Thirdly, estimated heroin use is much more substantial than might be expected on the basis of its relatively low prevalence (particularly for juveniles), because of the very high average frequency of use reported by heroin users.

Figure 4.2: Estimated composition of the market by population group (shares of expenditure for each drug and for the total market)



The measurement approach involves a mix of statistical survey-based estimation and *a priori* assumption only loosely related to external evidence. The standard theory of confidence intervals does not adequately cover this sort of complex process. The authors have developed a procedure (set out in Appendix 3, section A3.4) that takes account of both sources of uncertainty by using a bootstrap algorithm to estimate the extent of sampling error and a Monte Carlo simulation to allow for the uncertainty associated with the *a priori* assumptions. The resulting margins of error are intended to be interpreted in much the same way as a conventional 90 per cent confidence interval, but they should only be seen as rough indicators of reliability. In practice, these error margins are dominated by the uncertainty associated with *a priori* assumptions, rather than survey sampling error. Table 4.15 shows these very wide margins of error, which emphasise the basic problem of measurement in this field: that not enough is known about some aspects of the drug market, in particular, prices and quantities per use-day.

One, possibly encouraging, feature of the error margins is the fact that they are dominated by the uncertainty arising from the authors' assumptions about mean prices, usage rates per use-day, etc. For example, the mean per capita volume and value of cannabis consumption in the juvenile population are estimated to be 8.1 ± 4.1 grams and £20.26 ± £12.36. If no allowance is made for the uncertainty associated with price, quantity and purity assumptions, then these margins of error shrink to 35 per cent and 18 per cent respectively of their original width. This shows that further research, aimed at improving the available evidence on prices, purity and usage patterns, could make possible a big improvement in the precision of market size estimates.

Adjustments for under-reporting

Misreporting is always a potential difficulty in survey data, particularly when the subject of the survey is illicit or anti-social behaviour. There is a definite link between the accuracy of survey responses and the survey context and design of questions. For example, experience from surveys of smoking and drinking behaviour suggests that the specificity and design of questions has a large influence on mean responses (Pierce *et al.*, 1987; Lemmens *et al.*, 1992; Klesges *et al.*, 1995; Knibbe and Bloomfield, 2001; Stockwell *et al.*, 2004). However, the authors' is a secondary analysis and they have no control over the design of questions in the OCJS, AS and SS.

It is rarely possible to validate individual survey responses directly, but some drug surveys do allow this by providing both interview responses and independent tests for drug residues in biological samples. There is a substantial US literature comparing self-reported drug use with bio-assay test results. Magura and Kang (1996) made a meta-analysis of 24 studies published since 1985, finding only moderate agreement between self-reports and drug tests. More recently, Lu *et al.* (2001) found that fewer than 50 per cent of arrestees who tested positive for methamphetamines, cocaine or opiates in the US ADAM survey reported drug use in the relevant period. Mieczkowsli *et al.* (1998) found still higher rates of under-reporting for a small sample of homeless adults in New York in 1994. Less is known about misreporting rates for the UK, but there is some evidence from the NEW-ADAM survey (Bennett, 2000), suggesting some under-reporting but at lower rates than some of those quoted for the USA.

The aim here is to analyse NEW-ADAM and AS data, comparing self-reports with the results of drug tests carried out on urine or saliva samples. The details of this analysis and the way it is used to modify the market size estimate, are set out in Appendix 8. Here the authors summarise the main features of the method and present the results. The assumptions on which the adjustment method is based are as follows:

- The authors make no adjustment for under-reporting in the juvenile and adult non-arrestee populations since, without independent drug tests, there is no convincing basis for such an adjustment.
- The authors assume that misreporting takes only one form: the complete denial of consumption during the last year, rather than understatement of the intensity of consumption. This assumption also rules out any significant degree of over-reporting (such as bragging by respondents). This seems reasonable, given the private nature of CASI interviewing and the inclusion in the OCJS and SS of questions (about a non-existent drug, semeron) to detect false consumption claims. The authors also assume that, conditional on other observed personal characteristics, misreporters and accurate reporters share the same distribution of frequency-of-use.
- The authors assume there is a non-zero probability of a false positive test result (in the sense of a positive result for someone who has not consumed the drug during the last year) and this probability is fixed *a priori* at 2.5 per cent. Without this assumption, the authors' statistical model cannot be identified. Models which assume perfect accuracy in drug testing tend to produce implausibly high under-reporting probabilities for some groups of individuals, which generates instability in the adjustment process.
- The model of misreporting behaviour is estimated for cocaine and opiates using AS data and additionally for cannabis using NEW-ADAM data. The estimated model is then used to construct predicted consumption levels for those who self-report no consumption in the last year, using all available AS data about the individual.

At this stage, the authors regard these adjustments for under-reporting as experimental and not to be regarded as a definitive estimate. The issue of misreporting is an important topic which requires future research.

The results of carrying out this adjustment are given in Table 4.16 below. The impact of the adjustment is significant but moderate. The proportionate adjustment in the estimate of aggregate quantity consumed by the arrestee population ranges from 5.5 per cent for cannabis to 17.9 per cent for heroin. This produces a seven per cent increase in the estimate of total market size.

Table 4.16: Adjustments for under-reporting

	Unadjusted street quantity (England & Wales arrestees)	**Adjusted street quantity** (England & Wales arrestees)	**Un-adjusted total pure quantity** (UK)	**Adjusted total pure quantity** (UK)	**Un-adjusted value** (UK, £m)	**Adjusted value** (UK, £m)
Cannabis	62.7 ±18.9	66.1 ±25.6	412.4 ±94.5	416.2 ±96.5	1031.0 ±262.9	1040.6 ±262.6
Ampheta-mines	22.6 ±5.9	23.9 ±10.7	4.0 ±2.6	4.2 ±2.8	312.0 ±81.9	324.3 ±110.4
Ecstasy (millions of tabs)	29.6 ±7.5	31.4 ±6.1	15.5 ±9.2	16.0 ±9.2	267.8 ±85.9	276.6 ±90.2
Powder cocaine	8.6 ±2.8	9.4 ±5.1	8.8 ±6.9	9.3 ±7.4	973.3 ±267.3	1024.1 ±318.8
Crack	11.4 ±6.5	12.8 ±6.2	10.1 ±8.7	11.1 ±8.6	1480.4 ±394.3	1624.5 ±374.7
Heroin	11.9 ±3.1	14.0 ±3.1	8.0 ±6.1	9.0 ±6.1	1206.7 ±227.7	1349.0 ±265.1
Total market value					**5,271** ***±1,310***	**5,639** ***±1,458***

Comparison with earlier estimates

Table 4.17 compares the baseline estimates with those made by Bramley-Harker (2001). It should be noted that the Bramley-Harker estimates relate to the year 1998 and that there has been a substantial price fall since that time. In that sense, the estimates are not directly comparable.

Table 4.17: Comparisons of new estimates with Bramley-Harker estimates for 1998

	Average price (per gram or tab)	Street quantity (tonnes or million tabs)	Market size (£m)
Bramley-Harker estimates for 1998			
Cannabis	£3.29	401.1	1,318
Amphets	£10.00	21.6	216
Ecstasy	£11.00	22.4	246
Cocaine	£77.00	3.8	296
Crack	£100.00	15.2	1,520
Heroin	£74.00	26.0	1,922
England & Wales total (£m)			**£5,519**
UK total (£m)			**£6,595**
New estimates for 2003/04			
Cannabis	£2.50	360.3	900.8
Amphets	£8.50	32.7	277.8
Ecstasy	£4.50	52.8	237.5
Cocaine	£55.00	15.7	863.4
Crack	£95.00	13.8	1309.8
Heroin	£60.00	17.6	1055.9
England and Wales total (£m)			**£4,645**
UK Total (£m)			**£5,271**

The major structural differences between the authors' estimates and the earlier Bramley-Harker study are the lower aggregate volumes and values for cannabis (respectively 10% and 32% lower) and heroin (32% and 45% lower); and the much larger volume and value of the powder cocaine market (309% and 192% higher). Even allowing for the large margins of error in these estimates, these are large differences.

Comparison with other evidence

Drug seizures

The volume and occurrence of drug seizures can be used as an indicator of the relative sizes of the markets for different drugs. Statistics on seizures made by HMRC and by the police and other authorities are published annually. The most recent figures on police seizures (Ahmad and Mwenda, 2004) relate to 2002 and are reproduced in Table 4.18, together with those for 2001. The authors use data on police rather than HMRC seizures because the former are closer to street level and should reflect better the relative sizes of markets at the level of the consumer.

Table 4.18: Drug seizures made by the police and other authorities (excluding HMRC) in 2001 and 2002

	Amount seized by police 2001 (tonnes)	Amount seized by police 2002 (tonnes)	Number of police seizures 2001	Number of police seizures 2002
Cannabis	55.999	64.010	3,410	4,540
Amphetamines	0.627	0.588	70	70
Ecstasy *	2.350	1.062	10,320	8,210
Cocaine	1.024	0.222	5,470	5,240
Crack	0.034	0.044	3,650	4,230
Heroin	2.802	1.689	18,080	15,270

* Quantity of ecstasy in thousands of doses

Normal use-quantities vary considerably across drug categories, so the authors rebase the quantity data before calculating the total market shares shown in Table 4.19 and Figure 4.3. The volume data in Table 4.19 and Figure 4.3 is first converted to use-units by dividing each volume by the assumed quantity per use-day (for high-frequency users) given in Table 4.10.

Table 4.19: Market share indicators

	Number of police seizures 2001	Number of police seizures 2002	Amount seized by police 2001*	Amount seized by police 2002*	Market size estimate (quantity share)
Cannabis	4.9%	7.0%	83.3%	91.2%	69.2%
Amphetamines	0.1%	0.1%	1.1%	1.0%	7.4%
Ecstasy	9.4%	10.2%	2.8%	1.2%	6.0%
Cocaine	11.8%	12.2%	2.3%	0.5%	4.5%
Crack	9.0%	11.2%	0.1%	0.1%	4.5%
Heroin	64.8%	59.2%	10.4%	6.0%	8.4%

* Quantities converted to use-units (“hits”) before calculating share

The correspondence between the composition of the authors’ market size estimates and the composition of police seizures is reasonable. Cannabis dominates the market, with roughly a 70 per cent estimated share and accounts for around 80 to 90 per cent of seized amounts. The next largest in use-volume terms is heroin with over eight per cent of the market and from six to ten per cent of seizures. The remaining drugs are small in volume share terms and considerably smaller still in terms of seized amounts.

__Figure 4.3: Composition of the drugs market in adjusted volume share terms: market estimates compared to police seizures__

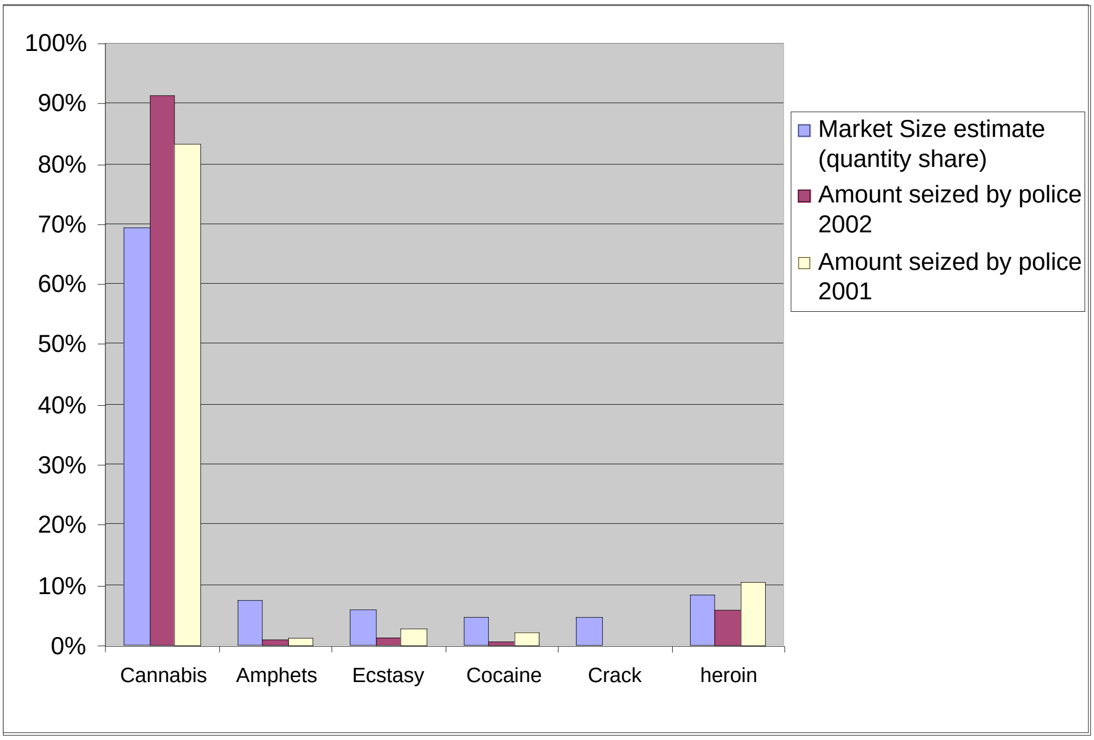
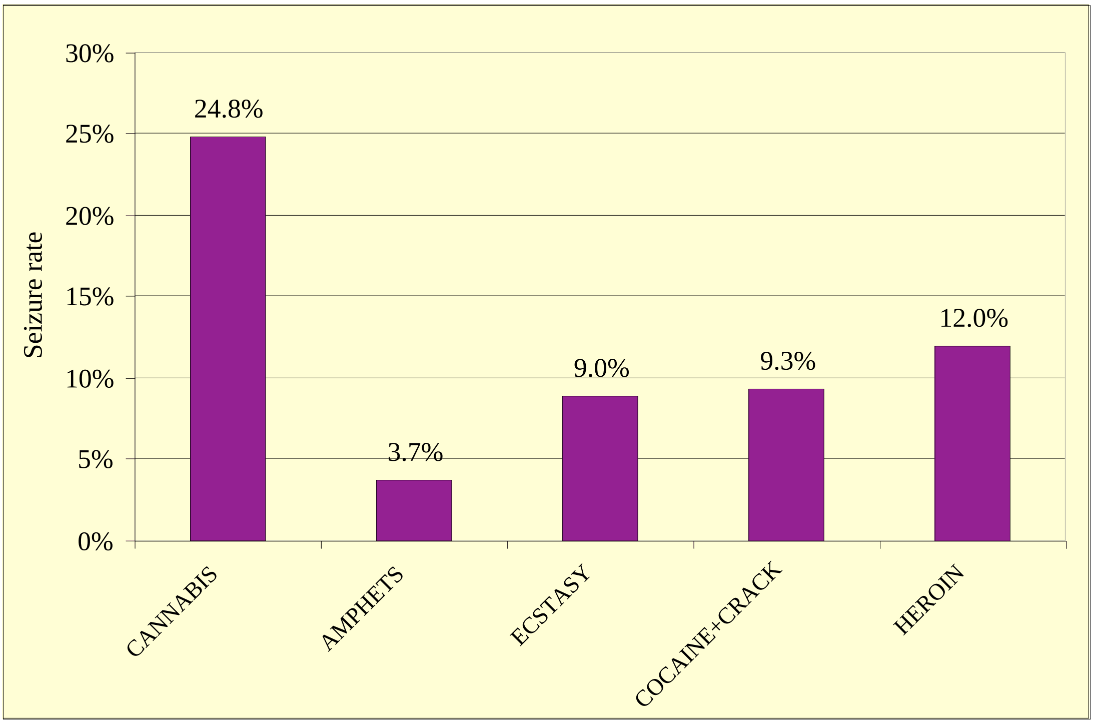


