

21st century Drugs and Statistical Science in UK

*Surveys, Design and Statistics Subcommittee of
Home Office Scientific Advisory Committee*

Foreword

Summary and Recommendations

1. The Landscape: Now

- 1.1 Surveys with or without biological samples
- 1.2 Databases and biological samples
- 1.3 Cohorts with or without biological samples
- 1.4 Biological sample collections
- 1.5 Tangle of technologies

2. Methodology Matters

- 2.1 Databases and ‘virtual’ cohorts
- 2.2 Using capture-recapture to estimate the number of current injectors
- 2.3 Epidemics: initiations into, and removals from, heroin use or injecting
- 2.4 Evidence-synthesis, and biases
- 2.5 Formal experiments – heed randomization and cost-effectiveness
- 2.6 Genetics.

3. Essential New Questions

- 3.1 Duration of injecting career: age at/year of starting to inject and at off-injecting
- 3.2 Number of periods ‘off-injecting for at least 1 year’ since injecting debut
- 3.3 Number of new initiates to injecting, in your presence, in the past year
- 3.4 Number of injectors, known to you, who gave up injecting in the past 2 years versus injectors who died in the past 2 years
- 3.5 Dependent use of cannabis
- 3.6 Number of HCV-contaminated injections since last HCV negative test
- 3.7 Drug users’ offending, drugs spend, and treatment/interventions – in the community and during incarceration
- 3.8 Parental use of alcohol or drugs

4. New Prospects

- 4.1 Evidence-synthesis across surveys
- 4.2 Database linkages and ‘virtual’ cohorts
- 4.3 Youth cohorts: evidence-synthesis-protocol and resource for randomized trials
- 4.4 Formal experiments in criminal justice
- 4.5 Genome-wide association scans

References

Acknowledgments

Statistical ANNEX

APPENDIX of TABLES (provisional: for completion).

Foreword

Surveys, Design and Statistics Subcommittee acknowledges not only those in the UK whose presentations, writings^{41, 66} and expertise have directly informed our thinking but a series of prescient reports by UK's Advisory Councils on the Misuse of Drugs⁴⁻⁸.

Growing, pruning, and challenging the UK's evidence-base on Drugs Science and Statistics is our brief. The methodologies for doing so are international in their reach and derivation. Our referencing does not adequately reflect international indebtedness.

The report by the Academy of Medical Sciences on "Brain science, addiction and drugs" set addiction in a robust scientific, evidential, genetic and pharmacological context¹.

The Data Sharing Review⁷⁴ in July 2008 acknowledged that developing an evidence base to improve health and social policy depends on using data derived from personally identifiable material; and made recommendations to allow this important statistical research to proceed.

In its report on "Performance Indicators: Good, Bad, and Ugly", the Royal Statistical Society recommended a wider role for formal experiments⁶⁵ and Scotland's "Road to Recovery" has espoused evidence workshops⁶⁶. Meanwhile, the National Institute for Health and Clinical Excellence has appraised the cost-effectiveness of both harm reduction and treatment interventions for dependent or injecting drug use⁵³⁻⁵⁷.

Local variation in crime was a key theme in "Crime Statistics: An Independent Review"²¹ which was chaired by Professor Adrian FM Smith FRS. We consider its implications for drugs and statistical science.

The government committed itself in its new drugs strategy to developing a cross-government research programme on drugs, with the aim of better coordinating science in this area.⁴¹ The Medical Research Council, together with National Institute for Health Research, has addictions as a strategic priority. We hope that our vision for "21st century Drugs and Statistical Science in the UK" may be of service to these endeavours.

Professor Sheila M. Bird
Chair: Surveys, Design and Statistics Subcommittee of the Home Office's Scientific Advisory Committee.

Summary and Recommendations

Our remit is growing, pruning, and challenging the UK's statistical evidence-base on Drugs Science in the 21st century. The statistical methodologies for doing so are international in their derivation and reach, and have a track record for success in monitoring and control of epidemics of chronic transmissible diseases in the late 20th century. Their deployment to tackle drug use in the 21st century could be no less remarkable. We could improve substantially and cost-efficiently the UK's quantitative understanding of, and effective interventions in, the many facets of epidemic and endemic drug use.

1. Epidemics are controlled by decreasing incidence (in this case, new initiates to injecting) and increasing recoveries (off-injecting). ***To understand changes in drug use and what interventions are effective to reduce drug misuse and its consequences, Government (the Home Office, Department of Health, and Department for Children, Schools and Families) needs to increase its efforts to measure incidence and recoveries, not just prevalence.*** Currently in the UK, there is too little of the former, and too much of the latter. Remedies are three: additional questions in surveys; new linked uses of biological samples and databases; and methodological developments.
2. Problematic or dependent drug use has many facets: criminality, infectious diseases, mental and physical morbidities, unemployment, dependant children, mortality. An intervention designed to moderate criminality may not positively (and may even negatively) affect other facets (e.g. Hepatitis C infection). ***As a rule, interventions for problematic drug users have different short and longer-term impacts so that plausible effect sizes on major outcomes should be documented a priori.***
3. ***Government funded research must make greater use of formal experiments, including randomisation, to determine “what works”, “how well”, and “whether cost-effectively” in the sentencing, treatment and rehabilitation of problematic or dependent drug users; and how best to communicate public health campaigns in relation to drug misuse.*** We note, however, that both the internal and external validity of all evaluations matter, including those that were randomized. Notably, judges lack a robust evidence-base on the effectiveness and costs of sentencing.
4. Cohort and intervention studies with problematic or dependent drug users are made less effective and more costly by many individuals being 'lost' between follow-up interviews. Without breach of client confidentiality, probabilistic database linkage can establish the dates of major events across criminal justice and health registries cost-efficiently. ***It is in the public interest that drugs science should make maximal use of approved database linkages, and Government should investigate ways that such linkages can be facilitated in a robust and timely manner.*** Such database linkage can be used to track, and analyse, the inter-

dependent educational, health, drug referral, criminal justice and benefit trajectories of ‘virtual cohorts’. ***Methodology matters in avoiding biased inference.***

5. ***To avoid deductive disclosure about individuals, the Home Office should investigate the role of ‘safe havens’ where both linkages and analyses of longitudinal data for ‘virtual cohorts’ are conducted.*** Practically, we recommend the funding of analysis secondments.
6. Historically, the UK has invested in the individualised follow-up of birth-cohorts, and some ‘at-risk’ cohorts of younger people recruited via criminal justice, mental health or drug referral settings. ***We recommend that the Home Office, together with the research councils, organises a research workshop to consider questions that might be answered by pooled analyses across these cohorts, and by added database linkages to understand the effectiveness of policy interventions.*** Comparisons between birth-cohorts from different decades and between similarly-aged but differently-recruited cohorts could be insightful, and hypothesis-generating.
7. The public costs of problematic or dependent drug use are high but diverse. Harms (and interventions) are specific to drug, route and era of use, so that frequency of occurrence, consequences, and costs all need updating. ***With the National Institute for Health and Clinical Excellence (NICE), the Government should develop a reference table, which is updated periodically, that sets out, by drug and route of use, the main cost-generating events (harms and interventions), evidence on their frequency of occurrence and consequences, and the costs they generate. In particular, revision of the Drug Harm Index is required.***
8. To profit from the decade or more of governments’ investment in surveys relating to drug misuse, ***we recommend that survey teams agree, and implement, a meta-analysis protocol to investigate collaboratively questions about antecedents to, and trends in, problematic or dependent drug use and criminality.*** Insights from cross-sectional surveys versus cohorts or database linkage should be compared.
9. Methodology matters. ***There are consequences when simplifying assumptions are made in analyses. Some redundancy of data is needed properly to test for conflicts between estimated parameters and data-sources.*** Consequences can be dramatically different answers when different assumptions are imposed on the data. In estimating the burden of drugs-related crimes, multi-parameter evidence synthesis uses influence diagrams to display the propagation of evidence, such as between health and criminal justice. It invokes expert opinion about potential biases in data-sources, and seeks to reveal conflicts of evidence. ***The Home Office should ensure that it utilises the skills to identify, and appraise critically, key assumptions underlying statistical, mathematical and economic models.***

10. ***The potential for UK's biological sample collections to be informative about patterns of drug misuse should be realised.*** Hepatitis C virus (HCV) antigen should be tested for in blood samples submitted for HCV testing from injecting drug users to identify new HCV infections. Surveillance for infectious diseases, especially HCV, in injecting drug users has been slow to measure incidence rather than prevalence – change in biological sample or questions asked is needed. Surveillance settings should adapt to study drugs-related, not just injection-related, physical and mental health co-morbidities. Mandatory saliva collections by police could be utilized for unlinked anonymous HCV surveillance, and drugs finds by police submitted for forensic analysis in randomly selected surveillance weeks. Is random mandatory drugs testing of prisoners cost-effective?

11. ***Government should consider whether biological sample collections could be used in genome-wide association scans, for example on injecting or HCV carriage.***

21st century Drugs and Statistical Science in UK

*Surveys, Design and Statistics Subcommittee of
Home Office Scientific Advisory Committee*

1 The Landscape: Now

The *gamut of surveys, databases, cohorts and biological sample collections* which do, or could, relate to drugs science ranges widely: from antenatal to post-mortem²⁶; from births through school-children³⁶ to the general adult population¹⁶; from the unemployed to occupational cohorts¹⁵; and to variously “at risk” target populations^{50,51,58,64}.

Target populations are recruited in, or linkable through, a diversity of locations: households; schools; general practices, hospital settings, virology laboratories; criminal justice settings, drug treatment agencies, needle and syringe exchanges; prescription monitoring; unemployment benefits; children’s panels; and social services.

Biological samples may be available as well as, or instead of, self-reported data and event dates (for example: third arrest date, first prison reception date, date of death). Biological samples can be specifically volunteered (for example: saliva sample from injection drug user – to be tested anonymously for HCV antibodies), mandated (urine sample from randomly selected prisoners – to be tested attributably for illegal drugs); or available by serendipity in virtue of having been collected for another legitimate purpose (Guthrie heel-prick blood-spot from newborns – collected to screen for disease in the baby but can also reveal maternal antibodies) so that, with suitable safeguards about non-attribution, residual samples can be used for public health, or other public-interest, surveillance.

Table 1.1 in **APPENDIX** summarises the commonly-available biological samples and tests that pertain to drugs science.

National databases illuminate event data, only some of which are medical-in-confidence, and may record self-reported data (risk factors for HIV infection; year of starting to inject; dependent cannabis use). National databases within the UK vary in completeness, in how data are recorded (for example: age last birthday, date of birth, age-group), in whether data-fields are back-filled if information becomes available subsequently, and in logical checking on the integrity of the recorded data.

Cohorts conventionally comprise individuals who satisfy a set of eligibility criteria (born in the same week, or diagnosed with condition X in region R) and have given their informed consent for ongoing clinical or other follow-up and for re-contact for research purposes. Informative drop-out by hard-to-reach clients (truants, prisoners, or injectors) can compromise inferences on behavioural or biological antecedents, and sequelae.

Identifiers associated with individuals’ data in *surveys, databases, cohorts and biological sample collections* range from **none** (wholly unattributable and unlinkable – no names, no labels, “no DNA”, no deductive disclosure) through **classification** with a guaranteed minimum number so classified for analysis even to proceed (six classifications by sex

and age-group, say, with samples individually unattributable and unlinkable, and a guaranteed minimum of 500 samples in a class before its analysis proceeds) to *many-one identifiers* (such as initial of first name, initial or soundex of surname, sex, date of birth: known as *master-index* when soundex of surname is used).

Many-one identifiers are linkable across datasets, but with differential success rate according to the many-one identifier chosen, the match criteria used, and the datasets to be matched to. They are liable to deductive disclosure, as in the criminal case of Stephen Kelly, one of 14 prisoners HIV infected in Glenochil Prison in 1993¹⁴.

Master-index is a sufficient basis for probabilistic linkage, and so the potential for deducing a respondent's master-index should *either* be designed out *or* the potential for its use – such as to remedy losses from individual follow-up - made explicit.

Personal numbers – such as National Insurance or Police National Computer (PNC) number - are not a suitable basis for linkage unless used identically across different datasets. Without pseudonymisation, there is a risk of deductive disclosure by third parties who have access both to the individual's identity and to their personal number. Even with pseudonymisation, longitudinal data can disclose identities.

At the extreme, DNA is ultimately attributable.

1.1 Surveys with or without biological samples

Table 1.2 in **APPENDIX** summarises major UK surveys by their target population. Further details for representatively sampled surveys can be found within the Economic & Social Research Council Question Bank (<http://qb.soc.surrey.ac.uk/docs/surveys.htm>).

Health surveys in hospital, laboratory, treatment and other settings have generally not attempted to sample settings representatively^{11,22,38,39}. Surveys under the auspices of Home Office, Ministry of Justice, or Office for National Statistics generally use representative sampling of settings and properly document response rates (by schools, prisons, police stations; as well as by individuals within settings).

Biological samples are unusual in surveys of school-children. Testing for illegal drugs, but not for injection-related infectious diseases, has been the rule for samples obtained in surveys that asked primarily about experiences of crime – as a victim or as offender.

Despite concern about dual diagnoses - mental health and addictions⁷⁶ including cannabis dependency^{4,42} – surveys have not focussed on linking the reason(s) for psychiatric admission with the presence of drugs, alcohol or infectious disease in biological samples obtained at psychiatric admission.

Despite concern about drugs-related deaths^{8,47,48,66} and suicides⁹, which are subject to forensic post-mortem, no surveillance has been designed to link no-names questionnaire-data and results in biological samples from forensic post-mortems.

Some surveys re-contact respondents on a series of subsequent (panel) dates. If the initial response was 80%, the proportion of eligible respondents who complete four interviews may be as low as two-fifths ($0.8 * 0.8 * 0.8 * 0.8 = 41\%$) or up to 70% with loyalty ($0.8 * 0.95 * 0.95 * 0.95$) but, in practice, is unlikely to exceed 60%. Cost-efficiency is compromised without database linkage as back-up.

1.2 Databases and biological samples

Databases can be a route by which to locate, and access, biological samples that were obtained for diagnostic or other clinical purposes. Database linkage can track clients' service trajectories; and be revealing about the quality of individual databases.

Knowing that a deceased Scottish patient was both HIV and HCV infected probably means that an annual series of CD4 lymphocyte counts, before and after AIDS diagnosis, has been notified to Health Protection Scotland by the relevant immunology laboratory. Liver function tests and hospitalisations were also likely in monitoring HCV progression. However, HCV may not be mentioned as an underlying cause either in relation to every hospitalisation or even when cause of death is coded. If the patient died from a heroin overdose, there is likely to have been a post-mortem at which cirrhosis of the liver would have been assessed and neuropathology or other tests done to assess brain-involvement with HIV disease. Because of his history of injection drug use, there is likely also to be a Scottish prison record for the deceased.

Registration of date and cause of death is an essential statistical function. NHS Scotland's Morbidity Register also attributes International Disease Classification codes to hospitalisation dates for individuals by name (and master-index) and is maintained by the Information Services Division, itself a 'safe haven' for conducting database linkages. Correspondingly, Hospital Episode Statistics for England and Wales are maintained by the Health Information Centre in Leeds.

National Treatment Agency (NTA) in England & Wales and Scottish Drug Misuse Database (SDMD) separately record new client episodes, including in prisons, by those seeking treatment for dependencies on illegal drugs. The two drug treatment databases are differently defined and indexed.

