



Ministry of
JUSTICE

**A compendium of research and analysis on
the Offender Assessment System (OASys)
2006-2009**

Mia Debidin (Editor)

Ministry of Justice Research Series 16/09
December 2009

A compendium of research and analysis on the Offender Assessment System (OASys) 2006-2009

Mia Debidin (Editor)

**This information is also available on the Ministry of Justice website:
www.justice.gov.uk/publications/research.htm**

The OASys Data Evaluation and Analysis Team (O-DEAT) supports effective policy development and delivery within the National Offender Management Service and Ministry of Justice by providing high-quality social research and statistical analysis, and aims to publish information to enable informed debate.

Disclaimer

The views expressed are those of the authors and are not necessarily shared by the Ministry of Justice (nor do they represent Government policy).

© Crown Copyright 2009.

Extracts from this document may be reproduced for non-commercial purposes on condition that the source is acknowledged.

First Published 2009

ISBN: 978-1-84099-328-8

Foreword

The Offender Assessment System (OASys) was introduced in 2001, building upon the existing 'What Works' evidence base. It combines the best of actuarial methods of prediction with structured professional judgement to provide standardised assessments of offenders' risks and needs, helping to link these risks and needs to individualised sentence plans and risk management plans.

OASys has improved and helped to join up assessment practice across prisons and probation, providing a basis for defensible decision making and supporting effective Offender Management. By identifying offending-related needs and assisting with the targeting of offenders to interventions, OASys has contributed to the reduction of reoffending. OASys data has been used at the local, regional and national levels for resource planning, with more than 3.5 million OASys assessments now collated within the joint OASys database.

While OASys was piloted prior to implementation, the intention was that as the evidence base developed the system would be improved over time. I am therefore pleased to publish this research compendium produced by the OASys Data Evaluation and Analysis Team (O-DEAT) which has contributed to some important revisions to OASys.

OASys will continue to play a key role in the delivery and evaluation of interventions and Offender Management, while ensuring that resources are used efficiently and effectively across the Prison and Probation Services.

Phil Wheatley
Director General
National Offender Management Service

Contents

Foreword	i
1 Introduction to OASys and research on OASys 2006 to 2009 Mia Debidin and Laura Fairweather	1
2 The current evidence base for the offending-related risk factors included in OASys Sarah Morton	10
3 Measurement of OASys rates of completion Sarah Morton	25
4 The internal reliability and construct validity of OASys Robin Moore	37
5 Inter-rater reliability of OASys Sarah Morton	56
6 Predictive validity of OASys – Improving prediction of violent and general reoffending Philip Howard	78
7 Measuring changes in risk and need over time using OASys Philip Howard and Robin Moore	108
8 Predictive Validity of the OASys self-assessment questionnaire (SAQ) Robin Moore	136
9 The coverage and representativeness of OASys risk and need offender profiles: 2007 probation commencements and sentenced prisoner receptions Robin Moore	152
10 Exploratory research on the evidence boxes in the OASys ‘relationships’ and ‘lifestyle/associates’ sections Wendy Smith-Yau	180
11 Exploratory research on the evidence boxes in eight OASys sections assessing offending-related problems Sarah Morton and Wendy Smith-Yau	191
12 OASys statistics: 2008 probation and prison assessments OASys Data Evaluation and Analysis Team	200
13 Use of OASys data elsewhere and compendium conclusions and recommendations Mia Debidin and Robin Moore	222

References	235	
Appendix 1	Example of a Quality Assessment Tool and guidance notes on completion	251
Appendix 2	OASys scored items	254
Appendix 3	Underlying factors of current scored questions	257
Appendix 4	Revised underlying factors	259
Appendix 5	Scored OASys questions included within revised risk factors, criminogenic need scales, OGP and/or OVP (OASys sections 3 to 12)	260
Appendix 6	Logistic regression results and 100-point scales for OGP and OVP	262
Appendix 7	Distributions and correlations of existing and new predictors	266
Appendix 8	Diversity issues: accuracy and validity of prediction by age, sex and ethnicity	269
Appendix 9	Changes in individual scored items	272
Appendix 10	The OASys Self-Assessment Questionnaire (paper version)	274
Appendix 11	A restructured Self-Assessment Questionnaire	275
Appendix 12	OASys representativeness at regional level	276
Appendix 13	Probation commencements' classification decision tree models	284
Appendix 14	Risk and need residual values and percentages correctly predicted for classification decision tree models (probation commencements)	297
Appendix 15	Prison receptions' classification decision tree models	298
Appendix 16	Risk and need residual values and percentages correctly predicted for classification decision tree models (prison receptions)	311
Appendix 17	Example of the relationships textual analysis matrix	312
Appendix 18	Example of the lifestyle and associates textual analysis matrix	313