Figure 4.4 compares the authors' market size estimates, in terms of street quantities, to total seizures by all authorities, including police and HMRC. This is presented in the form of an implied seizure rate, calculated as estimated volume consumed as a proportion of the amount consumed plus the amount seized. In this calculation, the authors have combined cocaine and crack, since much of the powder cocaine seized by HMRC would have been converted into crack form before reaching the retail market. For ecstasy, cocaine and heroin, the seizure rates are in line with the assumptions made by Groom *et al.* (1998). The amphetamine seizure rate of four per cent is low, whilst the 25 per cent implicit seizure rate for cannabis is high. There is a need for caution in the latter case, since no purity analysis is available for cannabis seizures and some cannabis seizures will consist of the drug after preparation for smoking, mixed with a larger volume of tobacco. Thus seized amounts may overstate true quantities. It is also worth noting that the penalty for possession of a small amount of cannabis[38] is often nothing more than a caution or informal warning, so there is less incentive for users to take concealment measures to prevent detection. Thus equilibrium consumption behaviour might be expected to involve a higher seizure rate for cannabis than for 'harder' drugs.

[38] Although the maximum penalty is high, the average penalty for moderate trafficking in cannabis is also low relative to other drugs.

Figure 4.4: Implied approximate seizure rates (volume seized by all authorities as a percentage of estimated market size + seized volume)



Conclusions and implications for further research

The authors' measurement attempt has some weaknesses that are impossible to avoid with currently available data resources. There are consequently several recommendations for future improvement.

Survey questionnaire design

Comparability of survey questions

The principal surveys the authors have used (the OCJS, SS and AS) were designed independently, without consideration for their joint analysis. Consequently there are differences in questionnaire design that cause serious analytical difficulties. The main problems here are that the SS has no question on the frequency of use for each drug and that the frequency questions asked in the OCJS and AS are not comparable. The OCJS question relates to drug use during the previous year, while the AS frequency question relates to "usual" behaviour. It is recommended that future reviews of Home Office and HMRC data requirements should consider the full range of survey resources and emphasise the advantages of comparability of questionnaire content.

Design of survey questions

There are many potential pitfalls stemming from the design of existing survey questions on drug use. Firstly, survey questions on frequency of drug use relate either to a general pattern of use during the previous year (OCJS) or to a usual pattern of use over an unspecified period (AS). Such questions are widely felt to be unsatisfactory and there is some evidence that they may systematically misrepresent the actual frequency of use during the one-year measurement period. Most survey designers would prefer questions which ask for factual information relating to a specific, fairly short, recent reference period. However, there are complications here, since recent use may be directly linked to the arrest event and because respondents may be more reluctant to report recent drug abuse accurately.

Secondly, the responses to 2003/4 AS questions on the quantities of, and expenditures on, specific drugs are suspect because of the presence of large outliers, which may be the result of confusion between different units of quantity and between purchases and consumption. Recent revisions to the AS questionnaire may overcome this problem to some extent.

Thirdly, if surveys are to continue concentrating mainly on usual frequency of use, it is important that there is direct information on the quantity of drugs used per day of use. The Australian National Drug Strategy Household Survey (Table 4.9) may offer a good way of phrasing these questions. In the absence of this information, the authors have been forced to use a degree of guesswork.

The authors' view is that there may be scope for significant improvements in the design of the existing survey questions on drug use. They recommend that this possibility is reviewed, paying attention to the wide range of question designs used in different countries and to the large body of evidence on question design for surveys of smoking and drinking (Pierce *et al.*, 1987; Lemmens *et al.*, 1992; Klesges *et al.*, 1995; Knibbe and Bloomfield, 2001; Stockwell *et al.*, 2004).

Drug testing

Comparison of self-reports with saliva test results is a potentially useful basis for adjusting for the effects of under-reporting. The authors have demonstrated one way of doing this. However, the scope for this is very limited at present. The context and design of arrestee surveys is so different from that of general population surveys that adjustment of general-population results based on arrestee misreporting rates is not valid. However, there do exist general population surveys that take biological samples (the Health Survey for England, for example) and there would be scope for investigating misreporting if the HSE questionnaire and bio-assay were extended to cover drug use. The AS does not test for cannabis and amphetamine use[39] and this limits the scope for investigating misreporting. The introduction of amphetamine tests into the AS is worth considering.

Surveys of juveniles

A particularly important issue raised by this work is the large difference between surveys in the reported levels of drug use among juveniles. The authors believe that the Schools Surveys are likely to be the most accurate source but this should be investigated. This may be an area where qualitative follow-up of survey respondents could contribute effectively to survey design.

Price data

Currently available data on street prices available from NCIS are not adequate as a basis for research of this kind. In comparison with the USA, where the Drug Enforcement Agency's STRIDE price monitoring system operates (Reuter and Caulkins, 2004), the system for monitoring UK drug price data seems underdeveloped. If the monitoring of market size (and the impact of interdiction activity) is to continue as a policy concern, a high priority should be attached to improving the process of price measurement. The authors note that a new initiative (project PARAMOUNT) has been developed by NCIS, on behalf of Concerted Inter-Agency Drugs Action group (CIDA), as a means for improving the quality of information available on Class A drug prices and purities and enable systematic monitoring of changes. The project is currently being piloted by a number of UK police forces, as well as NCIS, NCS and HMRC. National rollout is planned for later this year.

Scotland and Northern Ireland

There is a striking disparity between the data resources available for England and Wales and those available for Scotland or Northern Ireland. There are major differences in the design of the available surveys and a particular problem is the absence of a current arrestee survey in

[39] Saliva-based tests are not generally regarded as suitable for cannabis.

Scotland and Northern Ireland, and of police arrest statistics in Scotland. As a consequence of this, the authors have only been able to produce clear estimates for England and Wales and they have had to resort to fairly crude extrapolation to extend this to a UK basis. If UK-wide estimates are required, then it is important to fill the gaps in statistical coverage of Northern Ireland and Scotland.

Arrest statistics

It is surprising how little is known statistically about the arrest process. In England and Wales there are data on the volume of notifiable arrests but no systematic data on the numbers of non-notifiable arrests. There is also no data on the number of individuals who experience arrest during the course of the year – a figure which is required when combining surveys of arrestees with surveys of the general population. The authors recommend that a review be undertaken to assess the feasibility of measuring regularly the volume of non-notifiable arrests[40] and the number of individuals who experience one or more arrests during the course of the year.

Monitoring change over time

A potential policy use for this measurement effort is in monitoring change over time in the scale of drug abuse. There is no difficulty in repeating this analysis in future years, provided the core data sources (the OCJS, SS and AS) continue to be conducted in the future. However, the margins of error for the authors' estimates are large and will become larger still for estimated changes in market size. There have been proposals for this type of monitoring to be done on a quarterly basis. The authors would advise against this, since the estimated quarterly changes will be completely dominated by sampling error. Another proposal has been to use annual totals, rolling on a quarterly basis. If the sequence of quarterly market size estimates is represented as S_1, S_2, S_3, ..., then the rolling annual analysis would produce a sequence of annual figures:

$$(S_1+S_2+S_3+S_4),\ (S_2+S_3+S_4+S_5),\ (S_3+S_4+S_5+S_6),\ \ldots\ etc.$$

This does not solve the problem of imprecise measurement of change over time – for example, the difference between the first two annual totals is:

$$(S_2+S_3+S_4+S_5) - (S_1+S_2+S_3+S_4) = S_5 - S_1$$

This is still the difference between two quarterly figures and would again be dominated by sampling error.

The authors recommend that aggregate market size estimates of this type should be used in sequence as an annual time series, interpreted only as a broad indicator of the trend in drug use over a long period of time (in practice, at least five years). Very little attention should be paid to the precise value of the market size measure in a particular year, or its change between any particular two years.

[40] An attempt at this has already begun in a Home Office sponsored study by NatCen, in parallel with the Arrestee Survey.

References

Ahmad, M. and Mwenda, L. (2004) *Drug Seizure and Offender Statistics, United Kingdom 2001 and 2002*. London: Home Office Statistical Bulletin 08/04. (http://www.homeoffice.gov.uk/rds/pdfs04/hosb0804.pdf)

Atha, M. J. and Blanchard, S. (1997) Regular users. Self-reported drug consumption patterns and attitudes towards drugs among 1333 regular cannabis users, Independent Drug Monitoring Unit (http://www.idmu.co.uk/pdfs/ru1.pdf).

Australian Crime Commission (2004) *Illicit Drug Data Report 2002-03*. http://www.crimecommission.gov.au/content/publications/iddr_2002_03/1-7.pdf

Australian Institute of Health and Welfare, Canberra (2002) *2001 National Drug Strategy Household Survey*. AIHW cat. No. PHE 35 (http://www.aihw.gov.au/publications/phe/ndshs01/ndshs01-020717.pdf)

Australian Institute of Health and Welfare, Canberra (2005) *2004 National Drug Strategy Household Survey*. AIHW cat. No. PHE 57 (http://www.aihw.gov.au/publications/phe/ndshs04/ndshs04.pdf)

Aye Maung, N. (1995) *Survey design and interpretation of the British Crime Survey* in Walker, M. (ed.) Interpreting Crime Statistics. Oxford: Oxford University Press.

Badillo Bautista, C., Bryan, M. and Burton J. (2005) *Measuring Illicit Drugs Consumption: A comparison of data sets*, unpublished research paper, ISER, University of Essex.

Bennett, T. H. (2000) *Drugs and Crime: The Results of the Second Development Stage of the NEW-ADAM Programme*. London: Home Office Research Study 205.

Boreham, R., Fuller, E. and Hills, A. (2005) *The Arrestee Study: Annual Report October 2003-September 2004*. London: National Centre for Social Research.

Bramley-Harker, E. (2001) *Sizing the UK Market for Illicit Drugs*. RDS Occasional Paper no. 74. London: Home Office.

Brand, S. and Price, R. (2000) *The economic and social costs of crime*. Home Office Research Study 217. London. Home Office. (http://www.homeoffice.gov.uk.rds/pdfs/hors217.pdf)

Bretteville-Jensen, A. L. (1999) Gender, heroin consumption and economic behaviour, *Health Economics* 8, 379-389.

Caulkins, J. P. and Pacula, R. (2002) L. Marijuana Markets: Inferences from reports by the Household Population, mimeo.

Caulkins, J. P. and Reuter, P. (1998) What Price Data Tell Us about Drug Markets, *Journal of Drug Issues*, 28: 593-612, Summer 1998.

Chaloupka, F., Grossman, M. and Tauras, J. (1998) *The Demand for Cocaine and Marijuana by Youth*, NBER working paper, No. 6411, February 1998.

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

Dodd, T., Nicholas, S., Povey, D. and Walker, A. (2004) *Crime in England and Wales 2003/2004*. Statistical Bulletin 10/04. London: Home Office.

Dubourg, R., Hamed, J. and Thorns, J. (2005) *The Economic and social costs of crime against individuals and households 2003/04*. Home Office Online Report 30/05. London. Home Office. (http://www.homeoffice.gov.uk/rds/pdfs05/rdsolr3005.pdf)

Edwards, G. (1969) The British approach to the treatment of heroin addiction, *Lancet* 1, 768-772.

Frischer, M., Heatlie, H. and Hickman, M. (2004) *Estimating the prevalence of problematic and injecting drug use for Drug Action Team areas in England: a feasibility study using the Multiple Indicator Method*. Online Report 34/04. London: Home Office.

Fountain, J., Bashford, J., Winters, M. and Patel, K. (2003) *Black and minority ethnic communities in England: a review of the literature on drug use and related service provision*. London: National Treatment Agency for Substance Misuse and the Centre for Ethnicity and Health.

Godfrey, C., Eaton, G., McDougall, C. and Culyer, A. (2002) *The economic and social costs of Class A drug use in England and Wales, 2000.* Home Office Research Study 249. London. Home Office.
(http://www.homeoffice.gov.uk/rds/pdfs2/hors249.pdf)

Godfrey, C., Stewart, D. and Gossop, M. (2004) *Economic analysis of costs and consequences of the treatment of drug misuse: two year outcome data from the National Treatment Outcome Research Study*, Addiction, June.

Groom, C., Davies, T. and Balchin, S. (1998) Developing a methodology for measuring illegal activity for the UK national accounts, *Economic Trends*, 536, 33-71.

Hamlyn, B., Maxwell, C., Hales, J. and Tait, C. (2003) *2003 Crime and Justice Survey (England and Wales) Technical Report*. London: Home Office.

Hartnoll, R. L., Mitchelson, M. C., Battersby, A., Brown, G., Ellis, M., Fleming, P. and Hedley, N. (1980) Evaluation of Heroin Maintenance in Controlled Trial, *Archive of General Psychiatry* 37, 877-884.

Hay, G., Gannon, M., McKeganey, N., Hutchinson, S. and Goldberg, D. (2005) *Estimating the National and Local Prevalence of Problem Drug Misuse in Scotland*. Online Report, Available:
http://www.drugmisuse.isdscotland.org/publications/abstracts/prevalence3.htm [31/10/2006].

Hickman, M., Higgins, V., Hope, V. and Bellis, M.A. (2004) *Estimating prevalence of problem drug use: multiple methods in Brighton, Liverpool and London*. Online Report OLR 36/04. London: Home Office.

Hook, E. and Regal, R. (1997) Validity of methods for model selection, weighting for model uncertainty and small sample adjustment in capture-recapture estimation. *American Journal of Epidemiology*, 145, 1138-1144.