Separate entries for the same individual on the Police National Computer (PNC), and Scotland's equivalent, are indexed by a PNC number which is intended to be unique to the offender. Using PNC numbers, prisons in England and Wales could have a reliable dated database on serial receptions and releases for the same individual. Scottish Prison Service has had a unique prisoner-number system in 1995. See **Table 1.3** in **APPENDIX** for other prison databases.

1.3 Cohorts with or without biological samples

Under Medical Research Council (MRC) auspices, a survey in spring 2006 documented the key characteristics of 91/118 major cohort studies (typically of 1,000+ subjects) on age-related mental health. They ranged from birth-cohorts, through school-based, to recruitment in middle or older age; and from population-based to at-risk cohorts

(<http://www.mrc.ac.uk/OurResearch/ResearchPortfolios/MentalHealthResearch/CohortStudies/index.htm>). Other cohorts were identified from the Social Contexts of Pathways in Crime (SCOPIIC) website (<http://www.scopic.ac.uk/studies.htm>). **Table 1.4** in **APPENDIX** summarises younger-age and at-risk UK cohorts, together with their funding about which research ethics committees and participants may legitimately ask.

Notable is the dearth of at-risk youth cohorts^{28,50}. Youth cohorts have mainly been recruited at birth or in school, so that tens of thousands need to be studied for a few hundreds of incident events to be observed in late teenage years.

Biological samples are almost always obtained when a cohort is physician-led or epidemiological, but seldom if led by criminologists, sociologists or educationalists. Two decades ago, the MRC's firm stance on HIV disease was: no funding of behavioural questionnaires or interviews unless linked to a biological sample that was testable for HIV antibodies. Whether a similar ruling should be adopted for drugs science in the 21st century merits consideration.

Whereas most surveys report on sampling strategy and response-rates, cohort studies may overlook to do so. School-based cohorts' follow-up rates drop off very sharply once members of the cohort leave school. See SHARE on teenage pregnancies⁴⁰ and **Figure 1.1** in **APPENDIX** for an illustration of how vital database linkage can be for recovering correct inferences.

With expensive exceptions, the cost of re-contacting by phone or postal questionnaire or face-to-face interview ranges up to £200 per subject-year. Higher costs may be justifiable in at-risk cohorts, but research efficiency may also be questionable⁶².

Some £60 million pounds has been invested in at-risk or population-based cohorts, some terminated, with members who were under 40 years of age in 2000. These cohorts could yield insights on drugs issues by 2013-18, although many were designed with a different focus. Nested case-controls studies are embedded in some cohorts such that, when an event of interest occurs (a case, however defined), both the case and contemporaneous controls can be invited to provide relevant self-report data, biological samples, specifically-permitted access to medical files, or to attend for a clinical or other review.

Also insightful for drugs science is to map issues such as poly-drug use or transitions between specific drugs which are addressed both in cross-sectional surveys and in cohort studies: and to check whether cross-sectional snapshots readily translate longitudinally.

1.4 Biological sample collections

Biological sample collections may be centralised^{15, 39} or diffusely located²²; and may relate to obligatory¹⁵, volunteered³⁹, clinical⁵¹ or research-specific samples.

Biological samples are required by law in particular circumstances. DNA is held on some 4 million persons who have been either convicted or arrested for the purpose of being interviewed by the police. In England and Wales, a saliva sample - to be tested on-site for

the presence of heroin or cocaine - is required from those arrested in relation to acquisitive crime. The likelihood of testing positive depends on the time since last use of heroin/cocaine. Lacking is any centrally-maintained database (labelled by master-index or other identifier) on these mandatory test results for either audit or linkage to future criminal career.

Table 1.5 in **APPENDIX** summarises the main national databases and biological sample collections pertaining to drugs science.

1.5 Tangle of technologies

A tangle of technologies underpins quantitative drugs science. We focus first on two aspects: *questions posed*, and *unrepresentative sampling*. Other issues relate to: *oversight-by-whom*, *consent* and *linkage of acquired data*. And finally, *costs*.

Questions posed: The number of respondents likely to report a particular behaviour – injection drug use, for example – determines whether it is worth asking subsidiary questions on injecting. **Table 1.6** in **APPENDIX** reviews studies' first three questions on injecting and three about heroin use. Dependent use of cannabis has not generally been ascertained, except when inquiring into psychiatric morbidity, but does feature among the questions asked of those referred to drug services.

Sampling: Arguments against representative sampling of locations are difficult to sustain when surveillance is no longer experimental or proof-of-concept¹¹. Unrepresentative surveillance risks biased estimation and under-dispersion. See surveys of injectors' risk behaviours and infectious disease prevalence³⁸ in **Table 1.2** in **APPENDIX**.

Oversight: Frank answers require that studies are designed to protect respondents from dual risks: i) of answers being attributable to them or ii) of repercussions on those like them (fellow-prisoners or fellow-pupils) from frankness by their peer-group.

Consent: Unwilling gate-keepers (parent, headmaster, prison governor, doctor) disenfranchise potential respondents without their knowledge, but often do so for understandable reasons: from fear of breaching data protection, confidentiality, or an individual's peace-of-mind.

Linkage and deductive disclosure: Whenever initial of first name, surname, sex and date of birth are known, a master-index can be created by the data-holder, which enables probabilistic linkage - without disclosure of the client's identity – to other databases. To be approved⁷⁴, both linkage and analyses of the linked data may have to be done *within 'safe havens'* so that longitudinal data are not actually returned to a data-holder. If they were, the data-holder might inadvertently discover new information about 'their' named clients (by re-matching of longitudinal strings) which identifiable clients had not disclosed to them¹⁶.

Linkage of information that has been volunteered in surveys; compulsorily acquired; or derives from a biological sample that was given voluntarily to establish an individual's

diagnosis has a range of implications for: public health; public-interest; research; accountability by Ministers; professional codes-of-conduct; and ethics. This diversity needs to be bridged if properly-collected data are to be well analysed in the public interest, to optimise their scientific-effectiveness, to safeguard frankness, properly to challenge policy, and to assure cost-effectiveness.

Costs: Just as periods on and off injecting may have time-dependent influences on different morbidities (reconvictions; dependent use of alcohol; unemployment) or causes of death (drugs-related death; suicide; liver-related)⁵¹, time-dependency in cost-generating events^{33,34} induces time-dependencies in overall costs. For such reasons, it is timely to revise the Home Office's Drugs Harm Index.

Criminal justice implications for acquisitive crime may not be the same when the route of heroin use is by injection versus not, just as the public health implications differ – for blood-borne infections and overdose deaths, in particular. See Arrestee Survey (<http://www.esds.ac.uk/findingData/snDescription.asp?sn=5807>).

In short, the public costs of problematic or dependent drug use are high but diverse. Harms (and interventions) are specific to drug, route and era of use, so that their frequency of occurrence, consequences, and reference costs all need updating over time. A reference table, which is updated periodically, that sets out, by drug and route of use, the main cost-generating events (harms and interventions), evidence on their frequency of occurrence and both short and longer-term consequences, together with the costs they generate would greatly assist in making comparable cost-effectiveness calculations^{18,19,33,34,57,68}.

2 Methodology Matters

The public cost of data is a reason for their better exploitation in statistical analysis^{2,18,71}. Of course, how data were obtained matters – by legal compulsion, volunteered, by unexercised opt-out, or unconsented. Public-interest arguments for database linkage are: *firstly* the importance of the questions that will be answerable if linkage is permitted, *secondly* the safeguards against deductive disclosure about individuals who did not give explicit consent for linkage, and *thirdly* the linkage-protocol being open to public and professional scrutiny.

Methodologies beyond cross-tabulation of data matter for value-added analysis. Key methods are précised in this section, some others in the statistical annex.

2.1 Databases, and 'virtual' cohorts

Professional codes of conduct require scientists to consider the propriety of analysing data that were obtained by compulsion or under duress. This may partly account for limited analysis of prisoners' random mandatory drugs tests (rMDTs) for opiates by weekday when compulsory drugs tests in the British Army (a condition of employment) has been highly informative about soldiers' weekend pattern of cocaine use¹⁵. See **BOX 1**.

Obtaining approvals for database linkages can be lengthy – more so than the analysis itself. Data-holders need well-trained staff to enact the approved database linkages in reasonable time, or should be affiliated to a ‘safe haven’ whose staff can be called on.

BOX 1: Number of random mandatory drug tests performed by weekday in 3-year eras; with **P = number positive for prescribed methadone**, **O = number opiate positive** ^{rate per 1,000}, and **C = number cannabis positive** ^{rate per 1,000} (95% CIs for rates)

Financial years	Totals	Monday to Wednesday	Thursday + Friday	Saturday+ Sunday
Prisons which elected <i>for</i> 5% rMDTs (95% CIs for weekday positive rate per 1,000)				
2000/01 to 2002/03	87,300 P 12 O 4,298⁴⁹ (48, 51) C 6,906⁷⁹ (77, 81)	48,996 O 2,360 ⁴⁸ (46, 50) C 3,935 ⁸⁰ (78, 83)	26,169 O 1,404 ⁵⁴ (51, 56) C 2,079 ⁷⁹ (76, 85)	12,135 O 534 ⁴⁴ (40, 48) C 892 ⁷⁴ (69, 78)
2004/05 to 2006/07	110,204 P 419 O 4,739⁴³ (42, 44) C 7,503⁶⁸ (66, 70)	58,614 O 2,528 ⁴³ (41, 45) C 4,201 ⁷² (69, 76)	32,108 O 1,412 ⁴⁴ (42, 46) C 2,146 ⁶⁷ (64, 70)	19,482 O 799 ⁴¹ (38, 44) C 1,156 ⁵⁹ (56, 63)
Prisons which elected <i>against</i> 5% rMDTs (95% CIs for weekday positive rate per 1,000)				
2000/01 to 2002/03	70,997 P 4 O 2,449³⁴ (33, 36) C 4,670⁶⁶ (64, 68)	38,044 O 1,285 ³⁴ (32, 36) C 2,638 ⁶⁹ (67, 72)	21,301 O 735 ³⁵ (32, 37) C 1,321 ⁶² (59, 65)	11,652 O 429 ³⁷ (33, 40) C 711 ⁶¹ (57, 66)
2004/05 to 2006/07	66,113 P 332 O 2,040³¹ (30, 32) C 3,277⁵⁰ (48, 51)	35,137 O 1,079 ³¹ (29, 33) C 1,870 ⁵³ (51, 56)	18,352 O 599 ³³ (30, 35) C 887 ⁴⁸ (45, 52)	12,624 O 362 ²⁹ (26, 32) C 547 ⁴³ (40, 47)

Opiate positive rates decreased by a tenth, and very significantly, between eras. Markedly different from Scotland²⁵, even in the later era, there was minimal prescribing of methadone because the numbers testing positive for prescribed methadone were one ninth only of those positive for opiates. Opiate positive rates were largely uninfluenced by weekday. But cannabis rates were influenced, and tended to be lower **at weekends** than on Mondays to Wednesdays. Whether this pattern relates to the supply of cannabis into jails, or to recent outside-use by new receptions, is unclear.

Weekday of sample may matter when interpreting opiate and cocaine positive results obtained by the police in mandatory saliva testing of persons arrested for acquisitive crimes: results from the Arrestee Survey’s respondents who were, and were not, arrested for a trigger offence would be a good starting point.

Provided that left truncation and ascertainment bias are correctly handled^{29,30,46}, database linkage studies are powerful and have already quantified: UK prisoners’ 7.5 times higher risk of drugs-related death in the first fortnight after release from prison^{13,27} (and 2 times higher in the second fortnight); cause-specific mortality⁵¹ and morbidity rates for HCV-

diagnosed patients; delay from injecting debut to first attendance for drug treatment; how predictive usual weekly alcohol consumption, as reported in health surveys, is of (time to) subsequent first alcohol-related admission to hospital.

New insights can be got from linkage studies for ‘virtual’ cohorts. Across several databases, the idea is to build up a record of event-dates for individuals in a ‘virtual’ cohort (of persons ‘referred for drug treatment in 2005’, say). See later in **Section 4.2**.

2.2 Using capture-recapture to estimate the number of current injectors

Database linkage also underlies capture-recapture methods for estimating the number of current injectors³⁷.

A priority for policy makers has been local estimates. Local estimation can be at the expense of more sophisticated understanding (and modelling) of capture propensities⁴⁷⁻⁴⁹ at regional or national levels. Broadening the set of log-linear models (or model space) that is explored means both longer computational time and that the uncertainty interval qualifying the eventual estimate is wider because, besides parameter uncertainty, model uncertainty is reflected. For England, the estimated number of current injectors was also much higher at 204,000 (uncertainty: 189,000 to 223,000) when capture propensities were allowed for hitherto: 137,000 (uncertainty: 133,000 to 149,000), see **BOX 2**.

BOX 2: A Bayesian capture-recapture approach which modelled capture propensities^{47,48} was taken to estimating the number of current injectors in England by age-group (15-24, 25+ years) and sex for each region. For each region, the Bayesian estimates were higher than the original. Model differences led to non-overlapping uncertainty intervals for the national total. Generally, more model uncertainty was reflected in the Bayesian analysis.

Estimated number of current injectors (to nearest 100): Bayesian posterior mean and 95% credible interval at regional level compared with localised, classical estimation with 95% confidence interval.

Region	Bayesian estimate	Localised, classical estimate
East of England	11,100 (9,600; 12,900)	9,400 (6,300; 13,100)
East Midlands	15,700 (13,900; 18,000)	11,800 (10,500; 13,500)
London	45,800 (34,800; 60,600)	17,900 (16,200; 24,000)
North East	12,300 (10,400; 15,300)	9,000 (7,600; 10,600)
North West	35,400 (31,500; 39,700)	22,100 (18,800; 25,200)
South East	15,500 (12,800; 26,800)	13,800 (12,000; 17,800)
South West	19,300 (16,800; 22,000)	17,400 (15,900; 19,500)
West Midlands	17,100 (15,300; 19,400)	14,700 (13,600; 17,000)
Yorkshire and the Humber	31,800 (28,400; 35,800)	21,000 (19,900; 22,800)
ENGLAND	204,000 (189,300; 222,700)	137,100 (133,100; 149,100)

Important differences in modelling were: i) the original analyses were performed at local level and aggregated up to regional, whereas the Bayesian analysis was done on regional counts, ii) the localised original analyses considered the simplest 22 models only - with a maximum of two 2-way interactions, whereas the Bayesian analysis considered all models with 2-way interactions, and iii) the original analysis ignored capture propensities by age-group and sex, which the Bayesian analysis accommodated.

A suite of programs to facilitate different technical approaches (including Bayesian), different constraints on model space, or on consistency with external data (such as drugs-related deaths) or with expert prior opinion about capture propensities is needed.

Planning services for current problem drug users, and estimation of their uptake, requires knowledge of how many problem drug users there currently are, where, and their propensity to access services.

The number of persistent injectors (listed as current injector both now and, say, three years ago) has not been estimated; but could, and should, be.

2.3 *Epidemics: initiations into, and removals from, heroin use or injecting*

Similar techniques as for Hepatitis C^{43, 63} and used earlier to estimate the HIV incidence pattern underlying AIDS diagnoses²³ have been applied to back-calculate from opiate-related deaths to heroin incidence by using external knowledge on the incubation time from heroin debut to opiate-related death^{24,70}. Simplifying assumptions were made – such as that the rate of non-fatal removal from heroin dependency did not increase over time.

Greater emphasis on harm reduction and treatment referrals should have increased the removal rate in the past decade. By how much is important in policy terms, but the change of assumption may also impact on the estimated pattern of heroin incidence⁷². Before analysis begins, expert opinion could be elicited⁵⁹ about how much heroin users' engagement in substitution therapy is likely to have increased their removal rate from heroin dependency or from injecting. Do the data analysis and opinions concur⁷²?