Tables

Table 1.1	The desired characteristics of risk-needs classification	5
Table 3.1	OASys items measured for data completion	27
Table 3.2	Completion rates for probation and prison assessments per year	29
Table 3.3	Percentage of prison assessments with insufficient scored questions completed for each section of the core OASys assessment	30
Table 3.4	The ten scored OASys questions with poorest completion rates for each service in each financial year	30
Table 3.5	Change in completion rates between 2005/06 and 2006/07 for each service	32
Table 3.6	Percentage of assessments that met minimum validity criteria in 2006/07	35
Table 4.1	Scored OASys scales	37
Table 4.2	Internal reliability of current OASys scales	41
Table 4.3	Item-scale correlations of current OASys scored questions	41
Table 4.4:	Scored questions in non-corresponding factors	43
Table 4.5	Internal reliability of revised OASys scales	46
Table 4.6	Item-scale correlations of remaining scored questions	46
Table 4.7	Associations with 24-month reoffending	47
Table 4.8	24-month reoffending rates by revised OASys scales	48
Table 4.9	None/low need / medium need cut-off points for revised OASys scales by gender and age	50
Table 4.10	Medium need / high need cut-off points for revised OASys scales by gender and age	51
Table 4.11	24-month reoffending rates by revised criminogenic need cut-offs	52
Table 4.12	Correlations with reoffending and odds ratios for clinical judgements, current scales and revised scales	53
Table 4.13	Current and revised criminogenic need prevalence rates	53
Table 5.1	IRR Case studies	59
Table 5.2	Number of respondents per IRR case study	60
Table 5.3	Likelihood of reconviction and risk of serious harm profiles for 2006/07 probation assessments and the respondent weighting	61
Table 5.4	Raw section scores for each IRR case study	64
Table 5.5	Percentage of assessors scoring no problem and a problem for each OASys question	66
Table 5.6	Normative, mean, median, standard deviation and likelihood of reconviction for each IRR case study	68
Table 5.7	Percentage of assessors scoring criminogenic needs for each IRR case study	70

Table 5.8	Percentage of assessors scoring no and yes for section links to serious harm and offending behaviour for each IRR case study	72
Table 5.9	Percentage of assessors scoring each risk of serious harm level for each IRR case study	73
Table 5.10	Percentage of assessors that decided the full risk analysis should be completed	73
Table 5.11	Results of intra-class correlation coefficients for each IRR case study	74
Table 6.1	Selection and matching of assessments	81
Table 6.2	Indications of violent content in a range of offences	87
Table 6.3	Criminogenic needs by offence type: possible violent offences and comparison offences	90
Table 6.4	Logistic regression models to predict proven reoffending based on previous sanction counts	92
Table 6.5	Scoring the OASys General reoffending Predictor (OGP)	94
Table 6.6	Likelihood of proven general reoffending by OGP score	94
Table 6.7	Scoring the OASys Violence Predictor (OVP)	95
Table 6.8	Likelihood of proven violent-type reoffending by OVP score	96
Table 6.9	Predicting proven reoffending using existing and new measures: Area Under Curve scores	98
Table 6.10	Accuracy and validity of OGRS 2 and 3, OASys score and OGP/OVP by age, gender and ethnicity	102
Table 7.1	Numbers of assessments by sentence, type of initial assessment and reconviction status	114
Table 7.2	Reasons for closure of assessment sequences	116
Table 7.3	Changes in individual items and criminogenic need scores	119
Table 7.4	Changes in likelihood of reoffending scores	120
Table 7.5	Prevalence of score changes for individual items	123
Table 7.6	Mean absolute and net changes for individual items	124
Table 7.7	Changes in criminogenic needs	126
Table 7.8	Changes in community risk of serious harm ratings	127
Table 7.9	Changes in highest community risk of serious harm rating by offender characteristics	128
Table 7.10	Overall OASys score changes between first and second assessment by reconviction status at third assessment	132
Table 7.11	Section score changes between first and second assessment by reconviction status at third assessment	133
Table 8.1	24-month reoffending rate by response to SAQ question 28	141
Table 8.2	Perceived likelihood of reoffending vs. actual 24-month reoffending	142
Table 8.3	24-month reoffending rate by responses to SAQ questions 1 to 27	143
Table 8.4	24-month reoffending rate by SAQ raw problems score	144
Table 8.5	24-month reoffending rate by SAQ question 28 response categories	145