Humphries, I. J. and Joyce, J. R. (1982). A survey of the cannabis content of unsmoked reefer cigarettes, *Journal of the Forensic Science Society* 22, 291-292.

Johnson, B. D., Goldstein, P. J., Preble, E., Schmeidler, J., Lipton, D. S., Spunt, B. and Miller, T. (1985). *Taking Care of Business. The Economics of Crime by Heroin Abusers*. Lexington, Mass: Lexington Books.

Klesges, R. C., Debon, M. and Ray, J. W. (1995) Are self-reports of smoking rate biased? Evidence from the Second National Health and Nutrition Examination Survey, *Journal of Clinical Epidemiology* 48, 1225-1233.

Knibbe, R. A. and Bloomfield, K. (2001) Alcohol consumption estimates in surveys in Europe: comparability and sensitivity for gender differences, *Substance Abuse* 22, 23-38.

Layne, M., Rhodes, W. and Johnston, P. (1999) Estimating Cocaine Flow: The Sequential Transition and Reduction (STAR) Model, 1996-1998, Abt Associates Inc., report submitted to the Office of National Drug Control Policy, December, 1999.

Lemmens, P. H., Tan, E. S. and Knibbe, R. A. (1992) Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. *Journal of Studies on Alcohol* 53, 476-486.

Lu, N T., Taylor, B. G. and Riley, K. J. (2001) The validity of adult arrestee self-reports of crack cocaine use, *American Journal of Alcohol Abuse* 27, 399-419.

Magura, S. and Kang, S. Y. (1996) Validity of self-reported drug use in high risk populations: a meta-analytical review, *Substance Use and Misuse* 31**,** 1131-1153.

MacCoun R., and Reuter, P. (2001) *Drug War Heresies. Learning from Other Vices, Times and Places*. Cambridge: Cambridge University Press.

Mieczkowski, T., Newel, R. and Wraight, B. (1998) Using hair analysis, urinalysis and self-reports to estimate drug use in a sample of detained juveniles, *Substance Use and Misuse* 33, 1547-1567.

Millar, T., Gemmell, I., Hay, G. and Donmall, M. (2004) *The dynamics of drug misuse: assessing changes in prevalence*. Online Report 35/04. London: Home Office.

Murray, L. and Fiti, R. (2004) *Arrests for Notifiable Offences and the Operation of Certain Police Powers under PACE. England and Wales, 2003/4*. Statistical Bulletin 18/04. London: Home Office.

OAS (2004) *Overview of Findings from the 2003 National Survey on Drug Use and Health: Department of Health and Human Services Substance Abuse and Mental Health Services Administration,* Office of Applied Studies, Rockville, MD.

ONDCP (2001) *What America's Users Spend on Illegal Drugs 1988-2000*. Washington, D.C. Office of National Drug Control Policy, 2001.

ONS (2004) *Mid-2003 Estimates of Household Population; Broad Age Groups and Sex for Government Office Regions in England and Wales (experimental)*. Office for National Statistics online dataset (release 29/9/2004): http://www.statistics.gov.uk/statbase/ssdataset.asp?vlnk=8570&More=Y.

ONS (2005). *England and Wales; Estimated Resident Population by Single Year of Age and Sex; Mid-2003 Population Estimates*. National Statistics online dataset (release 20/1/2005: http://www.statistics.gov.uk/statbase/ssdataset.asp?vlnk=8556&More=Y).

Pacula, R. L., Grossman, M., Chaloupka, F. J., O'Malley, P. M., Johnston, L. D. and Farrelly, M. C. (2000) Marijuana and Youth, NBER Working Paper 7703, National Bureau of Economic Research, Cambridge, MA.

Pierce, J. P., Dwyer, T., DiGiusto, E., Carpenter, T., Hannam, C., Amain, A., Yong, C., Sarfaty, G., Shaw, J. and Burke, N. (1987) Cotinine validation of self-reported smoking in commercially-run community surveys, *Journal of Chronic Diseases* 40, 689-695.

Pudney, S. E. (2001) Sizing the market for illicit drugs: referee's comments. In Bramley-Harker, E. *Sizing the UK Market for Illicit Drugs*, pages 35-46. RDS Occasional Paper No. 74. London: Home Office.

Reuter, P. (1996) The Mismeasurement of Illegal Drug Markets: The Implications of Its Irrelevance in Pozo, S. *The Underground Economy* Kalamazoo, Mich: WE Upjohn Institute for Employment Research, 1996, pp. 63-80.

Reuter, P. (2000) *Transnational Crime: Drug Smuggling*. Paper prepared for conference on Transnational Crime, University of Cambridge, 7 January 2000.

Reuter, P. (2005) *Assessing Alternative Methodologies for Estimating Revenues from Illicit Drugs*. Draft, University of Maryland.

Reuter, P. and Caulkins, J. (2004) *Illegal lemons: price dispersion in cocaine and heroin markets*. Working paper: University of Maryland.

Rhodes, W., Langenbahn, S., Kling, R. and Sheiman, P. (1997) *What America's Users Spend on Illegal Drugs, 1988-95*. Washington: ONDCP.

Rhodes, W., Johnston, P., Han S., McMullen, Q. and Hozik, L. (2000) *Illicit Drugs: Price Elasticity of Demand and Supply Final Report Prepared for National Institute of Justice* by Abt Associates Inc, Cambridge, MA.

Rhodes, W., Layne, M. and Johnston, P. (1999) *The Supply of Heroin to the United States*, Abt Associates, Inc., report submitted to the Office of National Drug Control Policy, December 1999.

Saffer, H. and Chaloupka, F. (1995) *The Demand for Illicit Drugs*, NBER working paper, No. 5238, August 1995.

Stockwell, T., Donath, S., Cooper-Stanbury, M., Chikritzhs, T., Catalano, P. and Mateo, C. (2004) Under-reporting of alcohol consumption in household surveys: a comparison of quantity-frequency, graduated-frequency and recent recall, *Addiction* 99, 1024-1033.

Schwarz, G. (1978) Estimating the dimension of a model. *The annals of statistics*, 6, 461-464.

Thoumi, Francisco, E. (2003) The Numbers' Game: Let's All Guess the Size of the Illegal Drugs Industry! *Transnational Institute, Crime and Globalisation Paper* December 2003.

Wilkins, C., Pledger, M., Bhatta, K. and Casswell, S. (2004) Patterns of amphetamine use in New Zealand: findings from the 2001 National Drug Survey. *New Zealand Medical Journal* 117.

Wilkins, C., Reilly, J., Rose, E., Roy, D., Pledger, M. and Lee, A. (2004) *The Socio-Economic impact of amphetamine type stimulants in New Zealand*. Final report for the Centre for Social and Health Outcomes Research and Evaluation, Massey University.

Wilkins, C., Reilly, J., Pledger, M. and Casswell, S. (2005) Estimating the dollar value of the illicit market for cannabis in New Zealand. *Drug and Alcohol Review*, forthcoming.

Appendix 1: MIM Indicators

In this appendix further information on the MIM indicators that are already in the public domain is presented.

Socioeconomic data

Table A.1.1: Number of Income Support Claimants

Source	Office of National Statistics
URL	http://neighbourhood.statistics.gov.uk/dissemination
Table Description	Income Support Claimants: Total
Dataset Date	28 August 2004
Publication Date (NeSS website)	11 October 2005
Dataset Description	This dataset provides information about people who claim for Income Support. A person is eligible to claim for this if they are unemployed (i.e. working less than 16 hours a week) and have less money coming in than the law says they need to live on.

Table A1.2: Number of Incapacity Benefit Claimants

Source	Office of National Statistics
URL	http://neighbourhood.statistics.gov.uk/dissemination
Table Description	Incapacity Benefit Claimants: Total
Dataset Date	26 August 2002
Publication Date (NeSS website)	-
Dataset Description	The Department for Work and Pensions (DWP) has provided this dataset which was drawn from a 100 per cent scan of Incapacity Benefit (IB) claimants in August 2002. Although a small number of claimants whose details are held clerically will be excluded, the information provides a good basis for assessing the overall number of claimants at this time and gives details of those receiving various levels of IB for a range of areas.

Table A1.3: Number of Jobseeker's Allowance Claimants

Source	Office of National Statistics
URL	http://neighbourhood.statistics.gov.uk/dissemination
Table Description	Jobseekers Allowance Claimants: Total
Dataset Date	28 August 2004
Publication Date (NeSS website)	11 October 2005
Dataset Description	This dataset provides information about people who claim for Jobseeker's Allowance (JSA). JSA can be claimed if a person is out of work (or working for less than 16 hours per week) and is actively looking for, and is available for working at least 40 hours per week.

Crime data

Table A1.4: Burglary

Source	Home Office
URL	http://crimestatistics.org.uk
Table Description	Total number of offences
Dataset Date	Jan-Mar 2004
Publication Date	July 2004
Dataset Description	Total number of burglaries (police-recorded crime figures) where burglary is defined as follows. **Burglary in a dwelling** Where an offender enters a dwelling as a trespasser to steal, rape or commit grievous bodily harm. **Aggravated burglary in a dwelling** Where an offender, armed with a firearm, weapon or explosive, enters a dwelling as a trespasser to steal, rape or commit grievous bodily harm. **Burglary in a building other than in a dwelling** Where an offender enters a building other than a dwelling as a trespasser to steal, rape or commit grievous bodily harm. **Aggravated burglary in a building other than a dwelling** Where an offender, armed with a firearm, weapon or explosive, enters a building other than a dwelling to steal, rape or commit grievous bodily harm.

Table A1.5: Criminal Damage

Source	Home Office
URL	http://crimestatistics.org.uk
Table Description	Total number of offences
Dataset Date	Jan-Mar 2004
Publication Date	July 2004
Dataset Description	Total number of criminal damage offences (police-recorded crime figures) where criminal damage is defined as follows. **Criminal damage to a dwelling** Deliberately, or recklessly, destroying or damaging somebody else's home. **Racially or religiously aggravated criminal damage to a dwelling** Deliberately, or recklessly, destroying or damaging somebody else's home, where there is a racial or religious motive to the offence. **Criminal damage to a building other than a dwelling** Deliberately, or recklessly, destroying or damaging a building that belongs to someone else. **Racially or religiously aggravated criminal damage to a building other than a dwelling** Deliberately, or recklessly, destroying or damaging a building that belongs to someone else, where there is a racial or religious motive to the offence. **Criminal damage to a vehicle** Deliberately, or recklessly, destroying or damaging a vehicle that belongs to someone else. **Racially or religiously aggravated criminal damage to a vehicle** Deliberately, or recklessly, destroying or damaging a vehicle that belongs to someone else, where there is a racial or religious motive to the offence. **Other criminal damage** Deliberately, or recklessly, destroying or damaging items, excluding dwellings, buildings or vehicles that belong to someone else (not counted elsewhere). **Racially or religiously aggravated other criminal damage** Deliberately, or recklessly, destroying or damaging items, excluding dwellings, buildings or vehicles that belong to someone else (not counted elsewhere), where there is a racial or religious motive to the offence. **Threat or possession with intent to commit criminal damage** Threatening or causing someone to fear that his or her property might be damaged, or possessing anything, without lawful excuse, intended to cause damage. **Arson** Deliberately destroying or damaging property, by fire.

Table A1.6: MIM indicator details – Drug Offences

Source	Home Office
URL	http://crimestatistics.org.uk
Table Description	Total number of offences
Dataset Date	Jan-Mar 2004
Publication Date	July 2004
Dataset Description	Total number of drug offences (police recorded crime figures) where drug offences are defined as follows. **Possession of controlled drugs** Being found in possession of controlled drugs. **Trafficking in controlled drugs** Trading illegally in controlled drugs. **Other drug offences** Various violations of drugs legislation (see Home Office Counting Rules for details).

Table A1.7: MIM indicator details – Fraud and Forgery

Source	Home Office
URL	http://crimestatistics.org.uk
Table Description	Total number of offences
Dataset Date	Jan-Mar 2004
Publication Date	July 2004
Dataset Description	Total number of fraud and forgery offences (police recorded crime figures) where fraud and forgery are defined as follows. **Fraud by company directors, etc.** Deceiving members or creditors of a company about its affairs. **False accounting** Manipulating or falsifying an account, record or accounting document for personal gain. **Cheque and credit card fraud** Gaining possession of property, goods, services or money through the fraudulent use of a stolen or forged cheque or credit card. **Other fraud** Obtaining property by deception, with the intent to permanently deprive the owner of it. **Bankruptcy and insolvency offences** Misconduct in the course of winding up a bankruptcy or insolvency case. **Forgery or use of false drug prescription** Forging or attempting to use a false copy of a prescription. **Other forgery, etc.** Making a false copy of a document intending it to be accepted as genuine. Forgery of vehicle documentation and medical prescriptions are recorded separately. **Vehicle/driver document fraud** Forging, altering, using or lending a document relating to a vehicle (e.g. a registration document or tax disc) with the intention to deceive.

Table A1.8: MIM indicator details – vehicle and other theft

Source	Home Office
URL	http://crimestatistics.org.uk
Table Description	Total number of offences
Dataset Date	Jan-Mar 2004
Publication Date	July 2004
Dataset Description	Total number of vehicle and other theft offences (police recorded crime figures) where vehicle and other theft are defined as follows. **Theft or unauthorised taking of a motor vehicle** Stealing a motor vehicle, with intent to permanently deprive the owner of it. **Aggravated vehicle taking** Stealing a motor vehicle, with intent to permanently deprive the owner of it, and driving that vehicle dangerously, or consequently causing injury or damage to property, including to the vehicle. **Theft from a vehicle** Where there is intent to steal from a vehicle rather than drive it away. **Vehicle interference and tampering** Interfering with a vehicle, its contents or anything carried in it. **Theft or unauthorised taking of a pedal cycle** Stealing a bicycle, with intent to permanently deprive the owner of it. **Theft from the person** Stealing property from another person, with intent to permanently deprive him or her of it. **Theft from shops** Stealing, with intent to permanently deprive, from retail premises (i.e. shoplifting). **Theft by an employee** Stealing from one's employer, with intent to permanently deprive the employer. **Theft or unlawful taking of mail** Unlawfully taking away or opening a mailbag. **Theft from an automatic machine or meter** Stealing money from a meter in a dwelling. **Theft in a dwelling other than from an automatic machine or meter** Stealing property from a dwelling where the thief has not trespassed to gain entry. **Proceeds of crime** Concealing, disguising, converting, transferring or removing property obtained through criminal activity (i.e. money laundering). **Abstracting electricity** Dishonestly using, wasting or diverting electricity, without authority. **Other theft or unauthorised taking** Stealing not classified elsewhere. **Handling stolen goods** Dishonestly receiving or handling goods, knowing or believing them to be stolen.