Being explicit about simplifying assumptions is crucial^{31,32,60,61, 69, 70, 75}: so that experts can consider how tenable or realistic any simplification is, what impact its relaxation could have on the results, and what additional data or methodology are needed to test assumptions.

2.4 *Evidence-synthesis*

At a basic level, meta-analysis weights randomized controlled trials (or epidemiological studies) proportionately to the information they contribute, and - formally or informally – can take into account the internal and external validity of each study⁷⁵ (its quality and generalizability).

In essence, evidence-synthesis quantifies what we know already; and sign-posts how (and how not) to design the next critical study to ensure that it is powerful enough to discern the effect size that is plausible for the *next intervention in a sequence* to achieve⁷¹.

Evidence-synthesis has been generalized. *First*, inferences can be drawn about {A versus B} when the only available randomized controlled trials compare {A versus C (control)} or {B versus C}. The two series of trials may have been conducted in the same patient population by same investigative team; or in different countries by different investigators. For the exchangeability of evidence, it matters which. The integrity of trials (their internal validity) also matters, and likewise their generalizability (to the patient population for whom inference about {A versus B} is required: external validity).

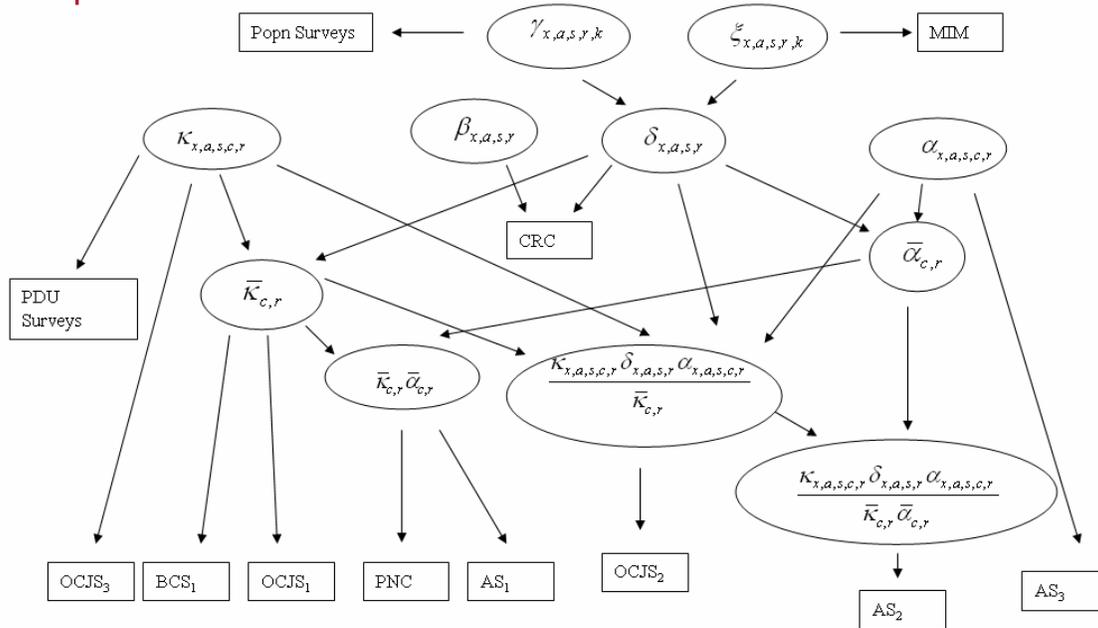
Over-emphasis on internal validity, including rigidity over randomization, can be at the expense of external validity. The judgments made in evidence synthesis need properly to balance external relevance against the risks of internal bias.

More generally still, multi-parameter evidence-synthesis combines evidence from a range of data-sources^{3,22,35,68} – to estimate the burden of drugs-related crimes for example; and uses influence diagrams, see **BOX 3**, and probability arguments to work out which data-sources are informative, directly or indirectly, about key parameters. Expert prior assessment about the internal and external validity of each data-source (biases) is sought, and simplifying assumptions (typically when to invoke exchangeability) play an even stronger role when it comes to multi-parameter evidence synthesis. Some redundancy of data-sources is needed for conflicts of evidence to be exposed.

BOX 3: Ades et al. have outlined a multi-parameter evidence synthesis which recognises the broad social, health and criminal justice contexts of, and evidence-sources on, problematic use in the UK of specific illegal drugs. The influence diagram shows that overlapping data-sources inform the key parameters of interest, and this is essential to be able to make progress analytically. Moreover, subject-matter specialists can give informed prior opinion about the potential biases between and within these data-sources.

	Offender Crime Justice Survey	British Crime Survey	PNC	AS PDU	surveys	Capture Recapture MIMs House-hold surveys
Crime rate by user type	x	x				
Prob (crime reported)	x	x				
Rate reported crime	x	x	x			
Rate arrested crime	x		x	x		
Prob (crime committed by user of given type, and arrested)	x					
Prob (arrested crime committed by user of given type)				x		
Prob (arrest crime and user type)	x			X	X	
Prevalence of drug use, by type						x

🌟 Schematic influence diagram: relation between parameters and data sources.



Much data already exist to populate the influence diagram but, behind it, quite taxing probability calculations are needed, and may entail simplifying assumptions. Formal elicitation techniques are needed because of the range of expertise - drugs-related, criminal justice, social, health, policing, statistical – that is knowledgeable on parts of the information or on the modelling of it.

Costs can be overlaid on a multi-parameter evidence synthesis, or its parameter estimations routed into cost-effectiveness studies.

Bias, when generalizing from cross-sectional (so-called snapshot) samples⁴⁶ or from cohorts^{29,30}, can be substantial. Simulation studies are useful²⁹ – seeing ascertainment bias in action is believing how large the biases can be! Methodology matters . . .

2.5 Formal experiments – heed randomization and cost-effectiveness

Elicitation of prior opinion^{25,59} and meta-analysis of related trials⁷¹ are two techniques for determining the a-priori-plausible effect size that a formal experiment – with randomization of clients – needs to be powerful in respect of.

Formal experiments on the effectiveness of criminal justice interventions for drug-dependent offenders are as essential as randomized controlled trials to estimate the efficacy and safety of pharmaceutical medicines^{12,67}.

Interventions’ extra cost relative to their likely extra effectiveness (at reducing re-convictions, say) should be instrumental in deciding which are worthy of evaluation, and which should be referred back to the drawing board – to prune costs or enhance their likely effectiveness before embarking on formal evaluation.

All well-designed experiments, whether randomized or not, ought to have a study protocol⁶⁵ that documents the prior considerations which determined the choice of intervention, its likely effect size and costs, how the study should be designed, baseline covariates, primary and secondary outcome variables (to be measured when, how, and by whom), and a statistical analysis plan. Data collection forms, or database specification, should be appended.

It is important to ensure that "the best is not the enemy of the good". Evaluative research can inform questions of effectiveness if there is genuinely a high signal to noise ratio; and qualitative research can be instructive about what can, or can't, work practically so that its findings inform how an intervention is to be designed which is subject to subsequent formal evaluation.

2.6 Genetics

Methodologies which minimise false-discovery rates in genome-wide association studies⁷³ have been deployed successfully with series of 2,000 to 10,000 cases and corresponding controls, see **BOX 4**. For these methods to be applicable in drugs science requires biological sample collections appropriate to the heritable susceptibility at issue: heroin-dependency, injection drug use, heroin-overdose-fatality, cannabis-dependency, cannabis-related psychosis, cocaine-dependency, or alcohol-dependency. And a well-reasoned study protocol.

BOX 4: Biological sample collections have been under-utilised for genome-wide association scanning. A scan which uses 2,000 -10,000 cases and similar number of controls (possibly shared with a previous study) would have a good chance to detect multiple common genetic variants - with small individual effects but reasonable cumulative effect - on the heritability of addiction-related outcomes. Major hurdles may be ethical concerns and cost. Such case series studies need around two million pounds to conduct and access to a suitable control series (see below). Costs relate to manual DNA extraction for case series, contributory costs of £3 per control series sample, genome scanning, and statistical analysis.

Biological collections which have stored or, in future, could store residual blood samples from ever-injectors and, as controls, non-injectors who have undergone HIV or HCV antibody testing could readily accumulate over 2,000 samples. Screening out repeat samples from the same individual is not difficult on a DNA-basis.

Sufficient DNA (preferably 4 to 5 micrograms; down to 200 nanograms can be sufficient for some techniques) for subsequent use in genome scanning by Affy-500K (500,000 snips) or its next generation of 1million snips can be extracted from 3-5ml of fresh blood (even after delays of up to 2 – 4 days), from Guthrie blood spots, and from saliva collected using ORAGENE kits (which cost £10-15). Genome scan results and subsequently-sourced behavioural, drug treatment, or forensic data can be linked at the time of statistical analysis, but in a manner that does not permit the resulting scan-analysis to be individually attributable.

With NHS Blood and Transplant in Cambridge, a joint Cambridge University and Wellcome Institute research team has established: *first, a control cohort of some 3,000 donors from England and Wales, classified by year of birth, sex and first 3-digits of postcode, who permitted DNA extraction from the lymphocyte-filter residue of their blood donation; secondly, a Cambridge Bio-Resource Cohort of 10,000 local donors who not only gave permission for DNA extraction but also to be re-contacted for secondary studies which may include MRI scans.* Research Management Committees sanction access to the protocolled use of both these control cohorts.

3 Essential New Questions

By broad consensus, new questions are needed about: injector incidence, cessations, number of novices initiated per injector, and initiation contexts; transitions to hazardous/dependent use, for example, of cannabis or cocaine; drug interventions during incarceration; spending on illegal drugs in past week or 4 weeks – in and out of prison; drug users' offending – by main drug or route of use; parental use of alcohol and drugs and perceptions of their child's use of drugs versus the child's self-reported (or detected) use.

Reservations were raised about: type of cannabis used (consumers are uncertain, and so better addressed by random weeks with full toxicology on all police finds); misuse of prescription drugs (whom to survey - prescribers or consumers); negative consequences perceived by drug users (balance by questions on positive consequences; also, perceptions are affected by time since last use); questions on drugs supply (frankness of answers; and whom to survey – street dealers, incarcerated dealers, or users few of whom may be 'in the loop').

3.1 *Duration of injecting career: age at/year of starting to inject and at off-injecting*

Duration of injecting careers has to be estimated, but how⁷⁰?

Data-sources include: i) current injectors attending needle and syringe exchanges, ii) new clients at treatment agencies who report having injected in the past 4 weeks, iii) ever-injector inmates in prisoner surveys, iv) former injectors attending drug treatment agencies who have ceased injecting and are on the road to recovery, or v) former injectors in the British or Scottish Crime Surveys. Each is a biased snapshot but, suitably analysed^{46,30}, can yield an estimate for the duration of injecting careers which is generalizable to all injectors. Correction for ascertainment bias will be only approximate, so that apparently different estimates across the five data-sources illuminate residual biases, and how better to resolve them. Multiple data-sources are thus essential.

For the duration of injecting careers to be estimable, surveys and databases on injecting drug users (including registers of HCV diagnoses) should routinely ask questions on:

- a) age at/year of starting to inject {knowing year of birth allows 'age at' to be translated into 'year of' }
- b) age at/year of off-injecting {good intentions re 'off-injecting' may lapse and so there needs to be commonly-agreed definition such as 'at least 1 year since last injection'. Alternatively, respondents could be asked for year of last/most recent injection and whether their injecting drug use has ceased. }

For some analyses, knowledge of the month and year of starting/ceasing to inject is preferable – if recall of month would be reliable. However, overly precise questions do not necessarily yield the best-quality data. In practice, it may be better to trust to the respondent to work out whether s/he has been 'off-injecting' for at least 1 year.

3.2 *Number of periods “off-injecting for at least 1 year” since injecting debut*
Injection drug use is recognised to be a remitting-relapsing condition. Remissions may coincide with periods of incarceration. Thus, it is important to ascertain how many spells of being “off-injecting for at least 1 year” the respondent has achieved – not least because their number may be informative about the potential for a subsequent relapse from the current ‘off-injecting’ period.

3.3 *Number of new initiates to injecting, in your presence, in the past year*
Statisticians approach the epidemic of injection drug use as they do an epidemic of any transmissible infectious disease. What matters for bringing the epidemic under control is that, on average, a “newly-infected person” (that is: a new injector) is responsible for “transmitting disease” (injection drug use) to at most one other person. Injecting careers can be long, which makes epidemic control more difficult. Public health interventions for injectors aim to shorten the ‘infectious period’ (their injecting career) and to avert or limit new initiations – both to injecting and to HCV infection.

Rather little is known quantitatively about the context of initiations into injection drug use. For example, if three experienced injectors are present when a novice is initiated into injecting, then, to a first approximation, each experienced injector may be taken to be one-third responsible. The part that each played – buying the heroin, drawing it up into a new or shared syringe, demonstrating to the novice how to inject or administering the first injection – is ethnographically interesting, but less relevant statistically.

To limit errors of recall, surveys should ask about the respondent’s own initiation, and about the initiation of others in the past year. For some respondents, ‘the past year’ will be their own initiation year, but may be 3, 5 or 10+ years into their own injecting career.

3.4 *Number of injectors, known to you, who gave up injecting in the past 2 years versus injectors who died in the past 2 years*

Survey questions give pause for reflection. The balance of live cessations from injecting versus injector deaths in the past 2 years is important. The same cessations, and the same deaths, will be multiply reported but their ratio – regionally and nationally – remains informative and its increase over time is a measure of public health success.

Since injectors’ death-rate from all causes is around 1.5% per annum, there is only a 40% chance that no-one in an injector-network of 30 peers would have died in the past 2 years.

The balance of transitions into, and out of, dependent use of cocaine is equally important.

3.5 *Dependent use of cannabis*

With few exceptions, surveys have asked respondents about their use of cannabis (ever, in past year, in past month) without attempting to monitor trends in dependent or hazardous use of cannabis.

First, there needs to be an agreed definition of “dependent” cannabis (D-C) use before respondents can be asked about their D-C use (ever, in past year, in past month); and

cessation thereof. Secondly, since psychiatric morbidity is a primary concern in relation to D-C use, there needs to be agreement on the set of psychiatric diagnoses to be inquired about⁴² so that survey respondents can be asked to date: i) their 1st and ii) their most recent hospitalisation for a diagnosis that falls within the set. Thirdly, database linkage of new drug treatment clients whose main drug is cannabis can identify D-C users, their initiation year into D-C use, and track their prior and subsequent psychiatric admissions.

3.6 Number of HCV-contaminated injections since last HCV negative test

The risk of HCV transmission per HCV-contaminated injection is high - around 2% to 3% - and HCV carriage is highly prevalent among current injectors – around 30% to 50%.

Assuming 3% and 50%, an injector who was HCV antibody negative at his last test has an estimated 1 in 4 chance of being HCV-infected after 20 shared injections, or after 10 HCV-contaminated injections. To focus on injection-related HCV incidence, new questions are required, such as:

- a) month and year of last HCV antibody test; & test result
- b) since last HCV antibody test, number of shared injections {0; 1-5; 6-10; 11-15; 16-20; 21-30; more than 30; **don't know**}
- c) since last HCV antibody test, number of shared injections with persons known to be HCV carriers {0; 1-5; 6-10; 11-15; 16-20; 21-30; more than 30; **don't know**}

Questions b) and c) may only be answered reliably by those whose last HCV antibody test was in the past 2 years, but asking them of all injectors conveys a message about the counts that they should keep for themselves. Counting would be easier if a syringe colour – orange, say – was designated for use (optional) by those who know themselves to be HCV-infected. Extra questions could then be asked about the use of orange syringes.