Table 8.6	An SAQ logistic regression model for predicting reoffending	146
Table 8.7	A combined OGRS 3, core OASys and SAQ logistic regression model for predicting reoffending	148
Table 8.8	Accuracy of established predictors and logistic regression models	149
Table 8.9	Goodness-of-fit of combined OGRS 3, core OASys and SAQ model	149
Table 8.10	Residual values (from combined OGRS 3, OASys and SAQ model) for various offenders groups	150
Table 9.1	Offender characteristic variables	155
Table 9.2	OASys completion rates by post-sentence eligibility and region (probation commencements)	158
Table 9.3	Validity rates of completed OASys assessments by post-sentence eligibility and region (probation commencements)	159
Table 9.4	OASys completion rates by post-sentence eligibility and region (prison receptions)	160
Table 9.5	Validity rates of completed OASys assessments by post-sentence eligibility and region (prison receptions)	161
Table 9.6	OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements)	163
Table 9.7	Validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements)	165
Table 9.8	OASys completion rates by post-sentence eligibility and offender characteristics (prison receptions)	168
Table 9.9	Validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (prison receptions)	171
Table 9.10	Actual, predicted and adjusted risk and need levels by OASys post-sentence eligibility (probation commencements)	174
Table 9.11	Actual, predicted and adjusted risk and need levels by OASys post-sentence eligibility (prison receptions)	176
Table 10.1	The number and proportion of assessments for each stratum	184
Table 11.1	Summary of recommendations from textual analysis	199
Table 12.1	Criminogenic need prevalence rates by offender subgroups	204
Table 12.2	Planned intervention prevalence rates for offenders with scored criminogenic needs by offender subgroups	207
Table 12.3	Likelihood of reconviction and risk of serious harm levels by offender subgroups	211
Table 12.4	Risk of serious harm rates (full community breakdown) by offender subgroups	213
Table 12.5	Current concerns regarding risks to self by offender subgroups	216
Table 12.6	Number of self-assessed problems and perceived likelihood of reoffending by offender subgroups	219

Table 13.1	Recommendations from O-DEAT research findings and responses from OASys business team	228
Table A6.1	Ordinal logistic regression model of proven reoffending within 12 or 24 months: results for OGP offences	263
Table A6.2	Scaling OGP logistic regression results to produce a 100-point score	263
Table A6.3	Ordinal logistic regression model of proven reoffending within 12 or 24 months: results for OVP offences	264
Table A6.4	Scaling OVP logistic regression results to produce a 100-point score	265
Table A6.5	AUC scores for validation sample, comparing raw regression, original and revised score models	265
Table A7.1:	Descriptive statistics of risk predictors	267
Table A7.2:	Intercorrelations of new and existing measures of recidivism risk	268
Table A8.1:	Predictive validity for offences other than violence against the person, threats/harassment, weapon possession, violent acquisitive, criminal damage, and public order, by age group, gender and ethnicity	270
Table A8.2:	Predictive validity for violence against the person, threats/harassment, weapon possession, violent acquisitive, criminal damage and public order offences, by age group, gender and ethnicity	271
Table A12.1:	North East OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements)	277
Table A12.2:	London OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements)	279
Table A12.3:	North East validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements)	281
Table A12.4:	London validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements)	282
Table A13.1:	Likelihood of reconviction classification decision tree model (probation commencements)	285
Table A13.2:	Risk of serious harm classification decision tree model (probation commencements)	286
Table A13.3:	Accommodation need classification decision tree model (probation commencements)	287
Table A13.4:	Education, training and employability need classification decision tree model (probation commencements)	288
Table A13.5:	Financial management and income need classification decision tree model (probation commencements)	289
Table A13.6:	Relationships need classification decision tree model (probation commencements)	290