Table A1.9: MIM indicator details – violence against the person

Source	Home Office
URL	http://crimestatistics.org.uk
Table Description	Total number of offences
Dataset Date	Jan-Mar 2004
Publication Date	July 2004
Dataset Description	Total number of offences (police recorded crime figures) classified as violence against the person where it is defined as follows. **Homicide** **Murder** Premeditated and unlawful killing of another person. **Manslaughter** Unintentional killing of another person. **Infanticide** Intentional killing of an infant under one year old by a mother suffering from post-natal depression or other post-natal disturbance. **Causing death by dangerous driving** Killing another person by driving a motor vehicle dangerously on a road or in a public place. **Causing death by careless driving when under the influence of drink or drugs** Killing another person by inattentive driving, the driver having recently consumed alcohol or drugs. **Causing death by aggravated vehicle taking** Killing another person by dangerous or careless driving of a stolen motor vehicle on a road or in a public place. **More serious wounding or other act endangering life** Viciously intending to cause grievous bodily harm to another person. **Other more serious violence offences** **Attempted murder** Attempting to kill another person in a premeditated and unlawful manner. **Threat or conspiracy to murder** Stating an intent to kill or solicit, encourage, endeavour, or persuade someone to do so. **Child destruction** Intentional killing of an unborn child, capable of being born alive, by its mother. **Endangering railway passengers** Placing railway passengers in danger by interfering in any way with the railway system. **Less serious wounding (including any minor injury)**

	Maliciously inflicting grievous bodily harm, with or without a weapon; also assaulting someone and causing him or her actual bodily harm. **Racially or religiously aggravated other wounding** Maliciously inflicting grievous bodily harm, with or without a weapon; also assaulting someone and causing him or her actual bodily harm, where there is a racial or religious motive to the offence. **Common assault (no injury)** Assaulting another person where the victim receives a minor injury or, as of 2002/03, no injury. **Racially or religiously aggravated common assault** Assaulting another person where the victim receives no injury and there is a racial or religious motive to the offence. **Harassment** Putting people in fear of violence; also continual, persistent attacks causing alarm or distress. **Racially or religiously aggravated harassment** Putting people in fear of violence; also continual, persistent attacks causing alarm or distress where there is a racial or religious motive to the offence. **Possession of weapons** Being found in possession of an object or instrument. **Assault on a constable** When a police officer is assaulted in the course of his or her duty. **Endangering life at sea** Sending an unseaworthy ship to sea. **Cruelty to or neglect of children** Inflicting pain or failing to care for a child, young person, servant or apprentice. **Abandoning a child aged under two years** Deserting or exposing an infant to a situation where its life is endangered or it is likely to be permanently injured. **Child abduction** Illegal taking of a child under the age of 16 by a parent or person connected. **Procuring illegal abortion** Giving drugs to, or using instruments on, an expectant mother to cause an illegal abortion. **Concealment of birth** Hiding an infant that has died before, at or after birth.

Table A1.10: MIM indicator details – all crime

Source	Home Office
URL	http://crimestatistics.org.uk
Table Description	Total number of offences
Dataset Date	Jan-Mar 2004
Publication Date	July 2004
Dataset Description	The sum of all six offences outlined above plus other offences (police-recorded crime figures) defined as follows. **Violent Disorder offences** **Riot** A disturbance made by an unruly mob of 12 people or more, where their behaviour causes bystanders to fear for their safety. **Violent disorder** A disturbance made by an unruly mob of three people or more, where their behaviour causes bystanders to fear for their safety. **Other offences against the state and public order** Various offences prejudicial to public order, such as hoaxes, and acts inciting racial or religious hatred. **Other offences** **Indecent exposure** Exposing one's body in a public place with intent to insult or offend. **Going equipped for stealing etc.** When not at home, a person has with him or her an item for use in connection with a burglary (i.e. a jemmy/short crowbar). **Blackmail** Attempting to obtain money by intimidation. **Kidnapping** Taking another person by force or fraud, without consent and legitimate excuse. **Treason** Betraying Queen or country, usually in an attempt to overthrow the Government. **Treason felony** Intending to deprive or depose the Queen from the imperial crown of the United Kingdom. **Perjury** A lawfully sworn witness in judicial proceedings deliberately giving false evidence. **Libel** Publishing defamatory matter in a permanent form, by a written or printed statement. **Betting, gaming and lotteries** Various offences connected with betting, gaming, lotteries and amusements. **Aiding suicide**

	Assisting, encouraging, or advising someone to commit or attempt suicide. **Immigration offences** Assisting illegal entry into England or Wales. **Perverting the course of justice** Intimidating or threatening a witness or other person involved in criminal proceedings. **Absconding from lawful custody** Escaping following arrest. **Dangerous driving** Driving a vehicle dangerously on a road or in a public place. **Firearms Acts offences** Selling firearms without a certificate. **Customs and Revenue offences** Making untrue declarations to HM Customs & Excise. VAT, car tax or misuse of oil offences. **Bail offences** Agreeing to indemnify sureties in criminal proceedings. **Trade descriptions etc.** Displaying false particulars of goods. Infringing copyright. **Health and Safety offences** Breaching/neglecting safety at work. **Obscene publications etc.** Displaying indecent material. Keeping offensive or outrageous material for profit. **Protection from eviction** Unlawfully depriving the resident occupier of any premises of his or her occupation of those premises. **Adulteration of food** Adding, abstracting, or using any article or constituent to food whilst making it harmful. **Knives Act 1997 offences** Illegally publishing or marketing of knives suggesting suitability for use in combat. **Public health offences** Violating environmental protection, pollution and waste legislation. **Planning laws** Failing to comply with various kinds of planning and preservation legislation. **Disclosure, obstruction, false or misleading statements** Various kinds of legislation relating to anti-terrorism and proceeds of crime; or falsifying, concealing and destroying documents. **Other notifiable offences** Recorded crime offences not classified elsewhere.

Health data

Table A1.11: MIM indicator details – number of people with a limiting long-term illness

Source	Office of National Statistics
URL	http://neighbourhood.statistics.gov.uk/dissemination
Table Description	People with a limiting long-term illness
Dataset Date	April 2001
Publication Date (NeSS website)	-
Dataset Description	All people usually resident in the area at the time of the 2001 Census who had a limiting long-term illness. A limiting long-term illness covers any long-term illness, health problem or disability that limits daily activities or work.

Population data

Table A1.12: MIM indicator details – Population Density

Source	Office of National Statistics
URL	http://www.statistics.gov.uk/
Table Description	Population density, 2001: Regional Trends 37
Dataset Date	2001
Publication Date (NeSS website)	-
Dataset Description	This table shows population density for 2001 based on 2001 Census data. Population estimates exclude HM Forces stationed outside the UK, but includes foreign forces stationed here. Students are considered resident at their term-time address.

Appendix 2: Methodology for estimating the size of the UK illicit drugs market

Estimation using the OCJS and AS

Estimation is based on a combination of OCJS and AS survey data and a set of assumptions sufficient to give identification. Since only adult arrestees are covered by the AS, adult and juvenile drug use must be distinguished. Overall aggregate drug consumption can be defined as:

$$D = N\,[\,P(J)\mu(J) + P(NA,NJ)\mu(NA,NJ) + P(A)\mu(A)] \qquad \text{(A2.1)}$$

where: N is population size; J and NJ refer to juveniles and non-juveniles; A and NA denote arrest (as an adult) in the reference year and non-arrest; $P(.)$ is generic notation for a probability or population proportion; and $\mu(.)$ is the mean drug consumption rate for the population group defined by its arguments. Note that event A implies event NJ under this definition. Thus, to take two examples: $P(J)$ is the population proportion of juveniles and $\mu(NA,NJ)$ is the mean consumption rate for adult non-arrestees.

This report also distinguishes between survey respondents and non-respondents, since behaviour may differ between the two groups. For the OCJS, let R and NR denote response and non-response respectively. Then the juvenile component of (A2.1) is:

$$P(J)\mu(J) = P(R,J)\mu(R,J) + P(NR,J)\mu(NR,J) \qquad \text{(A2.2)}$$

The first component of (A2.2) can be estimated as follows:

$$\hat{P}(R,J)\hat{\mu}(R,J) = \frac{\sum_{i=1}^{n_J} w_i j_i r_i y_i}{\sum_{i=1}^{n_J} w_i} \qquad \text{(A2.3)}$$

where: sums are over all n individuals selected for OCJS sampling; j_i = 1 if individual i is a (respondent) juvenile and 0 otherwise; r_i = 1 indicates a respondent and r_i = 0 a non-respondent; y_i = observed (or imputed) drug consumption if individual i is a respondent and any arbitrary value, such as 0, otherwise and w_i is a survey design weight defined as the inverse of the selection probability for the ith individual.

The second component of (A2.2) requires *a priori* assumption, since it relates to OCJS non-respondents whose drug consumption is not observed. It cannot be observed whether a non-respondent would turn out to be adult or juvenile. To resolve this problem, make the following assumptions:

Assumption 2.1 There exists a set of (discrete) variables X, observable for both respondents and non-respondents in the OCJS. The support set of X is $S_x = \{x : \Pr(X=x)>0\}$. Conditional on X and on being a non-arrestee, respondent and non-respondent juveniles have the same mean level of drug consumption:

$$\mu(NR,J,x) = \mu(R,J,x) \quad \text{for all } x \in S_x$$

Assumption 2.2 Response is independent of adult/juvenile status, conditional on X:

$$P(NR \mid J, x) = P(NR \mid NJ, x) \quad \text{for all } x \in S_x$$

Assumptions 2.1-2.2 imply:

$$P(NR,J)\mu(NR,J) = E_x\{P(NR \mid X)P(J \mid X)\mu(R,J,X)\} \qquad \text{(A2.4)}$$

which suggests the following estimate:

$$\hat{P}(NR,J)\hat{\mu}(NR,J) = \frac{\sum_{i=1}^{n} w_i \left(1-\hat{P}(R \mid x_i)\right)\hat{P}(J \mid x_i)\hat{\mu}(R,J,x_i)}{\sum_{i=1}^{n} w_i} \qquad \text{(A2.5)}$$

where the sum is over all sampled OCJS subjects, whether respondent or not. $\hat{P}(R \mid x_i)$ is the fitted probability, estimated from a probit model of OCJS response; $\hat{P}(J \mid x_i)$ is the fitted probability, estimated from a probit model of the distribution of the juvenile population across cells defined by *X*; and $\hat{\mu}(R,J,x_i)$ is the predicted level of mean consumption for individual *i*, constructed using a consumption model estimated on the sample of OCJS juvenile.

For adult non-arrestees the contribution to (A2.1) is:

$$P(NA,NJ)\mu(NA,NJ) = P(R,NA,NJ)\mu(R,NA,NJ) + P(NR,NA,NJ)\mu(NR,NA,NJ) \qquad \text{(A2.6)}$$

The first component of (A3.6) is estimated as follows:

$$\hat{P}(R,NA,NJ)\hat{\mu}(R,NA,NJ) = \frac{\sum_{i=1}^{n} w_i r_i (1-a_i)(1-j_i) y_i}{\sum_{i=1}^{n} w_i} \qquad \text{(A2.7)}$$

where a_i = 1 for (respondent) adult non-arrestee respondents and 0 otherwise.

The second component of (A2.6) again requires *a priori* assumption, since it relates to OCJS non-respondents whose drug consumption is not observed. Make the following assumptions:

Assumption 2.3 Conditional on *X* and on being an adult non-arrestee, respondents and non-respondents have the same mean level of drug consumption:

$\mu(NR,NA,NJ,x) = \mu(R,NA,NJ,x)$ for all $x \in S_x$

Assumption 2.4 The probability of adult non-arrestee status, conditional on *X* is independent of OCJS response:

$P(NA, NJ \mid NR, x) = P(NA, NJ \mid NR, x)$ for all $x \in S_x$

These two assumptions imply $P(NR,NA,NJ)\mu(NR,NA,NJ) = E_x\left[P(NA,NJ \mid R,X)P(NR \mid X)\mu(NA,R,NJ,x)\right]$, which can be estimated from OCJS data as:

$$\hat{P}(NR,NA,NJ)\hat{\mu}(NR,NA,NJ) = \left(\frac{\sum_{i=1}^{n_A} w_i \hat{P}(NA,NJ \mid R,x_i)[1-\hat{P}(R \mid x_i)]\hat{\mu}(R,NA,NJ,x_i)}{\sum_{i=1}^{n_A} w_i}\right) \qquad \text{(A2.10)}$$

where $\hat{P}(NA,NJ \mid R,x_i)$ is the fitted probability of being an adult non-arrestee, estimated from the sample of OCJS respondents but then applied to the full set of sampled individuals, whether respondent or not; $\hat{\mu}(R,NA,NJ,x_i)$ is the imputed level of mean consumption for individual *i*, constructed using a consumption model estimated on the sample of OCJS respondent adult non-arrestees.

The final component of (A2.1) relates to (adult) arrestees. It is to be estimated from AS data, which allow a detailed treatment of response. Four categories of response are used, indicated by the symbols V_1 ... V_4, which represent AS response, refusal, non-availability for non-behavioural reasons and non-availability for behavioural reasons. Thus:

$$P(A)\mu(A) = P(A)\sum_{j=1}^{4}\left(P(R,V_j \mid A)\mu(A,R,V_j) + P(NR,V_j \mid A)\mu(A,NR,V_j)\right) \qquad \text{(A2.10)}$$

where *R* and *NR* again denote OCJS response and non-response. Thus, for example, $P(R, V_2 \mid A)$ is the proportion of the arrestee population who would refuse, if selected for the AS, but be a respondent, if selected for the OCJS. The following further assumptions are made:

Assumption 2.5 There exists a set of variables *Z* which is observable for both respondents and non-respondents in the AS. The support of *Z* is $S_z = \{z: \Pr(Z = z)>0\}$. Conditional on *Z*, on being an adult arrestee and on response outcome, mean drug consumption is unaffected by the response outcome that would have been achieved by the OCJS:

$\mu(A,NR, V_j, z) = \mu(A,R, V_j, z)$ for all $z \in S_z$

Assumption 2.6 Conditional on *Z*, the mean drug consumption of adult arrestees in non-response category $j>1$ is the same as that of AS respondents:

$\mu(A, V_j, z) = \mu(A, V_1, z)$ for $j > 1$ and all $z \in S_z$

Assumption 2.7 Conditional on arrest and on AS response, *Z* is distributed independently of OCJS response. Thus:

$P(z \mid A,R,V_j) = P(z \mid A,NR,V_j) = P(z \mid A,V_j)$.