New HCV infections are, of course, most likely among injectors. Until HCV testing focuses on incident as well as prevalent infections, injectors won't either. HCV sero-conversions (antigen positive, antibody negative) currently go undetected when injectors come forward for testing because it is only those who are HCV antibody positive who are then tested for HCV-antigen (HCV-RNA)³⁸. All injectors, like all blood donors, should be tested for HCV-RNA.

3.7 Drug users' offending, drugs spend, injecting, and treatment/interventions – in the community and during incarceration

Drug users' offending pattern, their spending on illegal drugs, injecting, and drug treatments referrals were asked about quite extensively in the Arrestee Survey and, in the mid 1990s, in surveillance studies of injectors in the community in Glasgow⁴⁴. Willing anonymous surveillance of prisoners' HIV/HCV risk behaviours, prevalence and incidence¹⁷ has been in abeyance for a decade. New questions are now needed, together with re-consideration of saliva versus finger-prick-blood samples. Saliva assures high volunteer rate and minimized risk for staff, but blood allows recent HCV sero-conversions to be identified.

Questions about detoxification or other drug treatment have merit as not all prisoners will have received substitute prescribing in the outside community, let alone in prison: see **BOX 1**. Also important are questions which allow estimation of: a) injector incidence since last release from prison and b) HCV incidence since last self-reported HCV test date. Drugs spend in prison warrants a brief question, see **BOX 5**.

3.8 Parental use of alcohol or drugs

Surveys of school-children, householders, or injectors in the 21st century should ask about parental use of alcohol and drugs (as known to the respondent)⁴. In the British Crime Survey, comparison could be made between a parent's perception of their child's use of drugs with the child's self-reported use; and the parent's views on, and use of, alcohol and drugs with the child's self-reported use. The extent of alignment in parent-child perceptions and usage may be the basis for new public health campaigns.

BOX 5: Questions for prison-based willing anonymous HIV/HCV surveillance to learn about inside-injecting in the past 4 weeks, detoxification, dependencies, injector incidence since last release, and HCV incidence since last HCV test.

Question	Draft wording
A	have you been in prison for the past 4 weeks?
B	have you been in prison before this incarceration?
C	have you ever injected illegal drugs?
D	have you ever injected inside prison?
E	number of inside-injections in the past 4 weeks?
F	number of heroin-days in the past 4 weeks?
G	number of uses of sterilization tablets in the past 4 weeks to clean needles and works?
H	number of days of using illicit methadone in the past 4 weeks?
I	{E&W only} have you been subject to rMDT in the past 4 weeks? (<i>IF</i> yes, were you positive for i) cannabis, ii) heroin, iii) prescribed methadone, iv) illicit methadone?)
J	are you currently receiving prescribed methadone in prison?
K	did you undergo detoxification at the start of this incarceration?
L	were you dependent on heroin at the start of this incarceration?
M	were you dependent on cocaine at the start of this incarceration?
N	were you dependent on alcohol at the start of this incarceration?
O	how much did you spend on drugs in prison in the past 4 weeks?
P	{recidivists only} number of months on the outside between previous release and start of this imprisonment?
Q	{recidivists only} number of months between previous release and now?
R	{recidivists only} up to the time of your previous release, had you ever injected illegal drugs? <i>{in combination with earlier questions, P to R allow injector incidence to be estimated for recidivists}.</i>
S	have you ever had a personal test for Hepatitis C? (<i>IF</i> yes, month and year of your most last personal HCV test & test result)

4 New Prospects

Recent technical developments, together with essential new questions, open up *both* new linkages (of data-sources and analysts) *and* value-added analyses of existing data-sources. In broad policy terms, this means:

- i) better understanding of existing data, and especially their potential for generating and testing new hypotheses,
- ii) making better use of indexing methods to link health and other registers, including in criminal justice by use of PNC numbers; and ‘*safe-havens*’ to do it,
- iii) maximizing access to event dates by existing cohorts to compensate for their losses-to-follow-up; and cost-efficiently creating ‘virtual cohorts’ (which use only event-dates, without self-reported context) to generalize from cohort studies,
- iv) understanding trajectories better, and risk factors, in analysing increasingly problematic drug use, co-morbid heavy drinking and mental health problems,
- v) using insights gained from surveys, databases and cohorts to devise interventions that are likely to be cost-effective; and put them to test in formal experiments.
- vi) better use of collected biological samples for measuring incidence and for genome-wide scanning.

4.1 *Evidence-synthesis across surveys*

Together, UK survey teams can develop a joint-analysis protocol to resolve issues that are not powerfully, or generally, answerable within a single target population or survey year. Besides more methodological topics, questions of policy interest that could be addressed in this way include:

- i) identifying calendar trends in specific drug use by age, sex and deprivation or geography; likewise for multiple drug use and, by pooling several years’ data, for ethnic minority respondents.
- ii) deriving risk-scores to identify early those to whom interventions could be targeted with the aim of reducing future problematic use of alcohol or drugs and risk of mental health problems^{10, 20,45,52}; and assessing transferability across survey populations.
- iii) given the rate of change of survey estimates, what inter-survey intervals are appropriate in the 21st century.

Surprisingly, we do not know HCV prevalence in 13 year-olds - but we know very precisely prisoners’ largely unchanged opiate positive rate in mandatory testing. Consideration should be given to whether the likely benefits of *linking* of saliva samples to be tested unattributably for HCV antibodies (or drugs) *with* respondents’ answers in surveys of householders and school-children warrant the associated cost, and any negative impact on response rate. *Linkage* could be trialled with a random one third of households or school-children. Costs and information-yield could, for example, be compared with those for 60,000 rMDTs per annum from prisoners in England and Wales.

Surveys in addiction-at-risk settings, such as accident and emergency or psychiatric admissions, have been limited thus far. Valid consent may be problematic in both settings, so that an initial approach may need to be classification-only, by use of residual

blood or urine samples taken for another purpose (but testable for drugs/alcohol/ blood-borne infections) and linked only to age-group, sex and associated diagnosis.

4.2 Database linkages and ‘virtual’ cohorts

‘Virtual cohorts’ use database linkage to identify past and future event-dates. They can generalize existing cohorts which, due to costly follow-up, may have insufficient members for estimating low event-rates precisely. Database linkage can also be used to back-up existing cohorts whose members have been lost from individual follow-up; or to reduce the burden on respondents by not asking them to recall event-dates for which there is a more reliable, objective record elsewhere (such as dates of all incarcerations in Scottish prisons since 1995).

More extensive database linkages are in prospect than have been tried to date: Examples of the potential for this type of linkage include prisoner releases linkable to morbidity, infections and drug treatment registers as well as to the deaths register; and National Treatment Agency (NTA) or Scottish Drugs Misuse Database (SDMD) clients linkable to mortality, morbidity, infectious diseases, and criminal justice registers such as PNC. The second example can be seen as a generalization of the National Treatment Outcomes Research Study, its successor (DTORS) and Scottish counterpart (DORIS).

Although the same linkages may be involved (HCV diagnoses with drug treatment referrals, say), it is the master-file of clients about whom other dated events are ascertained that defines the study. The eligibility criteria which qualify an index (client) for being listed on the study’s master-file are therefore of key importance. Finding out about the morbidity and mortality of HCV diagnosed patients involves linking the master-file of HCV diagnoses to drug treatment referrals (on SDMD, say) but, for monitoring reduction of injection-related harms, a master-file of indices for eligible injector-clients (from SDMD, say) might be the starting-point, and matched to HCV diagnoses.

Linkage of mater-indexed health to PNC records could proceed if PNC had the facility to derive a master-index from the offender’s name, sex and data of birth.

Linkage of repeat incarcerations of the same individual in England and Wales has been a requirement for over a decade. A neat interim solution, would be for the PNC number, which is increasingly used by prisons in England and Wales, to be used to ‘compute and store’ the longitudinal record of receptions, transfers and releases for PNC-indexed individuals who were incarcerated from, say, 1 January 2009 onwards. This *PNC-for-prisons concept* is due to Professor Tim Millar at Manchester University.

BOX 6: Each prison maintains a ‘live’ Local Inmate Data System (LIDS). These local systems feed into the national (England and Wales) Inmate Information System (IIS). Although designed to provide ‘live’ information, the IIS maintains a permanent record of the basic information required to identify periods when individuals were incarcerated, namely: date of incarceration, date of final release, dates of temporary release, the individual’s Police National Computer (PNC) number (for linkage with other criminal justice data) together with information suitable to generate a **many-one** ‘attributor’ code (initials, sex, birth date) required for possible case-linkage with systems such as the National Drug Treatment Monitoring System.

By mid-2010, the above will be replaced by a national live system (C-NOMIS) which also includes a permanent record of the data items necessary for case-linkage. Implementation will be gradual, and with an interface to the legacy systems, in order to maintain a national picture during the roll-out phase. C-NOMIS has an extensive ad hoc reporting capability but the scope to develop tailored reports from IIS is more limited.

However, Ministry of Justice analysts are routinely provided with IIS extracts which, in principle, contain relevant information for case-linkage. For example, an extract describes all those incarcerated, and another all those released, during the year to date. Linkage of year-end extracts for, perhaps, the past five years could provide a sufficient record of both incarcerations and releases.

Historically, recording of PNC numbers within LIDS has been very patchy, and even the data items to generate 'attributor' codes may have been problematic. LIDS is currently being audited for data quality in preparation for the roll-out of C-NOMIS, with missing data being gradually 'back-filled'. Hence the quality and completeness of the underlying data held within the live system has recently improved.

4.3 Youth cohorts: evidence-synthesis protocol and resource for randomized trials

Evidence workshop: There is great potential for youth cohorts whose primary purpose was not drugs science to convene an evidence workshop to consider how, by modest additions to their current plans - in terms of actual questions posed, database linkages, or biological sample acquisition - they could collaborate to give more powerful insights than individually to, say, drug transitions for different birth-cohorts: born in 1970, in an intermediate year, in the millennium.

Do the same transitions, and prognostic factors^{10,20,45,52}, hold for cohorts of at-risk individuals (injection drug users, for example, among whom critical events are many) as are suggested by population-recruited cohorts (among whom the events of interest are relatively few: unless combined across cohorts)? The evidence workshop could also focus on how to design recruitment to a targeted, at-risk youth cohort in which, for research efficiency, incident-event rates are nearer 10% than 1%.

Resource: Youth cohort members, or those randomly selected to take part in surveys, may be ideally suited to be invited to take part in randomized "public-information" trials: their forthcoming cohort/survey interview offers the opportunity to assess the impact of a randomized intervention (different versions of a drugs harm reduction leaflet, say) to which they were recently exposed.

Virtual youth cohorts: Scotland has around 50,000 births per annum; England and Wales 10 times as many. Using only database linkage, different 'virtual' national birth-cohorts of 50,000 to 500,000 12 year olds could be followed for subsequent and antecedent dated events with relative ease. For example, resident children who were: *born in 1980 with follow-up from 12th until their 25th birthday in 2005; or born in 1985 with follow-up from 12th to 25th birthday in 2010; or born in 1990 with follow-up from 12th to 25th birthday in 2015.*

Their longitudinal event-date record could comprise, inter alia: 1st school exclusion, left-school date, 1st benefits claim, 1st live-born child, 1st marriage, 1st divorce, 1st alcohol-related hospital admission, 1st arrest, 1st community sentence and associated offence-type, 1st incarceration together with all subsequent release and reception dates, 1st psychiatric

hospital admission, 1st drug treatment referral by main drug of misuse (self; and 1st for spouse – if applicable), together with all hospital discharges and (if applicable) cause of death coded according to 9th or 10th International Classification of Diseases.

4.4 Formal experiments in criminal justice

Formal experiments, particularly on the impact of criminal justice-orchestrated interventions for drug-dependent clients, have been under-utilized so that the relative effectiveness and cost-effectiveness of interventions have not been quantified. For policy relevance, formal experiments need to have external, as well as internal, validity. Also, qualitative research can be designed into formal experiments.

There has been a lack of well-thought-through formal protocols. Pilot studies have focused on how to improve *the process of delivering* interventions rather than on assessing an intervention's *impacts on major outcomes*: reconvictions (cost and severity thereof), employment, and drugs-related deaths to name but three.

Drug treatment and testing orders, community orders with a drugs-rehabilitation requirement, electronic tagging in its various guises, drugs testing on arrest for acquisitive crime and drugs courts all lack formal experiments, yet address offenders who have high morbidity or mortality as well as being costly to the public purse in terms of their criminal careers^{12,57}. Reduced reconvictions – if reductions there be – may, or may not, be sufficient to deliver cost-effectiveness. Could cost-effectiveness be improved upon if the intervention were differently delivered? For example, an enhanced, more costly version of the intervention may sufficiently improve effectiveness that – despite costing more – it could be shown to be incrementally cost-effective in NICE terms. The criminal justice budget may not be able to afford to give the enhanced intervention to all (restorative justice⁶⁷, say). If so, then randomization is *the fair means* of allocating a scarce resource; and, as a scientific bonus, continues to grow an objective evidence-base.

The growing number of former heroin users maintained on substitution medication is evidence of ‘unfinished business’ for better approaches to rehabilitation to counter⁶⁶.

4.5 Genome-wide association scans

Genome-wide association scanning has looked into the heritability of alcoholism but dependent use of specific illegal drugs has received less attention. For the most part, *dependent use* has to be either a) self-reported (as yet, few surveys have inquired into it), b) identified by referral to a drug treatment agency, or c) deduced from a morbid or fatal consequence (such as drugs-related death or suicide). There are progression intervals and ascertainment biases which differentiate a) to c). The results of genome-wise association scans could be different in the three ascertainment settings, and so careful thought needs to be given to the sourcing of samples for case series.

Having decided *where* to source samples, consent for DNA-extraction and genome-wide scanning will generally be required for prospectively-collected voluntary biological samples. *With-consent* introduces its own potential bias but also enables optimal choice

of biological sample – currently, 3-5ml of fresh blood or ORAGENE-collected saliva sample.

With proper approvals, retrospectively-collected stored residual saliva or blood samples may be used *without consent if without attribution* (other than classification, based on database-linkage-ascertained minimal drugs or criminal justice history).

For injectors, a range of stored samples exists and also samples are being collected prospectively. And so, if genome-wide association scans in respect of heritable propensities to “injecting” (or HCV carriage) were the objective, blood donor control series (who are never-injectors) are certainly appropriate but pilot studies on the success rate of sufficient DNA-extraction from without-consent residual stored samples and with-consent prospectively-collected fresh samples is where to start.