Table A13.7:	Lifestyle and associates need classification decision tree model (probation commencements)	291
Table A13.8:	Drug misuse need classification decision tree model (probation commencements)	292
Table A13.9:	Alcohol misuse need classification decision tree model (probation commencements)	293
Table A13.10:	Emotional wellbeing need classification decision tree model (probation commencements)	294
Table A13.11:	Thinking and behaviour need classification decision tree model (probation commencements)	295
Table A13.12:	Attitudes need classification decision tree model (probation commencements)	296
Table A14.1:	Goodness-of-fit and accuracy of classification models by OASys post-sentence eligibility (probation commencements)	297
Table A15.1:	Likelihood of reconviction classification decision tree model (prison receptions)	299
Table A15.2:	Risk of serious harm classification decision tree model (prison receptions)	300
Table A15.3:	Accommodation need classification decision tree model (prison receptions)	301
Table A15.4:	Education, training and employability need classification decision tree model (prison receptions)	302
Table A15.5:	Financial management and income need classification decision tree model (prison receptions)	303
Table A15.6:	Relationships need classification decision tree model (prison receptions)	304
Table A15.7:	Lifestyle and associates need classification decision tree model (prison receptions)	305
Table A15.8:	Drug misuse need classification decision tree model (prison receptions)	306
Table A15.9:	Alcohol misuse need classification decision tree model (prison receptions)	307
Table A15.10:	Emotional wellbeing need classification decision tree model (prison receptions)	308
Table A15.11:	Thinking and behaviour need classification decision tree model (prison receptions)	309
Table A15.12:	Attitudes need classification decision tree model (prison receptions)	310
Table A16.1:	Goodness-of-fit and accuracy of classification models by OASys post-sentence eligibility (prison receptions)	311

Figures

Figure 3.1	Quarterly trend for percentage of assessments considered valid	33
Figure 3.2	Quarterly trend for percentage of assessments with all valid section scores	33
Figure 3.3	Percentage of valid probation and prison assessments by region	34
Figure 3.4	Percentage of probation and prison assessments with all SAQ questions completed by region	34
Figure 6.1	Proven reoffending for offences not covered by OVP, by OASys, OGRS 3 and OGP decile.	101
Figure 6.2	Proven reoffending for offences covered by OVP, by OASys, OGRS 3 and OVP score deciles.	101
Figure 9.1	National risk levels of probation commencements and prison receptions	177
Figure 9.2	National criminogenic need levels of probation commencements and prison receptions	178

1. Introduction to OASys and research on OASys 2006 to 2009

The Offender Assessment System (OASys) is a national risk/need assessment tool used across probation areas and prison establishments in England and Wales. This compendium presents the research and analysis conducted by the OASys Data Evaluation and Analysis Team (O-DEAT) over a three-year period from 2006 to 2009. It begins with an introduction to OASys, describing its development, content and uses, followed by a chapter updating the literature underpinning its development and several chapters presenting full research studies that have been previously published in summary form by the Ministry of Justice.¹ It then includes findings from previously unpublished exploratory qualitative analysis of OASys content, followed by a presentation of offender profiles using OASys data, and concludes with information about other research and usage of OASys, actions following from O-DEAT research recommendations and brief indications of future research by O-DEAT.