Expression (A2.3) can now be written:

$$\begin{aligned}
P(A)\mu(A) &= \sum_{j=1}^{4} P(A,R,V_j)\mu(A,R,V_j) + \sum_{j=1}^{4} P(A,NR,V_j)\mu(A,NR,V_j) \\
&= \sum_{z\in S_z}\sum_{j=1}^{4}\Big\{P(z \mid A,R,V_j)P(A,R,V_j)\mu(A,R,V_j,z) \\
&\qquad + P(z \mid A,NR,V_j)P(A,NR,V_j)\mu(A,NR,V_j,z)\Big\} \\
&= \sum_{z\in S_z}\sum_{j=1}^{4} P(z \mid A,V_j)\Big\{P(A,R,V_j) + P(A,NR,V_j)\Big\}\mu(A,V_1,z) \\
&= P(A)\sum_{j=1}^{4} P(V_j \mid A)\sum_{z\in S_z} P(z \mid A,V_j)\mu(A,V_1,z) \\
&= P(A)\sum_{j=1}^{4} P(V_j \mid A)\, E_{Z|A,V_j}\,\mu(A,V_1,z) \qquad \text{(A2.11)}
\end{aligned}$$

This can be estimated as:

$$\hat{P}(A)\sum_{j=1}^{4}\left(\frac{\sum_{i=1}^{n} w_i v_{ij}}{\sum_{i=1}^{n} w_i}\right)\left(\frac{\sum_{i=1}^{n} w_i v_{ij}\hat{\mu}(A,V_1,z_i)}{\sum_{i=1}^{n} w_i v_{ij}}\right) = \hat{P}(A)\sum_{j=1}^{4}\left(\frac{\sum_{i=1}^{n} w_i v_{ij}\hat{\mu}(A,V_1,z_i)}{\sum_{i=1}^{n} w_i}\right) \qquad \text{(A2.12)}$$

where: the sums are over all adult arrestees recorded by AS interviewers; $v_{ij} = 1$ if the *i*th arrestee sampled by the AS has response outcome *j* and $v_{ij} = 0$ otherwise; w_i is the AS design weight, equal to the inverse sample selection probability for arrestee *i*; and $\hat{P}(A)$ is the arrest probability, estimated in the manner described in Chapter 4.

Expression (A2.12) can be simplified still further, since the term v_{ij} is the only one that depends on *j* and since $\sum_j v_{ij} = 1$:

$$\hat{P}(A)\sum_{j=1}^{4}\left(\frac{\sum_{i=1}^{n} w_i v_{ij}\hat{\mu}(A,V_1,z_i)}{\sum_{i=1}^{n} w_i}\right) = \hat{P}(A)\frac{\sum_{i=1}^{n} w_i\hat{\mu}(A,V_1,z_i)}{\sum_{i=1}^{n} w_i} \qquad \text{(A2.13)}$$

However, this study works with (A2.12) to keep the AS response categories distinct and thus allow the possibility of making alternative assumptions to explore the sensitivity of the results to non-response.

Estimating mean consumption

Consumption is not observed directly. Instead, for any survey respondent *i*, the authors observe an ordinal frequency of use, f_i, taking values 1 ... *m*. For each frequency category, they have an external estimate, $\hat{c}(j)$, of the average amount used on a day of consumption. There is an external estimate, $\hat{p}$, of the average price of the drug. Consider the example of an OCJS respondent (for whom f_i is observed) with personal characteristics x_i. Then imputed mean consumption in expenditure terms is: $y_i = \hat{p}\hat{c}(f_i)$, which can be used in expressions (A2.8), (A2.10) and (A2.12) above. Now consider an OCJS non-respondent, for whom f_i is not observed. First an ordered probit model is fitted to the data on f_i for respondents. This model implies:

$$\Pr(f_i = j \mid x_i) = \Phi(\psi_j - x_i\beta) - \Phi(\psi_{j-1} - x_i\beta) \qquad \text{(A2.14)}$$

where $\Phi(.)$ is the distribution function of the N(0,1) distribution; β is a coefficient vector; and the ψ_j are a set of thresholds treated as constant parameters. Thus, in (A2.8), the imputed mean consumption levels (in expenditure terms) have the form:

$$\hat{\mu}(R,J,x_i) = \sum_{j=1}^{m}\hat{p}\hat{c}_j\left[\Phi(\hat{\psi}_j - x_i\hat{\beta}) - \Phi(\hat{\psi}_{j-1} - x_i\hat{\beta})\right] \qquad \text{(A2.15)}$$

with similar expressions for $\hat{\mu}(NA,R,NJ,x_i)$ in (A3.10) and $\hat{\mu}(A,V_1,z_i)$ in (A2.12). Consumption in quantity terms is treated identically, but with $\hat{p}$ defined as average purity rather than average price.

In addition to mean consumption, expressions (A2.8) and (A2.10) also require estimates of the conditional probabilities $\hat{P}(R \mid x_i)$ and $\hat{P}(NA \mid R, NJ, x_i)$. These are based on probit models fitted to OCJS respondents, but then used to impute probabilities for the relevant sets of non-respondents. The probit model implies, for example:

$$\hat{P}(R \mid x_i) = 1 - \Phi(x_i\hat{\gamma}) \qquad \text{(A2.16)}$$

with similar expressions for $\hat{P}(NA \mid R, NJ, x_i)$, where $\Phi(.)$ is the N(0,1) distribution function.

A2.3 Using the Schools Survey to estimate consumption by juveniles

It has been assumed that individual-level and school-level non-response in the Schools Survey (SS) are both independent of drug use. If the potential biases generated by the under-sampling of children aged 10 and 16 in the SS is also ignored, a conventional sample mean of the drug consumption variable y_i will be an adequate estimate of mean juvenile consumption. The main difficulty that remains is the lack of questions about frequency of use in the SS. Make the following assumption:

Assumption 2.8 Let f_i and d_i be the frequency of use and incidence of drug use for the *i*th individual sampled by the OCJS and r_i be the response indicator. Non-response (or under-reporting) is ignorable for the purpose of estimating the distribution of $f_i \mid d_i$ from OCJS data:

$\Pr(f_i \mid d_i, r_i{=}1) = \Pr(f_i \mid d_i, r_i{=}0)$ for all possible f_i, d_i

Thus an SS consumption variable can be constructed as in (A2.15), with the ordered probit model estimated from OCJS data individuals who report frequency of use.

Confidence intervals

The nonparametric bootstrap algorithm offers a relatively simple way of constructing indicators of the range of error associated with these estimates. Suppose a component of market size from a survey with a sample of *n* individuals is being estimated. The estimate is of the form $\hat{\mu} = g(d_1 ... d_n)$ where *g*( ) is some function and d_1 ... d_n represent the data available on each individual. The bootstrap uses an iterative resampling algorithm: at the *r*th of the *R* iterations, a random sample (with replacement) of *n* individuals from the original set of sampled individuals is drawn. Their data will be $d^*_{1r} ... d^*_{nr}$. Note that each of the d^*_{ir} will be identical to one of the original d_1 ... d_n. In general, some individuals will appear in the bootstrap sample multiple times and others will not appear at all. The estimate is recalculated for each of the *R* bootstrap samples, giving $\hat{\mu}_r = g(d^*_{1r} ... d^*_{nr})$ for each $r = 1 ... R$. The sampling distribution of the estimate $\hat{\mu}$ is then approximated by the empirical distribution of $\hat{\mu}_1 ... \hat{\mu}_R$. For large *R* (*R* = 500 is used here), this gives a good approximation.

The bootstrap can be extended by introducing random variation into parameters fixed by assumption. For example, between-individual random variations in frequency of use can be allowed for by generating a set of variables, one for each frequency category, drawn from a uniform distribution on the relevant interval from Table 4.8. Similar variation can be introduced for mean drug prices and quantities per consumption episode. The bootstrap algorithm is then executed, a new set of random consumption frequencies is generated, the bootstrap repeated, and so on. This nesting of a bootstrap within a Monte Carlo simulation works as follows:

Bootstrap algorithm

- For each of a series of R_1 Monte Carlo replications:
  - o Draw random values for mean price, quantity per episode and episodes per year by frequency class
  - o Calculate the relevant market size estimates, using the randomly generated price, quantity and frequency assumptions; call this estimate c_r ($r = 1... R_1$)
  - o Initiate the following bootstrap algorithm. For each of a series of R_2 bootstrap replications:
    - Draw (with replacement) a new sample of size *n* from the sample data
    - Recalculate the market size estimate for each replication; call this estimate c_{rs} where $r = 1... R_1$ and $s = 1 ... R_2$
  - o When the bootstrap replications are complete, calculate the bootstrap variance V_r
- When each of the Monte Carlo replications is complete, compute the final estimate as

$$\hat{C} = \frac{1}{R_1}\sum_{r=1}^{R_1} c_r$$

- Estimate the sampling variance of $\hat{C}$ as $V(\hat{C}) = \frac{1}{R_1}\sum_{r=1}^{R_1} V_r + Var(c_r)$

Appendix 3: Data sources on drug use

Table A3.1: Main survey sources of drug use data							
	Geographical coverage & timing	**Age range[1]**	**Survey mode[1]**	**Sample size**	**Sampling procedures**	**Response rate**	**Weighting**
OCJS 2003	England & Wales Jan-Jul 2003	10-65	CASI	10,085 (core) 1,886 (ethnic)	2-stage stratified sampling	74% (core)	To correct for clustering and differential selection
BCS 2002/3	England & Wales Apr 2002-Mar 2003	16-59	CASI	37,395 (core) 1,491 (youth) 2,964 (ethnic)	Multi-stage stratified (by Police Force Area) random sample	74% (core) 75% (youth) 50% (ethnic)	To correct for unequal selection probabilities
SS 2003	England Sep-Dec 2003	11-15	Paper self-completion	10,390	Multi-stage stratified random sample (schools then pupils selected)	65%	None
YLS 1998	England & Wales Oct 1998-Jan 1999	12-30	CASI	4,848	BCS 1998 + screening and focussed enumeration	69%	To correct for over-sampling big cities and non-response
AS 2003-4	England & Wales Oct 2003-Sep 2004	Over 16	CASI	7,549	Stratified random sample (60 custody suites[2])	28%	To account for survey design and non-response
NEW-ADAM 1992-2002	England & Wales various dates	Over 16	Face to face	Total 4,645 over 3 years	16 police custody suites	Under 30%	None

[1]Applicable to questions on drugs. [2] With a 61st replacement suite.

Table A3.2: Data available for OCJS non-respondents

Variable	Mean for non-respondents	Mean for respondents	Test for equality of means (P-value)
Area type: ACORN – wealthy subgroups	19.17%	19.97%	0.7146
Area type: ACORN – council areas	23.31%	19.78%	0.1448
Area type: LA with high cannabis use[1]	10.67%	09.12%	0.4108
Area type: LA with low cannabis use[1]	15.63%	15.81%	0.9531
Area type: rural LA	07.61%	10.26%	0.0661
Whole house – detached	18.97%	23.40%	0.0249
Maisonette or flat – purpose built	12.56%	08.50%	0.0004
Poor/worse house conditions	08.67%	07.80%	0.0415
Common litter/rubbish in the area	18.19%	15.21%	0.0207
Entry phone access visible	07.75%	04.96%	0.0041
No. members of the household	1.85	3.00	0.0000

Note: All estimates are corrected for survey design. [1] A preliminary cannabis prevalence regression on a set of LA dummies was used to divide LAs into three groups of high, medium and low cannabis use areas.

Table A3.3: Data available for AS non-respondents

Variable	**Mean for non-respondents**	**Mean for respondents**	**Test for equality of means** (P-value)
Gender: Female	14.09%	14.61%	0.3969
Ethnicity: White	83.90%	85.13%	0.2584
Ethnicity: Black	08.17%	08.58%	0.5389
Ethnicity: Asian	04.52%	04.07%	0.4174
Ethnicity: Mixed	03.40%	02.22%	0.0015
Interview time: 1:30-4:30	01.72%	00.94%	0.0075
Interview time: 4:30-7:30	24.32%	23.64%	0.6438
Interview time: 7:30-10:30	08.55%	11.24%	0.0052
Interview time: 10:30-13:30	14.40%	17.37%	0.0110
Interview time: 13:30-16:30	05.00%	06.07%	0.0173
Interview time: 16:30-19:30	28.62%	29.49%	0.3466
Interview time: 19:30-22:30	11.38%	09.31%	0.0034
Interview time: 22:30-1:30	06.02%	01.94%	0.0000
Interview day: Monday	12.01%	14.01%	0.0085
Interview day: Tuesday	14.93%	16.59%	0.1163
Interview day: Wednesday	14.58%	15.99%	0.0902
Interview day: Thursday	14.28%	14.37%	0.9140
Interview day: Friday	14.96%	12.63%	0.0032
Interview day: Saturday	13.30%	09.96%	0.0006
Interview day: Sunday	15.93%	16.44%	0.5897
Cause of arrest: Assault/other violent offence	24.32%	22.97%	0.1542
Cause of arrest: Burglary or other theft	21.31%	28.22%	0.0000
Cause of arrest: Shoplifting	11.72%	15.15%	0.0000
Cause of arrest: Criminal damage or disorder	16.90%	13.16%	0.0000
Cause of arrest: Drink driving, drunk/disorderly and other	11.88%	05.71%	0.0000
Cause of arrest: Drugs possession/other drugs offence	06.04%	07.21%	0.0083
Cause of arrest: Sex offence	02.22%	01.52%	0.0129
Cause of arrest: Driving while disqualified	02.30%	02.96%	0.0166
Cause of arrest: Other offence	03.32%	03.08%	0.5042
Age: 17-20	21.90%	26.27%	0.0000
Age: 21-29	34.23%	36.50%	0.0023
Age: 30-39	26.41%	22.95%	0.0000
Age: 40-85	17.46%	14.28%	0.0001

Note: all estimates are corrected for survey design.