References

1. Academy of Medical Sciences (chairman: Professor Sir Gabriel Horn FRS) *Brain Science, Addiction and Drugs*. The Academy of Medical Sciences, May 2008.
2. Ades AE, LuG, Claxton K. Expected value of sample information calculations in medical decision making. *Medical Decision Making* 2004; 24: 207 – 227.
3. Ades A, Sutton AJ. Multiple parameter evidence synthesis in epidemiology and medical decision-making: current approaches. *Journal of the Royal Statistical Society Series A* 2006; 169: 5 – 35.
4. Advisory Council on the Misuse of Drugs (chairman: Professor Sir Michael Rawlins). *Cannabis: Classification and Public Health*. Home Office, April 2008.
5. Advisory Council on the Misuse of Drugs (chairman: Professor Sir Michael Rawlins). Response to government’s consultation on *Drugs: Our Community Your Say*. ADMD: 19 October 2007.
6. Advisory Council on the Misuse of Drugs (chairman: Professor Sir Michael Rawlins). *Pathways to Problems: Hazardous Use of Tobacco, Alcohol and Other Drugs by Young people in the UK and its Implications for Policy*. ACMD, September 2006.
7. Advisory Council on the Misuse of Drugs (chairman: Professor Sir Michael Rawlins). *Hidden Harm: Responding to the Needs of Children of Problem Drug Users*. Home Office, May 2003. (See also *Hidden Harm: Three Years On*. ACMD, 2007)
8. Advisory Council on the Misuse of Drugs (chairman: Professor Sir Michael Rawlins). *Reducing Drug Related Deaths*. London, The Stationery Office: 2000.
9. Appleby L, Shaw J, Kapur N, Wingfuhr K. et al. *Avoidable Deaths: Five Year Report by the National Confidential Inquiry into Suicide and Homicide by People with Mental Illness*. December 2006 (http://www.medicine.manchester.ac.uk/suicideprevention/nci/information_sharing/avoidable_deaths_full_report.pdf).
10. Barnes G, Mitic W, Leadbeater B, Dhimi M. Substance use and mental health symptoms. *Canadian Journal of Community Mental Health* (in press).
11. Bird AG, Gore SM, Cameron S, Ross AJ, Goldberg DJ. Anonymous HIV surveillance with risk factor elicitation at Scotland’s largest prison, Barlinnie. *AIDS* 1995; 9: 801 – 808.

12. Bird SM. Prescribing sentence: time for evidence-based justice. *Lancet* 2004; 364: 1457 – 1459.
13. Bird SM, Hutchinson SJ. Male drugs-related deaths in the fortnight after release from prison: Scotland, 1996 - 1999. *Addiction* 2003; 98: 185 - 190.
14. Bird SM, Leigh Brown AJ. Criminalisation of HIV transmission: implications for public health in Scotland. *British Medical Journal* 2001; 323: 1174-1177.
15. Bird SM, Lynskey M, English C, Donnelly J, Michael M, Treble R. Revisiting British Army Drugs testing, 2003-07: test thresholds, efficient targeting, and underlying trends. *Journal of the Royal United Services Institute* 2008; 153: 16 – 24.
16. Boyle, P, Feitjen P, Feng Z, Hattersley L, Huang Z, Nolan J, Raab G. Cohort Profile: The Scottish Longitudinal Study. *International Journal of Epidemiology* 2008; in press (<http://ije.oxfordjournals.org/cgi/content/extract/dyn087>).
17. Champion JK, Taylor A, Hutchinson S, Cameron S, McMenamin J, Mitchell A, Goldberg D. Incidence of Hepatitis C virus infection and associated risk factors among Scottish prison inmates: a cohort study. *American Journal of Epidemiology* 2004; 159: 514 – 519.
18. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. *Health Economics* 1999; 8: 269 – 274.
19. Cooper NJ, Sutton AJ, Abrams KR, Turner D, Waioloo A. Comprehensive decision analytical modelling in economic evaluation: a Bayesian approach. *Health Economics* 2004; 13: 203 – 226.
20. Costa FM, Turbin MS, Jessor R. Transitions into adolescent problem drinking: the role of psychosocial risk and protective factors. *Journal of Studies on Alcohol* 1999; 60(4): 480-490.
21. *Crime Statistics: An Independent Review* (chair: Smith AFM). Home Office, November 2006 (<http://www.homeoffice.gov.uk/rds/pdfs06/crime-statistics-independent-review-06.pdf>).
22. De Angelis D, Sweeting MJ, Ades AE, Hickman M. Estimating Hepatitis C prevalence in England and Wales by synthesising evidence from multiple sources. *Statistical Methods in Medical Research* 2008: in press (& online in November).
23. De Angelis D, Day NE, Gore SM, Gilks WR, McGee MA. AIDS: the statistical basis for public health. *Statistical Methods in Medical Research* 1993; 2: 75-91.
24. De Angelis D, Hickman M, Yang S. Estimating long-term trends in the incidence and prevalence of opiate use/injecting drug use and the number of former users: back-calculation methods and opiate overdose deaths. *American Journal of Epidemiology* 2004; 160(10): 994-1004.
25. *Drug Misuse Statistics Scotland 2006; and 2007*. NHS Scotland Information Services Division, Edinburgh 2006; and 2007.
26. Editorial. In rude health. A treasure-trove of data in the UK National Health Service is set to energize biomedical research. *Nature* 2008 (3 July): 454: 1 -2.
27. Farrell M, Marsden J. Drug-related mortality among newly released offenders, 1998 to 2000. *Home Office Online Report 40/05*.
28. Farrington DP, Coid JW, Harnett LM, Jolliffe D, Soteriou N, Turner RE, West DJ. Criminal careers up to age 50 and life success up to age 48: new findings from the Cambridge Study in Delinquent Development. *Crime Reduction Home Office Report 299*, November 2006.

29. Fu B, Tom BDM, Delahooke T, Alexander GJM, Bird SM. Event-biased referral can distort estimation of hepatitis C virus progression rate to cirrhosis, and of prognostic influences. *Journal of Clinical Epidemiology* 2007; 60(11): 1140 – 1148.
30. Fu B, Tom BDM, Bird SM. Re-weighted inference about hepatitis C virus-infected communities when analysing diagnosed patients referred to liver clinics. *Statistical Methods in Medical Research* 2008 (in press & online in November).
31. Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association* 2005; 100: 680 – 700.
32. Garthwaite PH, O'Hagan A. Quantifying expert opinion in the UK water industry: an experimental study. *Journal of the Royal Statistical Society Series D* 2000; 49: 455 – 477.
33. Godfrey C, Eaton G, McDougall C, Culyer A. The economic and social costs of Class A drug use in England and Wales, 2000. *Home Office Research Study 249*, London: Home Office, 2002. (<http://www.homeoffice.gov.uk/rds/pdfs2/hors249.pdf>).
34. Godfrey C, Stewart D, Gossop M. Economic analysis of costs and consequences of the treatment of drug misuse: two year outcome data from the National Treatment Outcome Research Study (NTORS). *Addiction* 2004; 99(6): 697-707.
35. Goubar A, Ades AE, De Angelis D, McGarrigle CA, Mercer C, Tookey P, Fenton K, Gill ON. Estimates of HIV prevalence and proportion diagnosed based on Bayesian multi-parameter synthesis of surveillance data (with Discussion). *Journal of the Royal Statistical Society, Series A* 2008; 171(3): 541 - 580.
36. *Growing Up in Scotland: A Study following the Lives of Scotland's Children*. (<http://www.scotland.gov.uk/Publications/2007/01/17162004/> . . .)
37. Hay G, Gannon M, MacDougall J, Eastwood C, Williams K, Millar T. Capture-recapture and anchored prevalence estimation of injecting drug users in England: national and regional estimates. *Statistical Methods in Medical Research* 2008 (in press & online in November).
38. Health Protection Agency. *Shooting Up*. Infections among injecting drug users in the United Kingdom 2006. An update: October 2007 (http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1194947406808).
39. Health Protection Scotland and University of the West of Scotland. *Needle Exchange Surveillance Initiative (NESI). Prevalence of HCV, HIV and injecting risk behaviours among injecting drug users attending needle exchanges in Scotland, 2007*. Health Protection Scotland, May 2008.
40. Henderson M, Wight D, Raab GM, Abraham C, Parkes A, Scott S, Hart G. Impact of a theoretically based sex education programme (SHARE) delivered by teachers on NHS registered conceptions and terminations: final results of cluster randomised trial. *British Medical Journal* 2007; 334: 133 – 136.
41. HM Government. *Drugs: Protecting Families and Communities*. Home Office, February 2008.
42. Hickman M, Vickerman P, Macleod J, Kirkbride J, Jones PB. Cannabis and schizophrenia: model projections of the impact of the rise in cannabis use on historical and future trends in schizophrenia in England and Wales. *Addiction* 2007; 102: 587-606 (with discussion: 515-518).
43. Hutchinson SJ, Bird SM, Goldberg DJ. Modelling the current and future disease burden of Hepatitis C among injecting drug users in Scotland. *Hepatology* 2005; 42: 711 – 723.

44. Hutchinson SJ, Gore SM, Taylor A, Goldberg DJ, Frischer M. Extent and contributing factors of drug expenditure of injectors in Glasgow. *British Journal of Psychiatry* 2000; 176: 166 - 172.
45. Jessor R, Van Der Bos J, Vanderryn J, Costa FM, Turbin M. Protective factors in adolescent problem behaviour: moderator effects and developmental change. *Developmental Psychology* 1995; 31: 923 – 933.
46. Kaplan EH. Snapshot samples. *Socio-Econ. Plann. Sci.* 1997; 31: 281 – 291.
47. King R, Bird SM, Brooks SP, Hutchinson SJ, Hay G. Prior information on behavioural capture-recapture methods: demography influences injectors' propensity to be listed on data sources and their drugs-related mortality. *American Journal of Epidemiology* 2005; 162: 694 – 703.
48. King R, Bird SM, Hay G, Hutchinson SJ. Estimating current injectors in Scotland and their drug-related death rate by sex, region and age-group via Bayesian capture-recapture methods. *Statistical Methods in Medical Research* 2008 (in press & online in November).
49. King R, Brooks SP. On the Bayesian analysis of population size. *Biometrika* 2001; 88: 317 – 336.
50. McAra L, McVie S. *Criminal Justice Transitions*. The Edinburgh Study of Youth Transitions and Crime, The Centre for Law and Society, University of Edinburgh 2007.
51. McDonald SA, Hutchinson SJ, Bird SM, Mills PR, Dillon J, Bloor M, Robertson C, Donaghy M, Hayes P, Graham L, Goldberg DJ. A population-based record linkage study of mortality in hepatitis C diagnosed persons with or without HIV co-infection in Scotland. *Statistical Methods in Medical Research* 2008: in press & online in November.
52. Moore R. Adult offenders' perceptions of their underlying problems: findings from the OASys self-assessment questionnaire. *Home Office Findings* 284, 2007.
53. National Institute for Health and Clinical Excellence. *Methadone and buprenorphine for the management of opioid dependence*. (www.nice.org.uk/TA114).
54. National Institute for Health and Clinical Excellence. *Naltrexone for the management of opioid dependence*. (www.nice.org.uk/TA115).
55. National Institute for Health and Clinical Excellence. *Drug Misuse: Psychosocial Interventions*. (www.nice.org.uk/CG051).
56. National Institute for Health and Clinical Excellence. *Drug Misuse: Opiate detoxification*. (www.nice.org.uk/CG052).
57. National Institute for Health and Clinical Excellence. *Needle and Syringe Programmes: consultation on the draft guidance* (<http://www.nice.org.uk/guidance/index.jsp?action=byID&o=11829>).
58. Nevill C, Hales J. *Longitudinal analysis of the Offending, Crime and Justice Survey*. Report on the Descriptive Analysis. (prepared for Home Office, June 2008).
59. O'Hagan T. Elicitation. *Significance* 2005; 2: 84 – 86.
60. Ohlssen DI, Sharples LD, Spiegelhalter DJ. Flexible random-effects models using Bayesian semi-parametric models: applications to institutional comparisons. *Statistics in Medicine* 2007; 26: 2088 – 2112.
61. Prevost TC, Abrams KR, Jones DR. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. *Statistics in Medicine* 2000; 19: 3359 – 3376.

62. Raab G, Storkey H, Bird S. *Sample size calculations for a group-randomized trial to evaluate a drug prevention programme*. Report prepared for Home Office Blueprint Team, October 2002.
63. Razali K, Amin J, Dore CJ, Law MG on behalf of HCV Working Group. Modelling and calibration of the hepatitis C epidemic in Australia. *Statistical Methods in Medical Research* 2008: in press & online in November.
64. Roe S, Ashe J. Young people and crime: findings from the 2006 Offending, Crime and Justice Survey. *Home Office Statistical Bulletin 09/08*, July 2008.
65. Royal Statistical Society Working Party on Performance Monitoring in the Public Services (chair: Bird SM). Performance Indicators: Good, Bad and Ugly. *Journal of the Royal Statistical Society Series A* 2005; 168(1): 1- 27 (see also rss.org.uk).
66. Scottish Government. *Road to Recovery*. Edinburgh, May 2008.
67. Shapland J, Atkinson A, Atkinson H, Dignan J, Edwards L, Hibbert J, Howes M, Johnstone J, Robinson G, Sorsby A. Does restorative justice affect reconviction? The fourth report from the evaluation of three schemes. *Ministry of Justice Series 10/08*, June 2008.
68. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statistics in Medicine* 2003; 22: 3687 – 3709.
69. Sweeting MJ, De Angelis D, Hickman M, Ades AE. Supplementary material to: Estimating Hepatitis C prevalence in England and Wales by synthesising evidence from multiple data sources. Assessing data conflict and model fit. *Biostatistics* 2008; 9(4): 715-734 (see also doi: 10.1093/biostatistics/kxn004).
70. Sweeting MJ, De Angelis D, Ades AE, Hickman M. Estimating the prevalence of ex-injecting drug use in the population. *Statistical Methods in Medical Research* 2008: in press & online in November.
71. Sutton AJ, Cooper NJ, Jones DR, Lambert PC, Thompson JR, Abrams KR. Evidence-based sample size calculations based upon updated meta-analysis. *Statistics in Medicine* 2007; 26 (12): 2479-2500.
72. Sutton AJ, Gray NJ, Edmonds WJ, Andrews NJ, Hope VD, Gill ON. Modelling the characteristics of the male injecting drug user population in England and Wales. *International Journal of Drug Policy* 2005; 16: 176 – 182.
73. Thomas DC, Haile RW, Duggan D. Recent developments in genome wide association scans: a workshop summary and review. *American Journal of Human Genetics* 2005; 77(3): 337 – 345.
74. Thomas R, Walport M. *Data Sharing Review*. Ministry of Justice, July 2008.
75. Turner RM, Spiegelhalter DJ, Smith GCS, Thompson SG. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society Series A* (in press). (for pre-print, see <http://www.mrc-bsu.cam.ac.uk/BSUsite/Publications/Preprints/biasmodelling.pdf>).
76. Weaver T, Madden P, Charles V, Stimson G, Renton A, Barnes T, Bench C, Middleton H, Wright N, Paterson S, Shanahan W, Seivewright N, Ford G. Comorbidity of substance misuse and mental illness in community mental health and substance misuse services. *British Journal of Psychiatry* 2003; 183: 304-313.