OASys was first considered by the Home Office and Prison Service in 1998. An informative description of the origins of OASys in the Probation Service is described by Mair *et al.* (2006), including early resistance to the introduction of statistical prediction of likelihood of reoffending, which need not be reiterated in full in this compendium. In short, prior to the development of OASys, two main risk/needs assessment tools were being used with adult offenders in England and Wales: the Level of Service Inventory – Revised (LSI-R) developed by Andrews and Bonta (1995), and a tool known as Assessment, Case Management and Evaluation (ACE) developed by Roberts *et al.* (1996).² At that time some probation areas were not using any structured assessment tool in their work with offenders. In place of adopting either of the two assessments for use nationwide, the Home Office (1999) built on the existing evidence base in the literature, to devise a new assessment system as part of a national ‘What Works’ strategy (McGuire, 1995). OASys was developed with more individual questions than either of its predecessors; it contained sections for offending-related needs, risk of serious harm and sentence planning and it was amended over the course of three pilot studies between 1999 and 2001.³

OASys content

OASys is designed to fulfil the following purposes:

- to assess how likely an offender is to reoffend;
- to identify and classify offending-related needs;
- to assess risk of serious harm, risks to the individual and other risks;
- to assist with management of risk of serious harm;

1 See Howard, 2009, Moore, 2009 and Morton, 2009 in list of references.

2 For a comparison of criminogenic need areas between OASys, LSI-R and ACE, see Merrington (2004).

3 The Youth Justice Board has developed a separate risk/needs assessment tool for young offenders, known as Asset, which was introduced in April 2000 (Baker *et al.*, 2002).

- to link the assessment to the sentence plan;
- to indicate the need for further specialist assessments; and
- to measure change during the offender's sentence.

During the period in which the research reported in this compendium was undertaken, there were five main components to OASys.

1. Likelihood of reconviction and offending-related factors – the main body of OASys consists of 73 practitioner-completed questions across 12 scored sections each relating to different offending related factors: (i) offending information, (ii) analysis of offences, (iii) accommodation, (iv) education, training and employability, (v) financial management and income, (vi) relationships, (vii) lifestyle and associates, (viii) drug misuse, (ix) alcohol misuse, (x) emotional wellbeing, (xi) thinking and behaviour, and (xii) attitudes. Most sections contain both dynamic and historic data, and research has demonstrated that each area is closely related to the likelihood of reconviction (Howard, Clark and Garnham, 2006). The sections vary in how well they predict reconviction and the contribution each factor makes to the overall reconviction score is weighted accordingly.

For each factor, an offender is assessed as having a 'criminogenic need' or offending-related need if the score for the section exceeds a designated cut-off point. The sections vary in their strength of association with reconviction and are therefore weighted in their contribution to an overall likelihood of reconviction score. The final weighted scores range from 0 to 168, banded into low (0–40), medium (41–99) and high (100–168).

There are also questions at the end of each section that allow the assessor to link the offending-related factors to the risk of serious harm, risks to the individual and other risks, and offending behaviour.

2. Risk of serious harm, risks to the individual and other risks – OASys includes a comprehensive risk of serious harm assessment which draws together information from the earlier sections of OASys and has a range of questions pertaining to the risk an offender presents to others and themselves. Practitioners are required to make informed judgements as to whether various risks are low, medium, high or very high, and to complete a Risk Management Plan documenting how the risks will be controlled.
3. OASys summary sheet – a summary sheet enables practitioners to draw together key information from the assessment of the offender, providing a scoring schedule which encompasses the criminogenic needs scores, the likelihood of reconviction levels, and the risk of serious harm ratings.

4. Sentence planning – the sentence plan and risk management plan set out the activities to be undertaken to reduce the offender’s likelihood of reoffending and, where necessary, manage the offender’s risk of serious harm.
5. Self-assessment – OASys also includes a self-assessment questionnaire (SAQ) to provide a more complete picture by allowing offenders the chance to comment on how they see their lives. Questions 1 to 27 of the SAQ address a range of ‘external’ social problems and ‘internal’ individual characteristics, while question 28 asks offenders whether they think that they are likely to offend in the future. The questionnaire may identify issues that the offender has not raised in the interview or highlight differences of opinion that can usefully be discussed prior to completing the sentence plan.