Appendix 4: The impact of survey design on drug self-reports

Table A4.1: Drug use among 16- to 59-year-olds in the BCS 2003 and OCJS 2003

Drug	Unweighted		Weighted		H_0: equal proportions (p-value)[1]	Percentage of missing/invalid observations	
	BCS	OCJS	BCS	OCJS		BCS	OCJS
Ever taken:							
Cannabis	29.0	36.5	30.6	34.4	.000	0.5	0.6
Ecstasy	5.9	8.9	6.7	8.1	.000	0.2	0.4
Amphetamines	12.0	12.4	12.4	11.9	.255	0.4	0.4
Crack	0.9	1.4	0.9	1.1	.123	0.2	0.3
Cocaine	5.4	8.1	6.1	7.9	.000	0.3	0.3
Heroin	0.8	1.0	0.8	1.0	.103	0.2	0.3
Taken in last year:							
Cannabis	9.3	16.0	11.0	13.4	.000	1.1	0.7
Ecstasy	1.5	3.6	2.0	3.0	.000	0.7	0.5
Amphetamines	1.5	2.4	1.6	1.7	.555	2.1	0.5
Crack	0.2	0.4	0.2	0.3	.114	1.6	0.3
Cocaine	1.6	3.9	2.1	3.5	.000	0.6	0.4
Heroin	0.1	0.2	0.2	0.2	--	0.5	0.3
N	23,602	7,305					

[1] Of calibration-weighted sample; self-reported semeron users not excluded. [2] The OCJS has questions on drug use within last four weeks and last year (excluding last four weeks); a positive answer to either question is used to identify a user within the last year.

Table A4.2: Drug use among 11- to 15- year-olds in the Schools Survey 2003 and OCJS 2003

Drug	Unweighted[1]		H_0: equal means (p-value)	Proportion of missing observations	
	SS	OCJS		SS	OCJS
Ever taken:					
Cannabis	15.3	7.9	.0000	0.9	0.9
Ecstasy	1.7	0.5	.0000	1.5	1.0
Amphetamines	1.4	0.6	.0009	1.1	0.8
Crack	1.6	0.3	.0000	1.3	0.7
Cocaine	1.6	0.4	.0000	1.5	0.6
Heroin	1.1	0.1	.0000	1.5	0.6
Taken in last 12 months:					
Cannabis	12.5	6.2	.000	1.0	1.0
Ecstasy	1.2	0.4	.004	1.5	1.0
Amphetamines	1.1	0.2	.001	1.1	0.8
Crack	1.2	0.2	.000	1.3	0.7
Cocaine	1.2	0.2	.000	1.5	0.6
Heroin	0.8	.01	.000	1.6	0.6
N	10,028	1,646			

[1] The Schools Survey does not have weights

Table A4.3: Drug use among NEW-ADAM and AS respondents

Drug	Unweighted[1]		H_0: equal means (p-value)	Proportion of missing observations	
	AS	NEW-ADAM		AS	NEW-ADAM
Ever taken:					
Cannabis	75.3	82.5	0.000	0.78	0.0
Ecstasy	51.3	49.3	0.039	0.78	0.0
Amphetamines	45.2	58.3	0.000	1.03	0.0
Crack	36.3	36.9	0.487	0.73	0.0
Cocaine	44.7	42.8	0.045	0.81	0.0
Heroin	31.2	37.0	0.000	0.76	0.0
Taken in last 12 months:					
Cannabis	58.6	68.2	0.000	0.95	0.0
Ecstasy	22.7	26.9	0.000	0.91	0.0
Amphetamines	15.8	23.2	0.000	1.07	0.0
Crack	25.5	29.0	0.000	1.09	0.0
Cocaine	23.7	26.6	0.000	1.22	0.0
Heroin	24.2	30.9	0.000	0.89	0.0
N	7549	4645			

[1]NEW-ADAM does not have weights

Appendix 5: Prices and quantities

Figure A5.1: Price per gram distribution for heroin (AS 2003-4; trimmed sample)

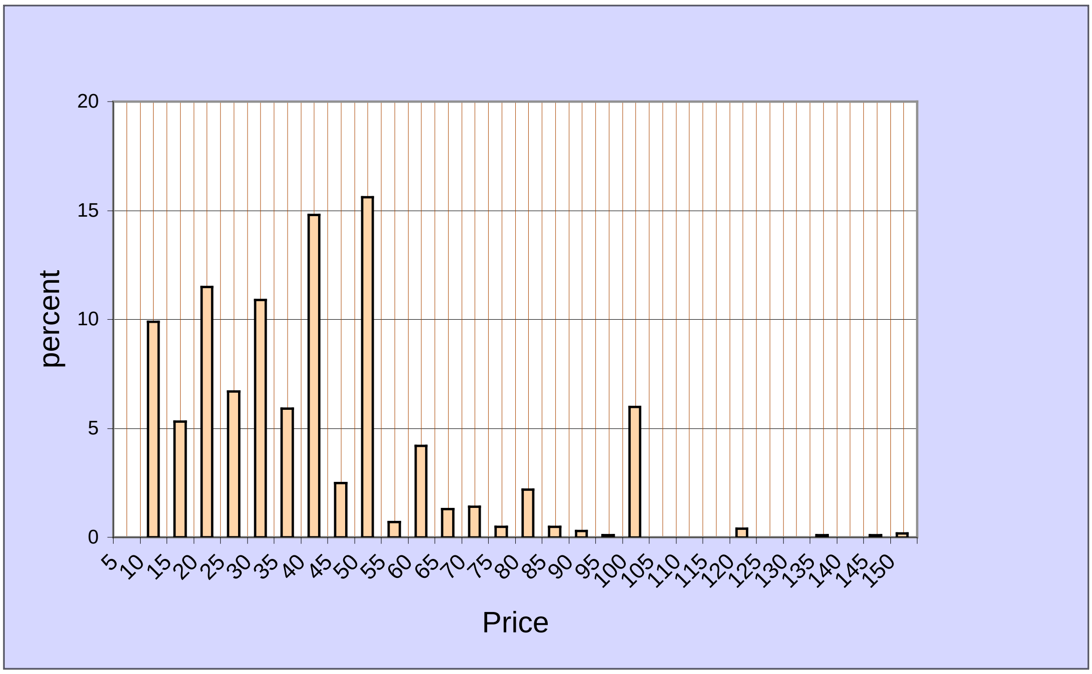


Figure A5.2: Price per gram distribution for powder cocaine (AS 2003-4; trimmed sample)

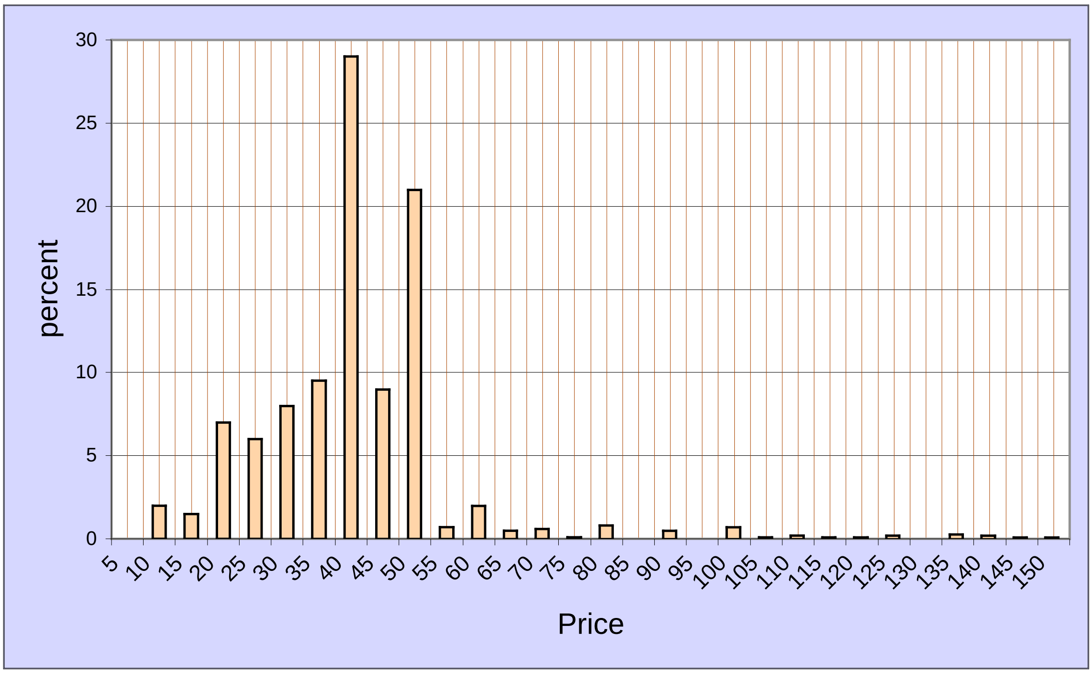
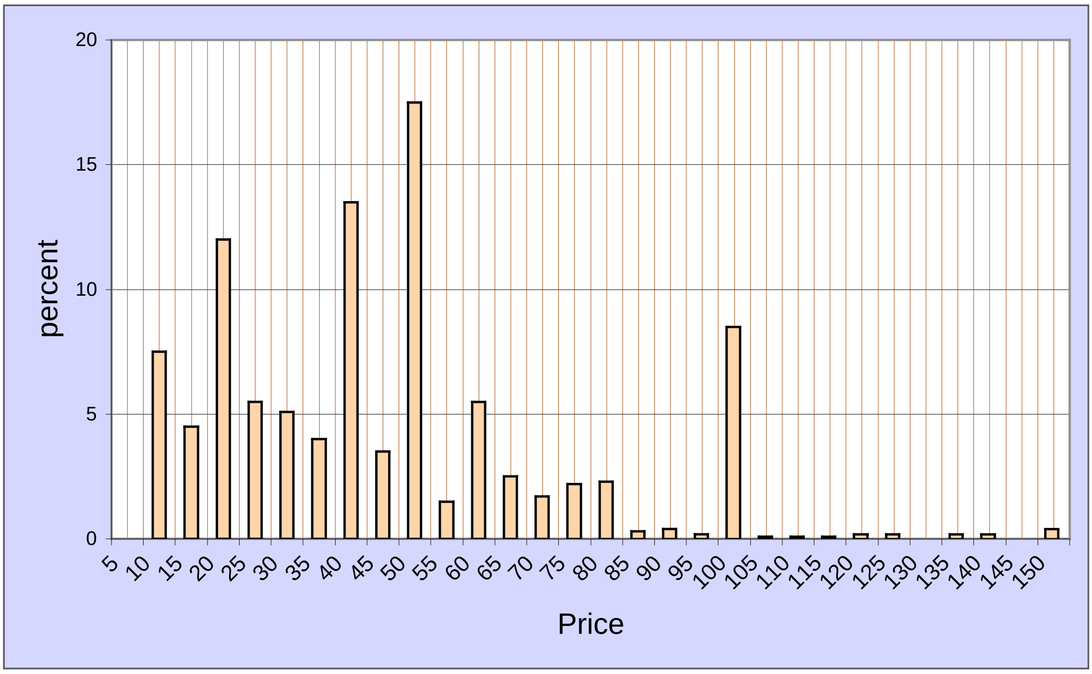


Figure A5.3: Price per gram distribution for crack (AS 2003-4; trimmed sample; assuming 0.2 gm per rock)



Nonparametric regression analysis of the relationship between unit value and transaction size suggests that there is some reduction in prices for large deals, implying that heavy users may spend less per unit than occasional users. Simple linear regressions of log price on log transaction quantity give overall elasticities of -0.21 for cocaine and -0.38 for heroin and crack. Thus, according to AS respondents, a doubling of the quantity transacted is associated with a unit price reduction of around 20 to 40 per cent. While these estimates seem plausible, they are very sensitive to the trimming policy applied to the sample and, henceforth, the authors do not attempt to adjust for variations in price by intensity of use.

Figure A5.4: Frequency distribution of heroin quantities consumed (grams) (AS 2003-4; trimmed sample

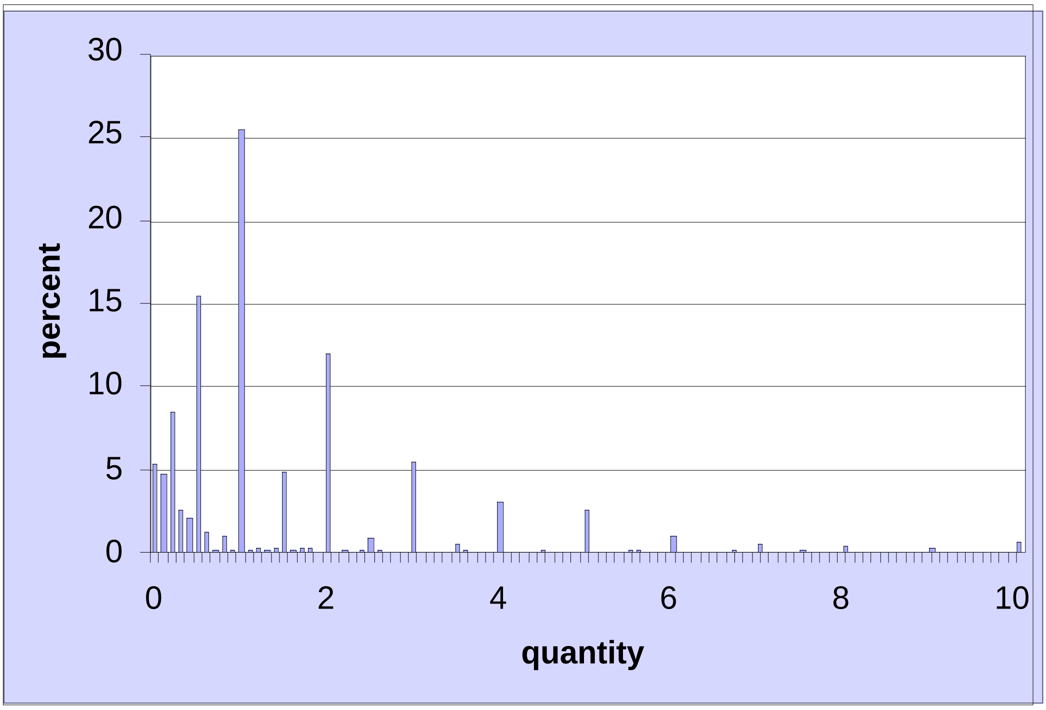
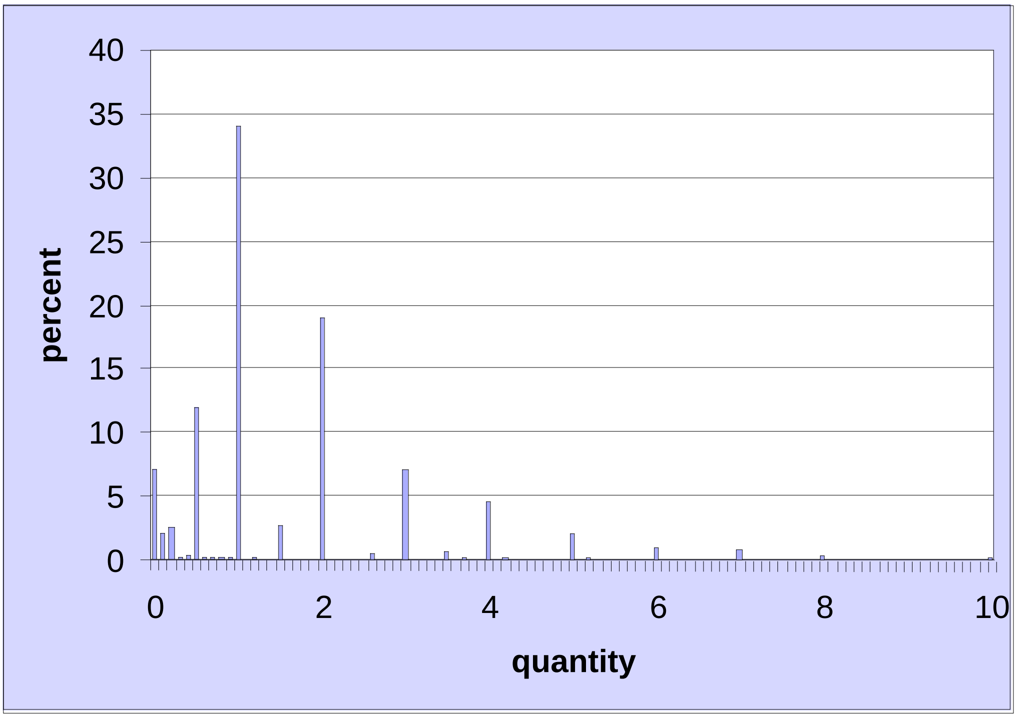


Figure A5.5: Frequency distribution of powder cocaine quantities consumed (grams) (AS 2003-4; trimmed sample)



Trimming the sample more radically reduces these levels somewhat. For example, in Table A5.1, the sample of responses in each case has been trimmed by arbitrarily dropping cases outside the range 0-2 grams, which reduces overall consumption levels to 0.82, 0.98 and 0.73 gm per episode.