Acknowledgements

We should like to thank the following people for their interest and valuable contributions:

Scientists:

Prof Mark Bellis	<i>Liverpool John Moores University</i>
Prof Colin Blakemore	<i>Oxford University</i>
Roger Bowles	<i>York University</i>
Prof Peter Boyle	<i>St. Andrews University</i>
Ms Sue Brooker	<i>NatCen, Home Office Surveys, Design and Statistics sub- committee</i>
Prof Valerie Curran	<i>University College London</i>
Dr Daniela de Angelis	<i>MRC Biostatistics Unit</i>
Dr Mandeep Dhami	<i>Cambridge University</i>
Prof Peter Diggle	<i>Lancaster University</i>
Mr Dave Elliot	<i>Office for National Statistics</i>
Prof Mike Hough	<i>King's College London</i>
Dr Sharon Hutchinson	<i>Health Protection Scotland</i>
Prof Brian Francis	<i>Lancaster University, Home Office Surveys, Design and Statistics sub- committee</i>
Dr Gavin Malloch	<i>Medical Research Council</i>
Prof Tim Millar	<i>Manchester University</i>
Prof David Nutt	<i>Advisory Council on the Misuse of Drugs</i>
Prof Sir Michael Rawlins	<i>Advisory Council on the Misuse of Drugs</i>
Prof John Strang	<i>National Addiction Centre, London</i>
Mr Andrew Sutton	<i>Warwick University</i>
Dr Michael Sweeting	<i>MRC Biostatistics Unit</i>
Prof Paul Turnbull	<i>King's College London</i>

Policy Leads and, particularly, *Science Secretariat at the Home Office*:

Ms Roisin Ash	<i>Scottish Government</i>
Mr Martin Barnes	<i>Druscope</i>
Ms Susannah Browne	<i>Home Office Crime, Drugs Analysis Research</i>
Ms Clare Griffiths	<i>Office for National Statistics</i>
<i>Ms Smita Kaur</i>	<i>Home Office Science Secretariat</i>
Mr Jonathan Knight	<i>National Treatment Agency</i>
Dr Marcus Roberts	<i>Druscope</i>
Prof Chris Robertson	<i>Health Protection Scotland</i>
Ms Nicola Singleton	<i>UK Drugs Policy Commission</i>
Mr Robert Street	<i>Home Office Crime, Drugs Analysis Research</i>
Ms Alison Walker	<i>Home Office</i>
Ms Claudia Wells	<i>Office for National Statistics</i>
<i>Dr Iain Williams</i>	<i>Home Office Science Secretariat</i>

Subject matter leads who presented:

Prof Tony Ades	<i>Bristol University</i>
Ms Laura Blackwood	<i>NHS Scotland</i>
Prof Sheila Bird	<i>MRC Biostatistics Unit</i>
Mr John Flatley	<i>Home Office</i>
Ms Elizabeth Fuller	<i>NatCen</i>
Dr Gordon Hay	<i>Glasgow University</i>
Dr Matt Hickman	<i>Bristol University</i>
Ms Lorna Jackson	<i>NHS Scotland</i>
Dr Ruth King	<i>St. Andrews University</i>
Dr Scott McDonald	<i>Health Protection Scotland</i>
Dr Fortune Ncube	<i>Health Protection Agency</i>
Ms Sara Skodbo	<i>Home Office Drugs Intervention Programme.</i>

Statistical ANNEX

'Virtual' cohorts using database linkage

Problem drug users often have lapsed contact with family and friends; involuntary contact with the criminal justice system; intermittent, and sometimes mandatory, contact with education, health, employment and social services. More or less chaotic life-styles mean that individualized follow-up of injectors or other problematic drug users is costly; and liable to informative drop-out.

The alternative is to use ethical database linkage to track the appearance of a client's master-index across a series of databases, and so build-up a dated-event history without recourse to individually-vouchsafed data (other than as recorded at unplanned event-dates – confidential data to which event-only database linkage would not give general access).

Probabilistic linkage across several databases inevitably has a higher cumulative risk of linkage-error in some part of the finally-linked record.

Particularly valuable would be to compare the finally-linked event-record for master-indices in a 'virtual cohort' which was defined, in fact, by the initial membership of an actual cohort of injectors or other problematic drug users who were enlisted to be followed-up by conventional interview or self-completion questionnaire techniques. DTORS may be an ideal platform for just such a study.

Elicitation methods

Blind alleys and naïve assumptions are avoided when statistical analysis proceeds in tandem with expert subject-matter knowledge.

Expert subject-matter knowledge is helpful in determining model structure: how complex does a statistical model need to be to have face value, let alone subtlety, for practitioners. Expert statistical knowledge may overlay additional complexity that practitioners had not realised a need for - such as anticipation that covariate influences are not constant over time (because they almost never are . . .).

Subject-matter expertise is especially valuable in providing plausible parameter values, such average duration of injecting career, the likely proportions of injectors who would permanently cease injecting within 3 years of their injecting debut, or be still injecting 20 years after commencement.

Methods have been developed for how best to elicit prior information from non-statistician probability assessors, and for how best to document the formal elicitation process⁵⁹. Some experts turn out to be poorly-calibrated assessors (their beliefs do not even accord with known data; or are not internally coherent). It is prudent to build redundancy into the elicitation – such as asking about injector incidence, prevalence and outcidence over time – so that that analyst is aware of incoherencies, and can adapt analyses to deal with them. Statisticians should also remember that the expert whose opinion is outlying may in time be shown to be correct.

Commonly-held beliefs are not necessarily correctly-held.

Formal elicitation of prior opinion is one of the range techniques used in determining, a priori, the plausible effect size that a sufficiently-powerful (say, 80% power) randomized controlled trial of a new treatment or criminal justice intervention for drug-dependent offenders should be designed to discern (by the yardstick of statistical significance at, say, the 5% level).

Projection methods – in epidemics of transmissible infections

Projection of the late sequelae of, say, injection-related Hepatitis C virus (HCV) transmission^{43,63} needs information or assumptions on injector-incidence and the age distribution at injecting debut (in different calendar periods), on the number (and switching) of needle-sharing partners and frequency of shared

injections with such partners (in different calendar periods); assumptions about off-injecting and about the death rates of current versus former injectors; information on the transmissibility of HCV, and on covariate-dependent progression from HCV carriage to late liver sequelae. Besides using elicitation and empirical data-sources to estimate prior parameter ranges in the stochastic model, other key data, such as how HCV prevalence in current injectors changes over time³⁸ and how the number of current injectors evolves⁴⁸, may be used in overall calibration. Thus, the posterior ranges for parameters which jointly yield stochastic scenarios that are consistent with the calibration constraints may differ importantly from the prior ranges. Clearly displaying these differences is important. Also inter-dependencies among the accepted parameters bear careful scrutiny.

Stochastic simulation models can be used to project forward the consequences of different public health scenarios, such as injector incidence being unchanged from 2000 but **x%** reduction in the number of shared injections per annum, **y%** shorter injecting careers and **trebling** of the HCV diagnosis and treatment rates for former injectors . . .

Overlay of costs on projection models allows the incremental cost-effectiveness of different interventions to be gauged according to NICE guidance.

Other less nuanced, mathematically-tractable approaches to projection can also be useful⁴², particularly if they incorporate different publicly-held or policy assumptions which give rise to widely different projections for what to expect in 2010 or 2015. Chickens soon come home to roost . . . and thereby we'll learn which scenario was correct.

There is no exclusivity of modelling approaches. Fundamental to all is the need to be crystal clear, and explicit, about simplifying assumptions: how tenable or realistic the simplification is, what impact its relaxation is likely to have on the results, and what additional data – or methodology - may be needed to permit its relaxation.

Multi-parameter evidence-synthesis

Multi-parameter evidence-synthesis comes into its own when there is a diversity of data-sources^{3,22,35, 69,75} which, together, are sufficient not only to estimate all key parameters but also sufficiently rich that there is redundancy, which means that several data-sources inform about the same parameter or set of parameters, so that conflicts within the data and between data and simplifying assumptions can be diagnosed, and resolved – either by acquiring new data or by realising that previously-accepted simplifying assumptions are not compatible with the evidence.

De Angelis et al.²² applied multi-parameter evidence synthesis to the complex problem of estimating HCV prevalence by sex, age, risk group and region in England. Complexity arose particularly because there is a dearth of direct information about former injectors^{69,70}, not least because HCV databases do not record year of starting to inject, let alone of cessation. England's register of HCV diagnoses by sex, age, risk group and region is incomplete; information on HCV prevalence in current injectors is from non-representatively sampled agency settings, and in heterosexuals from those attending genito-urinary medicine clinics or pregnant women, but some of each will have injected drugs; the size of risk-groups has to be determined, such as via capture-recapture based estimation of the regional numbers of current injectors; and representatively-sampled household surveys such as National Attitudes and Sexual Life-styles and British Crime Survey provided potentially-biased information on the ratio of current to former injectors.

Despite the complexity, impressive progress was made but key simplifying assumptions were also made about the duration of injecting careers being invariant whether the injector's debut was in the early 1980s or early 21st century. If true, the assumption's public health implication is profoundly depressing since dual goals for public health interventions should be to delay the age at starting to inject and to hasten off-injecting. A next iteration of the multi-parameter evidence synthesis could well consider whether the data are sufficient at least to support the estimation of a change-point after which injectors' careers were shorter. There is a strong prior presumption of the direction of change at least, albeit uncertainty about when harm reduction measures were demonstrably effective in shortening injectors' careers – an answer worth striving for.

Inverse-probability weighting methodologies to redress ascertainment bias

Cross-sectional, or snapshot, samples⁴⁶ – of ever-injector clients referred to drug treatment agencies or prison inmates or arrest events or HCV-diagnosed injectors referred to liver clinics^{29,30} – are common but require careful analysis if generalized inferences are required about the duration of *all* injecting careers (not just about the ascertained sample), those received into prison (not inmates), arrested persons (not arrest events), incubation period to cirrhosis for all HCV-diagnosed injectors (not just those referred to liver clinics).

Kaplan⁴⁶ defined a snapshot sample as “constructed at a fixed chronological time either by sampling only subjects where the initial event has occurred but the subsequent event has yet to occur (*active subjects*, for example: new client ever-injectors at drug treatment agency), or by sampling only subjects where both the initial and subsequent events have occurred (*inactive subjects*, for example: HCV-diagnosed injectors who have been referred to a liver clinic and have developed cirrhosis)”. Although snapshot samples are biased (*inactive subjects* towards shorter active times than occur in nature; *active subjects* towards longer active times than occur in nature), recognising the biases and dealing with them analytically enables more correct general inferences to be drawn.

Appropriate methodologies for sorting the ascertainment biases depend, for example, on whether “capture” (attendance at needle exchange; attendance for drug treatment; attendance at liver clinic) is *equally likely* throughout an injector’s career (effectively, *uniform sampling*); or is more likely the closer the individual is to the end of their incubation period (as applied for HCV-diagnosed injectors - who were referred to a liver clinic predominantly in the last half of their HCV incubation period to cirrhosis^{29,30}).

The probability arguments which underlie methods that correct for ascertainment bias are non-trivial: for example, it matters how tenable the assumption of *uniform sampling* is. Simulation studies are useful – seeing (ascertainment bias in action) is believing how large the biases can be; and to check the performance of analytical methods in recovering correctly the ‘true’ parameterisation. The best fitting of a suite of statistical models may fail to fit the data well⁷², if something is awry in the basic model structure. Statistical diagnostics matter.

Design effects in surveys matter. They are a related issue – in the sense that they use inverse-probability weighting to redress different sampling fractions, such as over-sampling of 16-24 years olds in British Crime Survey.

APPENDIX of TABLES

Table 1.1 Commonly-available biological samples and tests pertaining to drugs science

Biological Sample testable for:	Newborn's		Child/adult's (test sensitivity relative to blood)			
	Blood	Saliva or buccal swab	Blood	Saliva or buccal swab (sensitivity)	Urine (sensitivity)	Breathylser (sensitivity)
Examples of virological tests						
<i>Maternal HIV antibodies</i>	<i>M*</i>	?				
HIV antibodies			*	* (> 90%)		
HIV antigen		?	*			
<i>New HIV infection</i>	<i>I*</i>		<i>I*</i>			
<i>Maternal HCV antibodies</i>	<i>M*</i>	?				
HCV antibodies			*	* (> 85%)		
HCV antigen	<i>I*</i>	?	*			
<i>New HCV infection</i>			<i>I*</i>			
<i>HBV antigen</i>	<i>I* but can Immunize</i>		*	?		
Examples of dual tests – for medical or criminal justice purpose						
<i>DNA Genetics</i>	*	*	*	*	?	
<i>Maternal Alcohol</i>	<i>M*</i>					
Alcohol (recent use)			*		* (> 95%?)	* (> 90%?)
Alcohol (long-term heavy use)			*			
<i>Maternal Methadone</i>	<i>M*?</i>					
Methadone (recent use)			*	*	*	
Examples of tests for illegal drugs						
<i>Maternal Heroin</i>	<i>M*?</i>					
Heroin (recent use)			*	* (> 80%)	* (> 95%?)	
<i>Maternal Cocaine</i>	<i>M*?</i>					
Cocaine (recent use)			*	* (> 80%)	* (> 95%?)	
Ecstasy (recent use)			*	* (> 80%?)	* (> 95%?)	
Cannabis (10-day use)			*	* (> 90%?)	* (> 95%?)	

Table 1.2a Major UK surveys and associated biological samples: by target population

Target population	New-borns (re mother)		Antenatal women	Genito-urinary medicine clinic attenders		Psychiatric admissions; & Forensic post-mortems.
Residual sample survey	Guthrie heel-prick blood test		Booking blood test	Syphilis blood test		NIL
Setting	Scotland: hospitals	E&W: hospitals	E&W: hospitals	Scotland: hospitals	E&W: hospitals	Hospitals; & mortuaries
Representative of setting ?	Census	selected	selected	city census	Selected	
Age-group & sex	Reproductive age & female	Reproductive age & female	Reproductive age & female	Any age & both sexes: repeats within x months of previous test excluded		15-44 years & both sexes
Response rate within setting	Essentially 100%		Low opt-out rate	Essentially 100% as low opt-out rate		
Sample size	~ 50,000	~ xx,000	~ yy,000	~ zz,000	~ w,000	See Section 4
Identifier	C = age-group & region.	C = age-group, ethnicity & region.	C = age-group, ethnicity & region.	C = {?ethnicity in E&W} age-group, sexual orientation, gender, ever-IDU & region.		C = age-group, ICD10 diagnosis, gender, ever-IDU & region.
Qs re injecting	No		No	Yes		
Qs re drugs	No		No	Ever injected drugs? Ever prostituted?		
Qs re alcohol	No		No	No		
Biological sample(s)	Heel-prick blood-spot		Blood	Blood		
Tested for	Maternal HIV antibodies {& HIV-RNA in new-born if mother is HIV-infected}		HIV & HCV antibodies	HIV & HCV antibodies; & {E&W only} HIV-RNA to detect recent HIV infection		
Extension: 1 (see Section 4)	Maternal HCV antibodies			Saliva sample to be tested for illegal drugs		Saliva sample to be tested for HCV antibodies and for illegal drugs
Extension: 2 (see Section 4)	Opiates if mother is HCV-infected		Opiates if HCV-infected	HCV-RNA if ever-IDU to detect recent HCV infection		Urine sample to be tested for recent alcohol
Extension: 3 (see Section 4)	New-born's blood alcohol {depends on delay from birth to heel-prick blood sample}					Blood to be tested for long-term heavy alcohol & for HCV.