OASys use

OASys is viewed as an integral part of the management of offenders, identifying offenders’ risks and needs and linking and documenting these risks and needs into individualised sentence plans and risk management plans.

A full OASys assessment should be completed in the community for all cases designated at Offender Management Tier 2 and above, with the exception of Tier 2 cases in which there is a stand-alone unpaid work requirement. In prison establishments, all offenders serving a custodial sentence of at least 12 months should be assessed as well as all young adult offenders, regardless of sentence length (National Offender Management Service, 2007a).

OASys helps the Offender Manager to collate information from a range of sources in order to assess risk of serious harm and offending-related needs. These are used pre-sentence to prepare Standard Delivery Reports for judges and magistrates and are reviewed regularly after sentence. The ‘sentence planning’ component of OASys satisfies the purpose of linking sentence plans to offenders’ individual assessments. By integrating sentence planning into the overall process of assessment, OASys assists the practitioner in clarifying the links between these two essential aspects of case management.

An electronic form of OASys is used in both the prison and probation services. It aids the process of updating and reviewing individual assessments and also includes a reporting mechanism. The establishment of area-to-area ‘connectivity’ in early 2006, enabled OASys data to be shared between the prisons and probation areas. Every prison establishment in England and Wales is able to exchange OASys assessments with every other establishment and with all probation areas, allowing practitioners to view earlier assessments for individual offenders, irrespective of where they have been completed.

The electronic assessments of OASys are collated centrally within the Ministry of Justice in the OASys Data, Evaluation and Analysis Team (O-DEAT) database. The team conducts

research and provides management information to policy colleagues, regions, local probation areas and individual prison establishments. The potential benefits of using OASys data as a source of management information were noted during early stages of development:

OASys has the capacity to provide valuable management information, some of which will be used by practitioners to develop profiles of the offenders they are working with and to evaluate overall outcomes. Information will also be of use to local managers, to enable them to identify which risk factors are most common within their local offender population and to help ensure that adequate provision has been made for them. When applied on a national basis, OASys will provide a profile of offenders and their needs, and will permit resources to be allocated effectively.

(Home Office, 2002:3-4)

To assist the regional Directors of Offender Management with resource planning, offender profiles are routinely produced through a national reporting system, summarising the offender's risk and need levels. OASys data are also used to monitor a 'Sentence Plan Outcomes' Shadow Measure for the National Offender Management Service (NOMS), assessing the 'gap' between offenders' needs and provision.

The contribution of OASys to practising established 'What Works' principles (McGuire, 1995) is clear, with OASys data helping to identify:

- i. **which offenders** should receive the available interventions – the principle requiring that resources should be matched to risk; and
- ii. **which problems** should be addressed – the principle of offending-related need, requiring that interventions should be targeted at dynamic and changeable offending-related needs.

During 2007/08, around 700,000 assessments were completed on almost 350,000 offenders.

OASys value

Bonta noted, in 2001, that internationally, many correctional organisations were still classifying offenders using a subjective clinical approach and "*taking false comfort in the ability of staff to recognise a high risk criminal when they see one*", even though the research evidence clearly showed the superiority of actuarial assessment over clinical assessment. Also, at that time, actuarial assessments were in use in some systems to assess offender risk, but few services used research-based scales to assess need (Bonta *et al.* 2001).

The importance of accurate risk and needs assessments of offenders was highlighted in both the Halliday report (Home Office, 2001) and the Carter report (Carter, 2003). The use of OASys addresses both of the issues noted by Bonta, by providing a tool to standardise a practitioner's clinical judgement in assessing offenders' risks and needs.

Despite occasionally expressed doubts in the literature about the value of standardised assessment, research in England has confirmed that

“standardised risk assessment tools do provide some level of consistency and protection against the discriminatory tendencies of some workers” and that “practitioners in the probation context have acknowledged their helpfulness in respect of consistency, quality, professionalism and credibility”.