Figure A5.6: Frequency distribution of crack quantities consumed (grams) (AS 2003-4; trimmed sample)

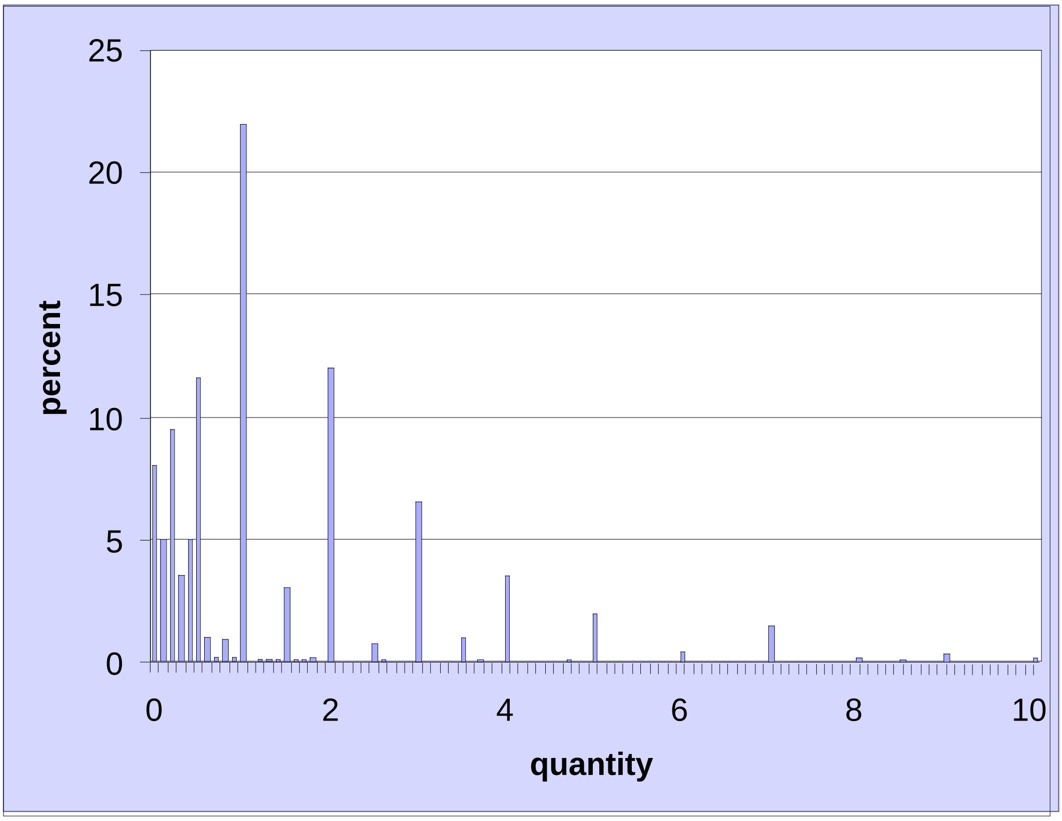


Table A5.1: Distributions of self-reported quantities consumed from AS data
(weighted for survey design and non-response; sample trimmed to give reported quantity under two grams; standard errors in parentheses)

	Heroin	**Powder cocaine**	**Crack**
Overall mean (gm)	0.82 (0.03)	0.98 (0.03)	0.73 (0.03)
n	1239	665	1016
Implied daily consumption (at average purity)	328mg (40%)	490mg (50%)	475mg (65%)
Mean daily consumption by frequency of use:			
a few times a year	0.69 (0.10)	0.88 (0.23)	0.56 (0.08)
1-2 days a month	0.78 (0.06)	1.09 (0.12)	0.60 (0.08)
1-2 days a week	0.71 (0.07)	1.28 (0.22)	0.89 (0.14)
3-4 days a week	0.72 (0.08)	0.71 (0.19)	0.78 (0.10)
5 or more days a week	1.00 (0.04)	1.06 (0.07)	0.90 (0.03)

Appendix 6: A method for estimating the size of the arrestee population

Divide the England and Wales population into four subgroups: (1) household residents aged 10 to16; (2) household residents aged over 16; (3) non-household residents aged 10 to 16; and (4) non-household residents aged over 16. Define N_j to be the size of the jth of these sub-populations and A_j to be the corresponding number of people who experience arrest during the reference period. The total number of arrestees to be estimated is $A = \sum A_j$. Let M_j be the number of arrests of members of sub-population j. Note that $M_j \geq A_j$, since arrestees may be arrested more than once during the reference period. Moreover, there is the identity $M_j = A_j \mu_j$, where μ_j is the mean number of arrests per arrestee. The imperfection of measurement and partial coverage of recording processes means that $A_1 \ldots A_4$, $M_1 \ldots M_4$, $\mu_1 \ldots \mu_4$ are not observed directly. The available measurements are given in Table 7.2 in the text.

The data in Table 7.2 raise some difficulties, since they breach the inequality: $M_{24} > N_2 P_2 \mu_2$, whose right-hand side is estimated to exceed the left by nearly 700,000 arrestees. A plausible explanation for this is that survey respondents use a much broader notion of arrest than the police statistics when reporting their arrest history. In addition to this apparent bias, there is an identification problem, since the 11 data items are insufficient to identify the 12 unknowns (A_j, P_j, μ_j, j = 1...4).

The problem of bias is resolved by assuming that $\hat{\mu}_2$, $\hat{\mu}_4$, $\hat{P}_1$ and $\hat{P}_2$ are all biased upwards by a common multiplicative factor $1/\lambda$, where $0 < \lambda < 1$. The identification problem is resolved by assuming that the mean arrest probability in the 10 to 16 non-household population is the same as that in the corresponding household population ($P_3 = P_1$). Thus the $\hat{A}_j$ can be solved as:

$$\hat{A}_j = \lambda \hat{P}_j \hat{N}_j, \qquad j = 1 \ldots 3 \qquad \text{(A6.1)}$$

$$\hat{P}_3 = \hat{P}_1 \qquad \text{(A6.2)}$$

$$\hat{Q} = \hat{A}_4 / (\hat{A}_2 + \hat{A}_4) \quad \Rightarrow \quad \hat{A}_4 = \lambda \hat{P}_2 \hat{N}_2 \hat{Q} / (1 - \hat{Q}) \qquad \text{(A6.3)}$$

The value of λ is determined using the definition of $\hat{M}_{24}$:

$$\hat{M}_{24} = \lambda \left(\hat{A}_2 \hat{\mu}_2 + \hat{A}_4 \hat{\mu}_4 \right) \qquad \text{(A6.4)}$$

This procedure gives a single point estimate. It is also possible to construct a confidence interval for the estimates $\hat{A}_j$ and their total, taking account of three sources of uncertainty:

(*i*) The variance of sampling error in the estimates $\hat{\mu}_2$, $\hat{\mu}_4$, $\hat{P}_1$ and $\hat{P}_2$, which can be estimated from the OCJS and AS survey data.

(*ii*) An *a priori* margin of error for the population totals $\hat{N}_j$ and $\hat{M}_{24}$; it is assumed that with 95 per cent probability, the true totals lie within two per cent of these estimates.

(*iii*) An *a priori* margin of error for the identifying assumption $P_3 = P_1$; it is assumed that, with 95 per cent probability, the discrepancy is within ± 0.005.

Given these estimates and assumptions, stochastic simulation is then used, making repeated draws from the distributions (*i*)-(*iii*) and solving (A6.1)-(A6.4) at each replication to produce a set of simulated values for $\hat{A}_1 \ldots \hat{A}_4$. The resulting distribution of values for the $\hat{A}_j$ are then used to construct confidence intervals. The authors do this separately for two cases. First, they allow only for sampling errors of type (*i*) and treat the population totals and identifying assumption as exact. This establishes the contribution of survey sampling error to the measurement process. Second, they repeat the process, allowing for all three sources of error, to give a better indication of the overall reliability of the estimates. The resulting confidence intervals are indicative only: in their view they are more likely to overstate than understate reliability.

Appendix 7: Conversion to a UK basis

Aggregate consumption for country c (= Scotland or Northern Ireland) can be written:

$$N_c P_c(J)\mu_c(J) + N_c P_c(NA,NJ)\,\mu_c(NA,NJ) + N_c P_c(A)\,\mu_c(A) \quad \text{(A7.1)}$$

Make the following assumptions:

Assumption 7.1 The mean drug consumption rates for juveniles, adult non-arrestees and adult arrestees is the same in country c as in England and Wales:

$$\mu_c(J) = \mu_E(J)\ ;\ \mu_c(NA,NJ) = \mu_E(NA,NJ);\quad \mu_c(A) = \mu_E(A)$$

where the subscript E denotes England and Wales

Assumption 7.2 The arrest rate for adults in country c is a known multiple, β, of that in England and Wales:

$$P_c(A|NJ) = \beta\, P_E(A|NJ)$$

After some manipulation, identity (A7.1) can be written:

$$N_c\left(P_c(J)C_1 + \frac{P_c(NJ)\left[1-\beta P_E(A\,|\,NJ)\right]}{P_E(NJ)\left[1-P_E(A\,|\,NJ)\right]}C_2 + \beta P_E(A\,|\,NJ)P_c(NJ)C_3\right) \quad \text{(A7.2)}$$

where: $C_1 = \mu_E(J)$, $C_2 = P_E(NA,NJ)\mu_E(NA,NJ)$ and $C_3 = \mu_E(A)$. There are estimates of the aggregates C_1 ... C_3 as set out in Appendix 2.
The other elements of (A7.2) are: $N_cP_c(J)$ and $N_cP_c(NJ)$ = the numbers of juveniles and adults in country c; $P_E(NJ)$ = the population proportion of adults in England and Wales; and $P_E(A|NJ)$ = the arrest rate in England and Wales.

Appendix 8: Adjustments for under-reporting

The authors' strategy proceeds in two stages: first, they attempt to infer from AS and NEW-ADAM data the extent and nature of misreporting, by comparing self-reports with drug test results. Secondly, they use the results of this analysis to predict, for each AS respondent, the true frequency-of-use probabilities, which can then be combined to construct an estimate of the average consumption level for each respondent.

Estimating the propensity to misreport

There are two sources of information on individuals' drug use. There are the results of a drug test, designed to have power to detect drug use within a standard 'detection window' (approximately 48 hours for the saliva tests used in the AS or 72 hours for the urine tests used in NEW-ADAM). The second information source is the response to a survey question on drug use within a specified reference period prior to interview. This reference period is chosen to coincide with the one-year period for which market size is to be measured, but also considered is an alternative reference period corresponding directly to the test's detection window of 48 or 72 hours.

Define C as a binary indicator equal to 1 if the respondent self-declares drug use during the reference period and 0 otherwise. $T = 1$ indicates a positive test result and $T = 0$ a negative result. The unobserved variable Q (= 1 or 0) represents underlying true drug use during the reference period.

There is some probability, $\rho^{+} = \Pr(T = 1 \mid Q = 0)$, that the test will give a false positive result when $Q = 0$. This might happen for technical reasons related to the test process or its calibration or, for a short reference period, because of physiological or behavioural factors that extend or shorten the detection window for a particular person. The test may also be sensitive to some kinds of licit consumption or prescribed medicines. Similarly, there is some probability, $\rho^{-} = \Pr(T = 0 \mid Q = 1)$, of a false negative result, when the test fails to detect drug use during the detection window. The false positive rate, ρ^{+}, is much more directly related to the technical characteristics of the test process than the false negative rate, ρ^{-}, since the latter depends on the timing, amount and purity of the drug consumption and the physiology and long-term pattern of drug use of the individual. Consequently, there is much more uncertainty about the size of ρ^{-} than of ρ^{+}.

There are three behavioural probabilities underlying the observed data. The probability of actual drug use $q = \Pr(Q = 1)$ represents consumption behaviour and the probabilities $\pi^{-} = \Pr(C = 0 \mid Q = 1)$ and $\pi^{+} = \Pr(C = 1 \mid Q = 0)$ represent intentional or unintentional misreporting by the survey respondent.