Table 1.2b Major UK surveys and associated biological samples: by target population

Target population	Current injectors		Prisoners: *minimum of 50 per classification		Prisoner surveys under ONS auspices		
	interview	interview	self-Q	self-Q	Interview re rMDT	Interview Physical Health	Interview Mental Health
Setting	Scotland, needle exchange schemes (NES)	E&W, drug treatment agencies, NES, other locations	Scotland's annual prisoner survey	Scotland; also E&W: neither since 1990s.	E&W, prisons {NB interview survey re opiates and injecting + hair sample in 1990s by Strang et al. [REF]}		
Representative of setting	Selected	Selected	Census	Major; or selected prisons	Yes	Yes	Yes
Age-group & sex	Any age & both sexes	Any age & both sexes	Any age & both sexes	Any age & both sexes	Any age & both sexes	Any age & both sexes	Any age & both sexes
Response rate within setting	Not reported	No reported	~ 70%	> 80%	75%	?	88%
Sample size	~ 600	~ 3,000	~ 5,000	~ 3,000; ~ 3,000 (cumulative)	2,270	?	3,142
Identifier	C = age-group, sex, & region	C = age-group, sex, setting & region	C = age-group, sex & prison	C = age-group, sex & prison*	?	?	See MDT survey response
Qs re injecting	yes	Yes	Yes	yes	Yes	?	Yes
Qs re drugs	yes	Yes	Yes	Yes	Yes	?	Yes
Qs re alcohol	?	?	?	No	?	Yes	Yes
Biological sample(s)	Saliva	Saliva & trialling finger-prick blood spot	<i>See, however, 1990s' saliva+self-Q WASH-C surveillance</i>		Saliva & hair	? blood	No
Tested for	HIV+HCV antibodies	HIV+HCV antibodies & trialling HIV+HCV antigen	Not Applicable	HIV+HCV antibodies	Cannabinoids & opiates	?	
Extension: 1 (see Section 4)			21 st C non-attributable saliva +self-Q to measure incidences (IDU & BBVs) & prevalences		Last done in 2001	Last done in 1994	Last done in 1997
Extension: 2 (see Section 4)			Trialling (un)attributed finger-prick blood re incidences + self-Q				

Table 1.2c Major UK surveys and associated biological samples: by target population

Target population	School children			Adult population			
Volunteer interview/self-completion Q surveys re nutrition, health or addictions	Self-Q	Self-Q	Self-Q Diet & Nutrition	Self-Q general health	Self-Q physical health	Self-Q mental health	Self-Q Food Standards Agency
Setting	Scotland, schools	E&W, schools	E&W, household	Scotland, household	E&W, household	Great Britain household	E&W, household
Representative of setting	Yes, xx% schools	Yes, xx% schools	Yes, xx% households	Yes, xx% households	Yes, xx% households	Yes, xx% households	Yes, xx% households
Age-group & sex							
Response rate within setting	YY% pupils	YY% pupils	ZZ% children	YY% adults	YY% adults	69% adults	YY% adults
Sample size							
Identifier							
Qs re injecting	No	No	No	no	No	? yes	
Qs re drugs	yes	Yes	No	? no	? no	Yes, including dependence	
Qs re alcohol	yes	Yes	No	yes	Yes	Yes, including Dependence	
Biological sample(s)	No	No	? blood	? blood	? blood	No	?
Tested for			?ask Dunn Nutrition				
Extension: 1 (see Section 4)	Consider saliva test for HCV at 11-13yrs						
Extension: 2 (see Section 4)							

Table 1.2d Major UK surveys and associated biological samples: by target population

Target population	Arrestees		16+ years of age in general population		Offender/Victim Surveys		National Survey of Sexual Attitudes and Lifestyles
Volunteer interview/self-completion Q surveys re criminal justice & addictions	Drugs or alcohol-related crimes and access to drug treatment.		Crime Surveys: see also 1998 Youth Lifestyles Survey (12-30 year olds) nested within British Crime Survey - it over sampled cities & high crime areas		Offender Crime & Justice Survey	Offender Crime & Justice Survey Panel re 10-25 year olds	
Setting	Scotland	E&W	Scotland	E&W	E&W		UK
Representative of setting	Selected, 199x only	Yes, xx% of eligible police suites	Yes, xx% households	Yes, xx% households			Yes, xx% households
Age-group & sex			16+ years of age & both sexes; ethnic & 16-24 year over-sampling in E&W		10-65 years	10-25 years	?
Response rate within setting	Approximately 1 in 4		YY% of eligible respondents	YY% of eligible respondents			YY% of eligible respondents
Sample size							
Identifier							
Qs re injecting	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Qs re drugs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Qs re alcohol	Yes	Yes	?	?	Yes	Yes	?
Biological sample(s)	Urine	Saliva	No	No	No	No	Urine (attributed)
Tested for	Drugs	Drugs					Chlamydia
Extension: 1 (see Section 4)	Nil: terminated	Nil: terminated		10-15 year olds in same-as-adult household	Nil: terminated	Longitudinal follow-up	Nil: terminated
Extension: 2 (see Section 4)	Further analysis of drugs spend and criminality by injection/non-injection use of drugs						

Other prison databases concern Counselling, Assessment, Referral, Advice and Throughcare services (CARATs), which provide basic drug intervention services for prisoners in England and Wales, and the Offender Assessment System (OASys) across prisons and probation services in England and Wales. The OASys self-assessment questionnaire asks simply if ‘taking drugs’; ‘drinking too much alcohol’ is a problem for you and, if so, is this problem linked to your offending. See **Table 1.3** for self-reported likelihood of further offending versus OASys reconviction score in just over 100,000 assessments. Actual re-conviction within 2 years of assessment is needed to complete this story . . .

Table 1.3 Perceived likelihood of re-offending versus OASys ‘likelihood of reconviction’ score (to nearest 100): but actual outcome matters!

Likely to offend again (self-perception)	OASys likelihood of re-conviction scored		
	Low: n = 31,600	Medium: n = 49,700	High: n = 20,000
Definitely not	20,900	19,400	3,600
Unlikely	10,100	23,300	9,400
Quite or very likely	1,000	7,000	6,800

Since 2007, the Drug Interventions Programme (DIP) has introduced an updated Drug Interventions Record and associated forms which aim to harmonize data collection across CARATs and Crime and Justice Intervention Teams.

Figure 1.1 illustrates how database linkage can shed light on *different* propensities to respond at age 20 years for 5,000 girls, 1,000 of whom were classified as at ‘high risk’ of teenage pregnancy (on the basis of smoking, alcohol, cannabis, truanting profile by 15 years of age). In contrast to naïve inference, the totals revealed by database linkage of high risk females who had had a live-born child or abortion (LBA) by 20 years of age argued more strongly for interventions to be targeted on girls classified as high risk by their profile up to 15 years of age. Comparison between self-reports and database linkage revealed a tendency for high risk women to under-report LBA if they responded, which only 20% did, and for their non-response to be associated with a substantially higher LBA rate, whereas non-high risk women who were non-respondents had by far the lowest LBA rate.

Figure 1.1 Uncovering the biases hidden behind non-response					
Risk category at 15 years	Age 20: Self-Q returned	By age 20: Self-Q confirms live-birth/abortion (SQ-LBA)	Naïve Inference	By age 20: database linkage live-birth/abortion (DB-LBA)	Revised Inferences (uses returns and database linkage)
1,000 high risk	200 returns (20% return-rate)	100 (50% SQ-LBA rate)	50% of high risk females have LBA by age 20.	120 (60% DB-LBA rate)	<p>A. 84% high risk females, {120+720}/1000, have LBA by age 20</p> <p>B. High risk females who did not return their self-Q had a substantially higher LBA rate (90%) than high risk respondents (60%)</p> <p>C. Even those high risk females who responded under-reported LBA (100 versus 120 by database linkage)</p>
	800 no self-Q	?	<p>Generalise to estimate a total of 500 high risk LBAs.</p> <p>{Database linkage reveals 840}</p>	720 (90% DB-LBA rate)	
4,000 others	1,800 returns (45% return-rate)	180 (10% SQ-LBA rate)	10% others have LBA by age 20.	180	<p>D. 7.3% of other women had LBA by age 20, namely {180+110}/4000</p> <p>E. Non-high risk women who did not return their self-Q had substantially lower LBA rate (5%) than respondents had (10%) – the opposite of high risk women.</p>
	2,200 no self-Q	?	<p>Generalise to estimate 400 LBAs among all others.</p> <p>{Database linkage reveals 290}</p>	110 (5% LBA rate)	

Table 1.4a Cohort studies, Scotland: 1

Cohort (MRC code)	Edinburgh Psychiatric Genetics Group (16)	Edinburgh High Risk (15)	Edinburgh Addiction Cohort (14)	Edinburgh Study of Youth Transitions and Crime (17)	Aberdeen Children of the 1950s Cohort Study (2)	Generation Scotland (21)
Population-based v. at-risk	At-risk cases, relatives, & controls	At-risk: well as recruited but 2 identifiable close relatives with schizophrenia	At-risk: injection drug users	Population-based: targeted whole school-year cohort – aged 12 yrs in 1998	Population-based: children born in Aberdeen in 1950-56 and in Aberdeen primary school in 1962	Adult population-based: Aberdeen, Edinburgh, Glasgow, Tayside
Geographical location	South Scotland	Scotland, but excludes Northern Isles	Muirhouse General Practice, NW Edinburgh	City of Edinburgh schools	City of Aberdeen schools	Aberdeen, Edinburgh, Glasgow, Tayside & community-based
Recruited at:	Any site	Any site	GP & community-based			(NB: family study: to be clarified)
Approximate age in 2000: & cohort-start year	Any age: & recruited since ~ 1996	Aged 16-24 years & well on recruit since 1994 (22-30 yrs in 2000): since 1994	Aged about 16-29 years on recruit since 1984 (32-45 yrs in 2000): since 1984	Aged 14 years in 2000: cohort start-year = 1998	Aged 44-50 years in 2000: cohort start = 1962, revitalised in 1998	Aged 18+ years in 2006 (15,000 to be recruited in phase 1, up to 50,000 by end phase2)
Number recruited @ baseline	500 schizophrenia 500 bipolar 500 depression	223 index cases (& 2 relatives each)	800 ever-injectors (depleted by HIV and IDU-related mortality)	Recruited 4,317= 89% of target cohort	12,150 (unknown % of target cohort)	Target of 50,000
Number @ follow-up in year y	1,000 controls 1,000 relatives	180 index cases in 2004		Self-Q replies by 3,525 in 2003 (80.5% of eligibles in 2003)	Self-Q in 2000-02 (unknown % response by eligible survivors)	As above because follow-up by national registers
Biological samples?	Blood; DNA	Blood; DNA (ongoing); Imaging	Blood	No	No	Blood; DNA
Addictions?	Not reliable	Smoking, alcohol, cannabis/ other drugs	Smoking, Alcohol, cannabis/ other drugs	Smoking, Alcohol, cannabis/ Other drugs	Smoking, Alcohol.	Smoking, Alcohol, Education.

Mental health?	Yes	Yes	Yes	Yes	Yes, by database linkage	Yes, access to patient records
Crimes?	No	No, but personality disorders	“life-style” & personality + conduct disorders	Yes, e.g. via Scottish Criminal Record Office	No	No
Database linkage?	No	Yes: national registers	Yes: national registers	Yes: census and ‘official records’	Yes: morbidity & mortality	Yes: national registers
Cost (e.g. per recruit; per incident).	Not reported	£4.5millions (£750 per person-year)	Latest grant was £200K (latest: ~ £160 per person-year)	£1.14m (~ £60 per person-year)	Latest grant was £500K (Latest: ~ £20 per person-year)	£6.2m over 3 years for Phase 1 (Phase 1: ~ £140 per index person-year)
COMMENT	Consider database linkages - : feasible re SDMD, Scottish Criminal Records & prison terms.	CLOSED to recruitment	High HIV-prevalence cohort in early/mid 1980s	ORAGENE (postal) saliva sample, or finger-prick blood as basis for DNA-extraction. Self-Q rate drops-off in school-leavers.	Unknown self-Q response rate after 38-40 years’ elapsed time. Cohort too old for Scottish prison terms to be retrieved by database linkage.	Potential for database linkage to Scottish Criminal Record Office, benefits, prescriptions (Tayside) & prison terms.

Table 1.4b Cohort studies, Scotland: 2

Cohort (MRC code)	Growing up in Scotland (23)	Sexual Health & Relationships – Safe, Happy And Responsible: SHARE (37)	West of Scotland 11-16 & 16+ Study (41)	West of Scotland Twenty-07 Study (42)	Drug Outcomes Research in Scotland (not MRC-listed).	
Population-based v. at-risk	Population-base with representative sampling; later waves to follow same design, eg 2007/08	Secondary school-based: cluster-randomized controlled trial	Population-based: age 11 at primary schools (1994/95)	Population-based: said to be community accrual re ‘social patterning of health’ – cohorts born in 1930s, 1950s, 1970s.	At-risk , because, inter alia, registered as new client with Scottish Drugs Misuse Database	
Geographical location	Scotland	East of Scotland	Central Clydeside Conurbation of 1.5m	Central Clydeside Conurbation	Scotland	
Recruited at:	Home, accessed via child benefit	25/26 secondary schools: target of 8,430 pupils aged 14 years	135 primary schools	? community-based	?	
Approximate age in 2000: & cohort-start year	N/A: 10month birth-cohort and 34month toddler cohort in 2005/06 (wave 1)	Youngest were aged 17 in late 2000/early 01: ~ 1993/94.	17-year olds in 2000/01.	In their 30s for youngest cohort born in 1970s; otherwise 50s or 70s.	? 30 years	
Number recruited @ baseline	5,000 birth-cohort; 3,000 toddler-cohort. Recruited as % invited not stated.	7,616 (95% of those targeted).	2,586 (93%) 11-year olds from target sample of 2,793 in 1994/95.	4,510 (assume roughly 1,500 per cohort; no volunteer rates cited.)	1,000	
Number @ follow-up in year y	N/A	Record linkage was achieved for 4,120/4,195 young women. Self-Q at age 20 was received back from 33% eligibles, or 37% of those recruited.	2,196 (85%) of recruits when followed-up at age 15 years in secondary school but down to 1,258 (49%)	2,661 (59% of originally accrued, but no allowance for deaths)	????	

			by ages 18-20 years.			
Biological samples?	Saliva	No (female age at 1 st pregnancy or termination was outcome)	No	No	No	
Addictions?	Smoking, Alcohol, cannabis/ other drugs	Smoking, Alcohol, cannabis/ other drugs	Smoking, Alcohol, cannabis/ other drugs	Smoking, Alcohol, cannabis/ other drugs	Smoking, Alcohol, cannabis/ other drugs	
Mental health?	Unclear	No	Yes	Yes	?	
Crimes?	Unclear	No	No	No	Yes, prison	
Database linkage?	Stated as 'patient; obstetric'	National registers, census & patient.	No	National registers & patient.	National registers	
Cost (e.g. per recruit; per incident).	£2m for development & wave 1+2 recruitment (£125 per wave 1+2 recruit)	£1m for development, RCT & analysis. (~ £20 or £40 per RCT child for follow-up of 6 or 3 years)	~ £220K, also MRC Unit	MRC Unit core funded	?	
COMMENT	Recruitment waves & follow-up planned to ages 16-20 years. Essential to establish database linkages.	Self-Q identified characteristics of those lost to attrition & enabled weighting to be used to address attrition. Data shared with SE England RIPPLE study. Consider impact of RCT assignment on males' SDMD, crimes, prison terms & early mortality.	Compensate for low self-Q response rate by database linkage re dated-events. However, low event-numbers to be expected because population-based and overall sample size therefore too low.	Compensate for low self-Q response rate by database linkage re dated-events. Prison terms only recoverable for the youngest birth cohort from 1995.	Biological samples likely to be stored for many DORIS participants.	