(Lancaster and Lumb, 2006)

Yet relatively recently, it has been claimed that *“better risk assessments arose from the consistent sustained relationship built up between the offender and Probation Officer or Case Manager before OASys.”* (Fitzgibbon, 2007). Apart from proving no supporting evidence of that claim, that conclusion mistakenly views OASys as a substitute to the development of good pro-social relationships necessary for effective offender management, instead of it being a rigorous tool to aid assessment and management of offenders, for which it was designed.

It has been suggested, however, that assessments such as OASys are adopted and promoted by correctional services, without any real evidence of them enhancing prediction of reoffending or improving knowledge in the management of offenders (Horsefield, 2003). This brings us to the importance of ongoing research on OASys to address such concerns and the points highlighted by Bonta (2001) about the research needed to support the use of assessment instruments.

Importance of research

Bonta (2001) described the desirable characteristics of a risk and needs assessment tool that are important to establish, whether or not the assessment is developed in-house, or chosen from an existing instrument or adopted from use elsewhere. Table 1.1 shows the features that should be present.

Table 1.1: The desired characteristics of risk-needs classification (Bonta et. al. 2001 p 233)

Characteristic	Description
Objective	Items described with publicly observable referents; structured administration and scoring rules.
Internal reliability	Items relate to each other and the total score
Inter-rater reliability	High agreement among test administrators; items are scored the same way producing similar results
Meaningful	Information makes sense; items consistent with the research on the prediction of recidivism
Predictive validity	Scores predict relevant outcomes (e.g. recidivism, prison misconduct, parole violation)
Dynamic validity	Changes in scores predict changes in outcome
Socially unbiased	Items do not violate constitutional/charter rights (e.g. ethnicity, gender)
Generalisation	Instrument applies well to other groups and settings beyond the initial construction sample.

Research is important for establishing the different types of reliability and validity of the assessment and for evaluating, improving and modifying the assessment for optimal use in the service with the intended population.

The value of OASys data is dependent upon the assessment tool being both reliable and valid. As stated by Chitty (2004:77) in a review of the impact of correction on reoffending: *“It is essential that these [risk and need assessment] systems are sufficiently reliable, valid and sensitive measures of risk factors so that they can perform their assessment and monitoring tasks effectively.”*

For OASys to be a reliable tool, it needs to produce consistent measurements, and to be valid it needs to be measuring what it is intended to measure. The two concepts have several types as noted in the Table 1.1 from Bonta. During the second OASys pilot study, assessment of 24-month reconviction data for 757 offenders with an OASys assessment found that the total OASys score performed fairly well in terms of predictive validity, although there was variance between sections (Howard, Clark and Garnham, 2006). The focus of this compendium is on the research conducted by O-DEAT on OASys, examining different forms of reliability and validity between 2006 and 2009, described below and detailed in further chapters. There has, however, been other research on OASys prior to this period, undertaken by external researchers, summarised briefly here.

Previous external research on OASys

While OASys was in the process of being rolled out to probation and prison services nationally, Robinson (2003) used qualitative methods to explore the experiences and views of stakeholders involved in the implementation of risk assessment tools such as LSI-R (the Level of Service Inventory), adopted from the Correctional Services Canada. His research predated OASys but had some application to OASys implementation. He found that whilst assessment tools were viewed as improving consistency, fairness, accuracy and effectiveness of probation officer practice, they also raised concerns about undermining professional skills and experience. This pointed to the importance of training as essential to getting practitioners on board. Robinson also noted the importance of giving practitioners feedback on assessments along with information about what ultimately becomes of the information recorded, and access to data collated for profiling the offender population (Robinson, 2003).

In 2008, Fitzgibbon published research on risk assessments used in parole board decisions. She found assessments of risk to be higher in Probation Officers' reports compared to OASys assessments except for certain categories of offender, which showed high consistency in assessing risk of serious harm. Other findings indicated shortcomings in the quality of OASys assessments and a difficulty in using a narrative style in OASys to set the context of the offenders' lives, which parole board members would have found helpful alongside the structured information recorded in OASys.

Research commissioned by the National Association of Probation Officers (Mair *et al.* 2006), described as the first national survey of Probation Officers' views about OASys, found a variety of concerns but overall noted that users were not opposed to OASys. Users with less experience of OASys had more negative views, while those with experience of more than 12 months had more positive views about the value of OASys, its purpose and advantages. Suggestions for improvement by reducing OASys length and repetition of questions within the assessment were amongst the opinions that would benefit from consideration alongside more empirically-based evidence on the content of OASys.