There are four possible observable outcomes in terms of C and T, whose probabilities, P^{00}, P^{10}, P^{01} and P^{11} are:

$$P^{00} = \Pr(C = 0, T = 0) = q\rho^{-}\pi^{-} + (1 - q)(1 - \rho^{+})(1 - \pi^{+}) \quad \text{(A8.1)}$$

$$P^{10} = \Pr(C = 1, T = 0) = q\rho^{-}(1 - \pi^{-}) + (1 - q)(1 - \rho^{+})\pi^{+} \quad \text{(A8.2)}$$

$$P^{01} = \Pr(C = 0, T = 1) = q(1 - \rho^{-})\pi^{-} + (1 - q)\rho^{+}(1 - \pi^{+}) \quad \text{(A8.3)}$$

$$P^{11} = \Pr(C = 1, T = 1) = q(1 - \rho^{-})(1 - \pi^{-}) + (1 - q)\rho^{+}\pi^{+} \quad \text{(A8.4)}$$

Since these exhaust the observable possibilities, they sum identically to 1 and there are therefore only three independent observable probabilities to determine the five underlying parameters, q, ρ^{-}, ρ^{+}, π^{-} and π^{+}. To solve this identification problem, further *a priori* information is needed, which is supplied by the following two assumptions.

Assumption 8.1: All misreporting is under-reporting: survey respondents do not claim drug use they have not made. Thus $\pi^{+} = 0$.

Assumption 8.2: The rate of false positives generated by the test process is a known small constant. The authors assume ρ^{+} = 2.5 per cent or five per cent.[41]

Note that, if there are individuals in the sample for whom no drug test is available, there are only two observable states, with probabilities:

$$P^{0\cdot} = \Pr(C = 0) = q\pi^{-} + (1 - q)(1 - \pi^{+}) \quad \text{(A8.5)}$$

$$P^{1\cdot} = \Pr(C = 1) = q(1 - \pi^{-}) + (1 - q)\pi^{+} \quad \text{(A8.6)}$$

The probabilities q, ρ^{-} and π^{-} can be treated as constants, allowing them to be estimated by solving the sample analogue of equations (A8.1)–(A8.4). Alternatively, the analysis can be extended to conditional modelling, by expressing some or all of q, ρ^{-} and π^{-} as functions of a set of observed explanatory covariates, X. This idea is implemented by specifying q and π^{-} as probits of the form $\Phi(X_q\beta_q)$ and $\Phi(X_\pi\beta_\pi)$, where $\Phi(.)$ is the normal distribution function and X_q and X_π are subsets of X. The authors treat ρ^{-} as a constant parameter of unknown size. The structure (A8.1)–(A8.4) can then be fitted to data on C and T by maximum likelihood. A large number of covariates X_q have been used in the drug use part of the model, including variables reflecting the arrestee's demographic characteristics, education, parenthood, labour market activity, age of onset of drug use, recent experience of prison or arrest, the timing and reason for arrest. A much smaller set of covariates has been found relevant to the misreporting part of the model, including age, parenthood, labour market activity, recent use of prescribed medicines and recent experience of arrest or prison. These sets of covariates have been reduced further in some cases, where coefficients were grossly insignificant.

A special-purpose Stata routine has been written to carry out the estimation. The top panel of Table A8.1 shows the results obtained when the probabilities q, ρ^{-} and π^{-} are treated as constants, using the 48-hour (AS) or 72-hour (NEW-ADAM) recall period for self-report and assuming a false positive rate of five per cent. Estimates are obtained for cocaine and opiates using AS data and for cannabis, amphetamines, cocaine and opiates using NEW-ADAM data. It proved impossible to find evidence of misreporting for amphetamines, presumably due to its relatively low rate of use rather than any specific feature of amphetamines that encourages honest reporting.

There is a reasonable degree of agreement between the AS and NEW-ADAM estimates: misreporting probabilities are found to be higher for cocaine than opiates for both the AS and NEW-ADAM data. The misreporting rate is also rather lower for cannabis than for cocaine, but not lower than for opiates. This is to be expected: both the AS and NEW-ADAM data reveal heroin users to have very regular, high-frequency consumption patterns and to experience frequent arrest. Since the drug is an obvious and central part of their lives, it is unsurprising that they appear more willing to admit their drug use to interviewers. More 'respectable' arrestees, who are unlikely to be heroin users, appear less willing. Even so, these misreporting rates are quite low compared to some that have appeared in the research literature, particularly for the USA. However, it should be noted that most published evidence comparing self-reports with screening test results do not make allowance for false positives. In fact, the estimates are highly sensitive to the assumed false positive rate. A lower error rate of 2.5 per cent (Table A8.1, panel 2) generates considerably higher estimated misreporting rates.

The authors have investigated the validity of the assumption that respondents have the same propensity to misreport recent drug use as they do less recent drug use, by repeating the analysis for a one-year self-report recall period. Note that this should reduce the extent of false positives. For example, if a respondent used heroin 49 hours before interview, the saliva test might give a positive result despite the 'typical' 48-hour detection window, purely for random physiological reasons. The positive result would not be 'false' in the context of a longer reference period. Consequently, we assume here a false positive rate of 2.5 per cent (note that, when ρ^{+} = 5 per cent was assumed, it was not possible to produce a positive estimate of the misreporting probability). The results are given in panel 3 of Table A8.1. For

[41] Confirmatory testing is in progress for a subset of the AS samples and this will eventually make it possible to refine this assumption.

this one-year recall period, much lower misreporting rates are found, ranging from under five per cent for cannabis to nearly ten per cent for cocaine, both in NEW-ADAM. These are strikingly low rates, which suggest that, for the arrestee population at least, the baseline market size estimates may not be too badly affected by misreporting bias.

Table A8.1: ML estimates of constant probabilities of drug use, misreporting and false negative test results (standard errors in parentheses)

	Consumption probability: *q*	Misreporting probability: π^-	False negative probability: ρ^-
48/72 hour self-report reference period; false positive test rate = 5%			
AS: cocaine	0.164 (0.014)	0.192 (0.021)	0.264 (0.018)
AS: opiates	0.196 (0.013)	0.089 (0.012)	0.117 (0.008)
NEW-ADAM: cannabis	0.549 (0.012)	0.182 (0.008)	0.177 (0.008)
NEW-ADAM: amphetamines	0.043 (0.0001)	-	0.341 (0.055)
NEW-ADAM: cocaine	0.220 (0.029)	0.222 (0.029)	0.146 (0.014)
NEW-ADAM: opiates	0.276 (0.023)	0.103 (0.009)	0.030 (0.003)
48/72 hour self-report reference period; false positive test rate = 2.5%			
AS: cocaine	0.194 (0.013)	0.315 (0.018)	0.264 (0.018)
AS: opiates	0.219 (0.013)	0.086 (0.013)	0.117 (0.008)
NEW-ADAM: cannabis	0.563 (0.012)	0.203 (0.008)	0.177 (0.008)
NEW-ADAM: amphetamines	0.045 (0.007)	-	0.340 (0.051)
NEW-ADAM: cocaine	0.244 (0.028)	0.298 (0.023)	0.146 (0.014)
NEW-ADAM: opiates	0.295 (0.023)	0.161 (0.010)	0.030 (0.003)
1-year self-report reference period; false positive test rate = 2.5%			
AS: cocaine	0.427 (0.015)	0.074 (0.010)	0.652 (0.039)
AS: opiates	0.260 (0.013)	0.087 (0.009)	0.252 (0.013)
NEW-ADAM: cannabis	0.724 (0.009)	0.046 (0.003)	0.354 (0.014)
NEW-ADAM: amphetamines	0.230 (0.025)	-	0.802 (0.069)
NEW-ADAM: cocaine	0.471 (0.025)	0.098 (0.011)	0.546 (0.054)
NEW-ADAM: opiates	0.341 (0.021)	0.059 (0.006)	0.159 (0.013)

There is likely to be considerable between-respondent variation in misreporting rates. To investigate this, the authors have re-estimated the model (A8.1) – (A8.4), allowing *q* and π^- to vary. Again, they have used the 2.5 per cent error rate assumption. Table A8.2 summarises the distribution of the conditional misreporting probability across individuals.

Table A8.2: Predicted misreporting probabilities from ML estimates of the model (A8.1) – (A8.4), with covariates

	Distribution of estimated π^-		
	Mean	Median	Std dev
48/72 hour self-report reference period; false positive test rate = 2.5%			
AS: cocaine	0.334	0.321	0.150
AS: opiates	0.179	0.149	0.096
NEW-ADAM: cannabis	0.201	0.196	0.032
NEW-ADAM: cocaine	0.348	0.356	0.110
NEW-ADAM: opiates	0.207	0.205	0.107
One-year self-report reference period; false positive test rate = 2.5%			
AS: cocaine	0.098	0.079	0.067
AS: opiates	0.087	0.045	0.108
NEW-ADAM: cannabis	0.039	0.042	0.016
NEW-ADAM: cocaine	0.093	0.064	0.050
NEW-ADAM: opiates	0.088	0.041	0.104

Adjustment for under-reporting

The econometric analysis based on the model (A8.1)–(A8.4) gives direct estimates of the four conditional outcome probabilities: $P^{00}(X)$, $P^{10}(X)$, $P^{01}(X)$ and $P^{11}(X)$ and three important unobservables: the misreporting probability $\pi^{-}(X)$; the true conditional prevalence rate $q(X)$ = Pr(Q>0 | X); and the false negative rate ρ^{-}.

Now extend Q and C to be ordinal, rather than binary, categories of frequency of use, ranging from 0 (no drug use) to J (highest frequency of use). Under the authors' assumptions, it is evident that the distribution of self-reported (positive) frequency of use is identical to the true distribution: $\Pr(Q = j \mid Q > 0, X) = \Pr(C = j \mid C > 0, X)$, for all j and X.

The nature of the adjustment process depends on what can be observed for each AS respondent. There are four cases to consider.

Case 1: positive self-reported drug use Under the assumption that respondents only ever conceal drug use, rather than exaggerate it, any positive response, $C = j > 0$, is accurate and no adjustment is required. It is only when $C = 0$ is observed that adjustment is required.

Case 2: No self-reported drug use, no test result available This case applies to cannabis, amphetamines and ecstasy in the AS, where respondents claim to be non-users, with a possibility of under-reporting. The same applies to some respondents for other drugs, where there was a refusal or other reason for not taking the saliva test. If the test result is not available because the drug in question is not tested for, the reduced model (A8.5) – (A8.6) is fitted to AS data, using external NEW-ADAM estimates of the misreporting probability $\pi^{-}(X)$. The adjustment is then based on the following relationships:

$$\Pr(Q=0 \mid C=0, X) = \frac{\Pr(Q=0 \mid X)}{\Pr(C=0 \mid X)} = \frac{1-q(X)}{P^{0\cdot}(X)} \qquad \text{(A8.7)}$$

$$\Pr(Q=j \mid C=0, X) = \Pr(Q=j \mid Q>0, X)\left[1-\Pr(Q=0 \mid C=0, X)\right]$$

$$= \Pr(C=j \mid C>0, X)\left(1-\frac{1-q(X)}{P^{0\cdot}(X)}\right) \qquad \text{(A8.8)}$$

where $P^{0\cdot}(X) = P^{00}(X) + P^{01}(X)$.

Case 3: Drug test result observed for a single drug This case applies only to heroin, for which the AS yields a test result. The following relationships hold:

$$\Pr(Q=0 \mid C=0, T=0, X) = \frac{\Pr(Q=0 \mid X)\left[1-\rho^{+}\right]}{\Pr(C=0, T=0, X)} = \frac{\left[1-q(X)\right]\left[1-\rho^{+}\right]}{P^{00}(X)} \qquad \text{(A8.9)}$$

$$\Pr(Q=j \mid C=0, T=0, X) = \Pr(Q=j \mid Q>0, X)\left[1-\Pr(Q=0 \mid C=0, T=0, X)\right]$$

$$= \Pr(C=j \mid C>0, X)\left(1-\frac{[1-q(X)][1-\rho^{+}]}{P^{00}(X)}\right) \qquad \text{(A8.10)}$$

$$\Pr(Q=0 \mid C=0, T=1, X) = \frac{\Pr(Q=0 \mid X)\rho^{+}}{\Pr(C=0, T=1, X)} = \frac{\left[1-q(X)\right]\rho^{+}}{P^{01}(X)} \qquad \text{(A8.11)}$$

$$\Pr(Q=j \mid C=0, T=1, X) = \Pr(Q=j \mid Q>0, X)\left[1-\Pr(Q=0 \mid C=0, T=1, X)\right]$$

$$= \Pr(C=j \mid C>0, X)\left(1-\frac{[1-q(X)]\rho^{+}}{P^{01}(X)}\right) \qquad \text{(A8.12)}$$

Expressions (A8.7), (A8.9) and (A8.11) can be implemented directly from the ML estimates of $P^{00}(X)$, $P^{10}(X)$, $P^{01}(X)$, $P^{11}(X)$, $\pi^{-}(X)$, $q(X)$ and ρ^{-}. Expressions (A8.8), (A8.10) and (A8.12) also required an estimate of $\Pr(C = j \mid C > 0, X)$. This is readily provided by an ordered probit model of frequency-of-use.

Case 4: Drug test result observed for group of drugs. This applies to powder cocaine and crack, for which there is a single test, which does not discriminate between powder cocaine and crack. Here, it is assumed that the misreporting probability is identical for crack and cocaine and this is estimated using a combined analysis based on (A8.1)–(A8.4). The authors then ignore the drug test result for adjustment purposes, using (A8.7)–(A8.8) directly.

Case: Non-respondents. For non-respondents there is no self-report or test result. However, the misreporting model estimated for respondents can be used to construct an estimate of the true consumption probability $q(X)$ and then use (A8.7)–(A8.8) to adjust the predicted frequency-of-use probabilities.

Produced by the Research Development and Statistics Directorate, Home Office

This document is available only in Adobe Portable Document Format (**PDF**) through the RDS website

http://www.homeoffice.gov.uk/rds

Email: public.enquiries@homeoffice.gsi.gov.uk

ISBN-13: 978 1 84726 123 6
ISBN-10: 1 84726 123 x

Erratum Online Report 16/06

Please note the following changes to this publication.

Page 43:

Text now reads: Total costs broken down by domain are presented in Figure 3.1. Drug-related crime accounts for 90 per cent (or £13.9 billion) of total costs. Drug-related deaths account for six per cent (or £923 million), health service use three per cent (or £488 million) and social care 0.4 per cent (or £69 million).

Text formerly read: Total costs broken down by domain are presented in Figure 3.1. Drug-related crime accounts for 90 per cent (or £13.9 billion) of total costs. Drug-related deaths account for six per cent (or £924 billion), health service use three per cent (or £488 million) and social care one per cent (or £69 million).

Page 45:

Text now reads: The first sweeps of estimates (for 2004/05) are now available and these have been incorporated into the revised costing model.

Text formerly read: The first sweeps of estimates (for 2003/04) are now available and these have been incorporated into the revised costing model.