Table 1.4c Cohort studies, England & Wales: 1

Cohort (MRC code)	Peterbr' Adolescent Development Study (not MRC-listed) follows cross-sectional Peterbr' Youth Study, 2002}	Offender Crime & Justice Survey Longit. Panel (not MRC-listed)	Sheffield Pathways Out of Crime Study (not MRC-listed)	Cambridge Study in Delinquent Development (not MRC-listed)	National Treatment Outcomes Study (32)
Population-based v. at-risk	?	Aged 10-25 yrs living in private households	At-risk: born ~ 1983 & at least 2 recorded conviction occasions.	School & area-based	At-risk, because registered as new client for drug treatment
Geographical location	Peterborough	England & Wales	South Yorkshire	Working class inner city, London	England & Wales
Recruited at:	? schools	Households, annually from 2003 with panel of respondents retained fro re-contact	?	6 state primary schools: boys aged 8-9 yrs in 1953-54.	Community, hospital, or drug treatment agency
Approx. age in 2000: & cohort-start year	~ 9 yrs old in 2000: & started in March 2003	Cohort-start was 2003	~ 17 yrs old in 2000: & started in 2003 {9mly interviews to 24 yrs}	~ 47 yrs in 2000: & cohort-start year = 1961	34.3 years in 2000: & started in 1995
Number recruited @ baseline	707 12-year old boys and girls	? for example, 4,554 panel respondents in 2006 + 799 new	Target of at least 250 males & 50 females	411 boys	1,075
Number @ follow-up in year y	? 4-year follow-up thro' 15 years of age (until 2007)	?	?	93% of survivors to 2001 interviewed	Stratified sample at 5 years: 496 (76% of target)
Biological samples?	No	No	No	No	No
Addictions?	?	? smoking Alcohol cannabis other drugs	?	?	Smoking Alcohol, cannabis Other drugs

Mental health?	?	? Yes	?	Yes	Yes
Crimes?	Yes	Yes	Yes	Yes	No, but 'life-style'
Database linkage?	?	Yes	Yes, to PNC Assumed	Mortality & PNC, assumed	National registers & patient
Cost (e.g. per recruit; per incident).	?	?	?	?	£1.7m (£1,600 per recruit re 5 years, or £320 per recruit-year)
COMMENT	Limited by initial low number of clients.		Limited by volunteer rate & losses to follow-up risk.	Highly successful follow-up. Limited generalization.	ENDED High cost, given no RCT intervention & no infectious disease focus.

Table 1.4d Cohort studies, England & Wales: 2

Cohort (MRC code)	National Treatment Outcomes Study (32)	Drug Treatment Outcomes Research Study (13)	Disability conditions & registration for Child Abuse & Neglect: West Sussex (12)	Blueprint Drug Education Research Programme (5)	Determinants of Adolescent Social Well-being & Health: DASH (11)
Population-based v. at-risk	At-risk , because registered as new client for drug treatment	At-risk, adult problem drug users at tx. services for new course of tx.	Geographic population-based: birth-cohort 01/83 to 12/01	School-based. Local Education Authority-comprehens. secondary schools	School-based in London boroughs
Geographical location	England & Wales	England (100 Drug Action Team areas)	West Sussex Primary Care Trust	Cheshire, Derby & D-shire, Lancashire: 23 intervention, 6 comparison secondaries.	Brent, Croydon, Hackney, Hammersmith & Fulham, Haringey. Lambeth, Newham, Southwark, Waltham Forest, Wandsworth
Recruited at:	Community, hospital, or drug treatment agency	Treatment services	Birth, or soon thereafter	Comprehens. secondary school at ages 11, 12 & 13.	School but recruitment age unclear: assume secondary
Approx. age in 2000: & cohort-start year	34.3 years in 2000: & started in 1995	18m follow-up, aged 25 years in 2000; start year 2005.	Various - up to 17 years: & cohort-start in 1983	~ age 10 in 2000: & cohort-start was 2002. Non-RCT.	Unknown: & cohort-start was 2003
Number recruited @ baseline	1,075	3,000 (accrual rate as % invitees not known)	119,729 births	4,500 pupils (? accrual rate as % invitees)	6,652 pupils (81% response rate)
Number @ follow-up in year y	Stratified sample at 5 years: 496 (76% of target)	Not known response-rate at 6m interview	Not stated	Not stated; 2006-outcomes due in spring 2008.	4,656 in 2005/06 (70% of initial recruits)

Biological samples?	No	No	No	No	No
Addictions?	Smoking Alcohol, cannabis Other drugs	Smoking Alcohol, cannabis other drugs & quality of drug tx.	No	Smoking Alcohol cannabis other drugs	Smoking Alcohol cannabis other drugs
Mental health?	Yes	Yes	Yes, eg ADHD and autism	No	Yes
Crimes?	No, but 'life-style'	Yes, in self-Q	Yes, child abuse/ Neglect	?	No
Database linkage?	National registers & patient	National registers (includes PNC, excludes prison?)	Child abuse/ Neglect registration in West Sussex: 1,853 registered.	?	Linkage to NHSCR for vital events, but ? re morbidities.
Cost (e.g. per recruit; per incident).	£1.7m (£1,600 per recruit re 5 years, or £320 per recruit-year)	£2.1m (£700 per recruit re 18m, or £460 per recruit-year)	Not stated	£7.5m includes BLUE-PRINT cost (£1,667 per child, or £400 per child-year)	~ £530K (£80 per child, or £30 per child-year)
COMMENT	ENDED High cost, given no RCT intervention & no infectious disease focus.	DAT has access to clients' results of random urinary drugs tests.	Develop-ment milestones were collected.	FINISHED High cost, given non-RCT & too few children to determine impact on class A. (scientific advice = DON'T.)	Ongoing and so costs may increase. Consider PNC-linkage & to child/other benefits.

Table 1.4e Cohort studies, England & Wales: 3; & Great Britain or UK-wide

Cohort (MRC code)	South-ampton Women's Survey (38)	Twins Early Development Study: TEDS (39)	Avon Longit. Study of Parents & Children ALSPAC (4)	Gates-head Millen Study: GMS (20)	Great Britain National Child Development Study (31)	Great Britain British Cohort Study 1970: BCS70 (8)	UK-wide: Millen Cohort Study: MCS (26)	UK-wide: BIO BANK (40)
Population-based v. at-risk	Population-based	Population-based	Population-based	Popul. based	Popul. based: born in 1 week March 1958	Popul based: born in 1 week April 1970	Popul based E&W = born Sept00 to end Aug01 S+NI = born 24N00 to 10Ja02	UK Popul based
Geographical location	South-ampton	England & Wales	Avon, Greater Bristol	Gates-Head Borough, Tyne & Wear	Great Britain	Great Britain	UK	UK
Recruited at:	Unclear	Via Twins Register: representative sample	All children born in 21months & their parents.	'normal infants' recruit at birth	? birth	? birth	? birth	? by postal contact; Aged 40-69 years
Approx. age in 2000: & cohort-start year	22-36 yrs in 2000: & cohort-start year =1998 (also their 1 st born after recruited)	5 years: & cohort-start = 1995 {E-risk sub-cohort of 1100 families : how genetic and environmental factors shape children's disruptive behaviour}	~ 10 years in 2000: cohort-start year = 1990	1 year: & cohort-start year = 1999/0	42 yrs in 2000: cohort-start year = 1958	30 yrs in 2000: cohort-start year = 1970	0 yrs in 2000: cohort-start year = 2000.	34-63 yrs in 2000: cohort-start year = 2006.
Number recruited @ baseline	12,500 women & 2,800 babies (target =3,000)	15,000 pairs of twins from infancy to adolescent, 16 yrs.	12,000 children (assumed) – follow-up to age 18 years at least	1,029 accrued (? % accept rate)	~17,500 (? % accept rate)	17,500 (? % accept rate)	19,245 (? % accept rate)	Target = 0.5m (accept rate ~ 10%). Number so far ??
Number @ follow-up in year y	? annual data collection	At 2, 3, 4, 7, 9, 10, 12, 14 &	8,000 children (67%) @	830 in 2005/06	10,000 in 2004/05 (phone	10,000 in 2004/05 (phone	15,511 at 6 years of	?

		16yrs in 2011.	13 years		interview); aim for 12,000 when f-to-f in 2008	interview); aim for 12,000 when f-to-f in 2008	age: 81% of recruits	
Biological samples?	Blood; DNA	DNA; imaging	Blood; DNA; saliva	Blood	Blood; DNA; saliva	No	Saliva	Blood; DNA; imaging
Addictions?	Smoking, Alcohol	Smoking Alcohol	Smoking Alcohol Cannabis Other drugs	No	Smoking Alcohol Cannabis Other drugs	Smoking Alcohol Cannabis Other drugs	Smoking Alcohol Cannabis Other drugs	Smoking Alcohol
Mental health?	depression	Yes, with ADHD	Yes	Eating disorder	Yes	Depressed	Depressed ADHD	Yes
Crimes	No	No, but education	No	No	No, but education	No, but education	No, but education	No, 'lifestyle' & environmental & education
Database linkage?	Obstetric	National registers, census, obstetric	National registers & census; patient	Not cited	Census, deaths	Deaths	National registers, obstetric, patient	National registers FULL PATIENT RECORD
Cost (e.g. per recruit; per incident).	£2.5m with ancillary projects, & core MRC Unit support (~ £160 per subject)	£4.6m in direct MRC costs in 1995-2010, plus other grants (~ £200)	£20m for 13 years follow-up (~ £50 per subject-yr; or ~ £150 per child-yr)	£800K (~ £100 per child-yr)	£400K re phone; £1.4K re f-to-f (~ £80 per recent contact)	£400K re phone; £1.7K re f-to-f (~ £95 per recent contact)	£12m (~ £80 per recent contact)	£58m (~ £120 per subject)
COMMENT	<i>Consider NDTA linkage, HES, benefits & PNC.</i>	<i>Longer-term: consider NDTA linkage, HES, benefits & PNC.</i>	Worrying one-third loss to follow-up by 13 years of age – even before out-of-school.	Too small to stand alone.	<i>Consider HES or morbidity linkage, PNC, SDMD or NDTA.</i> <i>Consider biological samples re 1970 cohort, if not already done.</i>		As for earlier birth cohorts	VERY LOW OPT-IN RATE

Table 1.5 National databases and biological sample collections on drugs science.

Database, or biological sample collection	Centralise or not?	Nations	Coverage?	Mandatory, or voluntary	Eligibility	Data-designer	New entries in 2007	Total to end 2007
Hepatitis C diagnoses (blood)	Yes	Scotland to HPS & two reference virus labs.	Near complete	Voluntary	HCV test request & confirmed HCV antibody positive	HPS	~ 1,500	~ 22K
	No	England to HPA	Incomplete			HPA	?	?
HIV diagnoses (blood)	Yes	Scotland to HPS & two reference virus labs.	Near complete	Voluntary	HIV test request & confirmed HIV antibody positive (HIV –ve samples & also database in Scotland)	HPS (all test requests)	?	?
	No	England to HPA	Near complete			HPA	?	?
Guthrie heel-prick blood spots	Yes	Scotland to HPS & two reference virus labs.	Near complete	Voluntary but very low opt-out	Live newborn to screen for cystic fibrosis inter alia – baby has maternal HIV/HCV Antibodies	HPS, classify-only	?	?
	No	England to HPA	Incomplete to HPA			HPA, classify-only	?	?
Syphilis blood test at GUM	Yes	Scotland to HPS & two reference virus labs.	Near complete	Mandatory	GUM clinic attenders, both sexes	HPS, classify-only	?	?
	No	England to HPA	Incomplete to HPA: selected GUMs only			HPA, classify-only	?	?
Injection drug users (saliva or finger-prick blood) at needle exchange	No	Scotland to HPS & two reference virus labs.	Non-representative	Voluntary	Those who attend venues & consent	Self-Q or interview	~ 2K in 2007/08	?
		England to HPA					~ 3,000 a year	~ 21K since 2000

or Drug Treatment Agencies								
rMDTs (urine)	Yes to LGC & results to MOJ	England & Wales	5% to 15% of inmates per prison per annum	Mandatory	Prison selects rMDT rate; inmates selected by NOMS-team?	LGC & MOJ	~ 60K	Nearly 400K since 2000
British Army's CDTs (urine) (& other services)	Yes to LGC & results to CDT-HQ	UK and overseas	Variable by rank and year	Mandatory	Mainly, CDT-HQ selects units & personnel to be tested	LGC & CDT-HQ	~ 90K	Nearly 700K since 2000 in British Army
Police (saliva drugs tests)	No	Local police suite in E&W; pilot in Scotland	Unknown	Mandatory	Offenders arrested for acquisitive crime	Not Known (NB: chain of custody)	?	?
Courts (random urinary drug tests wrt DTTOs & DRRs)	No	Local courts & DATs	Supposed to be 'random'; frequency set by courts.	Mandatory that results reported to court; testing by 'compact'	Offenders on DTTOs or DRRs	Not known – assume@ local DATs	?	?
							>> 6K DTTOs per annum	> 42K DTTO clients
Blood transfusion services (donors: ever-IDUs self-defer)	No, per centre	Scottish BTS	Complete	Mandatory HCV-RNA & HIV tests	New & repeat donors	Scottish BTS	?	?
	No, per centre	National Blood & Transplant Service				National Blood & Transplant Service	?	?
SDMD ?	No	Scotland	NIL	?	?	ISD	?	?
NDTA ?		E&W		?	?	NDTA	?	?

Table 1.6a First three questions about injection drug use.

Study	B = BBV surveillance S = Survey C = Cohort	Q1.	Q2.	Q3.	Comment
1994 WASH-C, HMP Barlinnie	B.	In which year did you first inject drugs (excluding insulin)?	In which year did you last inject?	Have you ever injected while inside?	Grouped answers: NEVER INJECTED In 1982 or earlier In 1983 to 1985 In 1986 to 1988 In 1989 to 1991 In 1992 or later NK
DTORS, 1st follow-up	C	First of all, have you injected any drugs in the past 4 weeks?	How often in the last 4 weeks did you inject heroin? Was it . . .	How often in the last 4 weeks did you inject unprescribed methadone? Was it . . .	Grouped answers: daily most days 3 or 4 days a week 1 or 2 days a week less than once a week Not injected heroin in the last 4 weeks
British Crime Survey {AW to check}	S	<i>Single question on injecting?</i>			Single question on injecting?
Survey of Smoking, Drinking and Drug Use: Schoolchildren in England, 2007	S				No question on injection drug use
Needle Exchange Surveillance Initiative (NESI)	B				

Table 1.6b First three questions about heroin use.

Study	B = BBV surveillance S = Survey C = Cohort	Q1.	Q2.	Q3.	Comment
1994 WASH-C, HMP Barlinnie	B.				No questions on heroin
DTORS, 1st follow-up	C	Which of these drugs are you receiving [FORM of TX] for?	During break from structured tx., which drugs were you using?	Which of the drugs on the card have you taken in the last 4 weeks? (then, how often each used)	Answer list: Heroin Methadone (not prescribed to you) Other opiates Crack cocaine ... etc
British Crime Survey {AW to check}	S	Have you EVER taken HEROIN (SMACK, 'H'. BROWN), even if it was a long time ago?	In the last 12 MONTHS have you taken HEROIN (SMACK, 'H', BROWN)? {later: asks how often}	In the LAST MONTH have you taken HEROIN (SMACK, 'H', BROWN)?	Answer list: Yes No Never heard of it (Q1 only) Don't want to answer
Survey of Smoking, Drinking and Drug Use: Schoolchildren in England, 2007	S	Have you ever heard of Heroin?	Have you ever been offered Heroin?	Have you ever tried Heroin (even if only once)?	Next 2 Qs: a) how old were you when you first tried Heroin? (age then [] years) b) when did you last use or take Heroin? (last month, year, more than a year ago)
Needle Exchange Surveillance Initiative (NESI)	B	Have you ever been prescribed methadone? {If yes: Has this been in last 6 months? If yes: for how many of the last 6 months}	How old were you when you first injected drugs? {& What year was that?}	In which year did you last inject drugs? {& did you inject in last 6 months, in last month?}	Next 2 Qs: in the months when you injected drugs, a) how often on average did you inject? & b) % injections with new & unused needle