A survey exploring risk of harm categorisation of offenders in the National Probation Service (Coulbeck, 2004) found a wide variation in the allocation of offenders to 'high' or 'very high' risk of serious harm. The author suggested that "*this level of variation implies the existence of radically different interpretation between Probation Areas of the proper application of the OASys risk of harm categories*", that certainly warrants further examination.

A small scale study in 2007 (Crawford) looked at OASys assessments and interviewed Probation Officers, in an exploration of what impacts on quality assessments using OASys. This research found a preference for the Risk of Harm assessment section over the scored sections in OASys which were viewed as too prescriptive, and that practitioners lacked confidence in completing OASys and were unable to make best use of their knowledge, experience and expertise. Problems identified included the limited time for assessment and the view of OASys as a form instead of a tool to aid assessment and case management. These findings supported previous research that highlighted the importance of ensuring staff understanding in the use of OASys.

While the previous research on OASys has been helpful in examining users' perspectives and the use of OASys in practice, research on OASys reliability and validity is still necessary for establishing it as an evidence-based assessment of offenders. Research and analysis on OASys content is essential for demonstrating how it is functioning as a valid and reliable assessment and where improvement or amendment to the content is needed to strengthen its value to offender management.

Research in this compendium

The second chapter in this collection begins the presentation of research on OASys conducted by the OASys Data Evaluation and Analysis Team (O-DEAT). It provides an updated review of the literature considered at the time of OASys development that underpinned OASys content. Chapter 2 adds findings from a rapid evidence assessment of research published recently to the findings from the earlier literature review and summarises the conclusions to provide the current picture from the literature on offending-related factors used in an assessment of offenders' risks and needs.

Chapter 3 presents the presentation of findings from direct analysis of OASys data, completed by O-DEAT. It reports on the rate of completion of OASys as one indication of quality and as information that is fed back to users to help improve proper completion of data recorded in OASys.

Chapter 4 reports on the internal reliability and construct validity of OASys, looking at how well the scales fit together to measure distinct offending-related factors and where possible amendments might reduce and improve OASys content and measurement of offending-related need.

Chapter 5 reports on the inter-rater reliability study, which examined consistency of assessment between different practitioners of individual offenders and provided findings on areas of OASys that might benefit from further guidance on completion following training, or possible amendment.

Chapter 6 presents findings on predictive validity of OASys and newly developed improved predictors for general reoffending (OGP) and violent reoffending (OVP), improving on the performance of existing predictors of reoffending and including factors that are subject to change.

Chapter 7 reports on the findings from analysis of OASys as a measure of change, a subject important to assessing progress with offender management, to evaluating interventions and to monitoring performance against aims to reduce reoffending.

Chapter 8 reports on the ability of the Self-Assessment Questionnaire (SAQ) to predict further offending, comparing offenders' self-assessment with practitioners' assessment of factors associated with further offending and identifying possible uses of combining the different sources of information.

Chapter 9 reports on the coverage and representativeness of OASys, examining its use with the population of offenders at which it is targeted, against the offender population processed by the courts and managed by NOMS. It identifies strengths and weaknesses in practice aimed at effective assessment in order to reduce reoffending.

While Chapters 2 to 9 present peer-reviewed research that has been separately published in summary by the Ministry of Justice, Chapters 10 and 11 present more exploratory research findings from qualitative analysis of the content of OASys found within the evidence boxes completed by assessors. These chapters identify common themes noted in evidence boxes that are not included in the content of scored items, suggesting new content for OASys for consideration.

Chapter 12 departs from the research sections in earlier chapters, to present analysis of OASys used to produce offender profiles, and reports findings from analysis of data from 2008.

Finally, Chapter 13 concludes with a list of the recommendations and actions that arose from O-DEAT research findings and discussion with policy leads and future research under consideration. It includes summary information about the use of OASys data in other research and analysis elsewhere.

2: The current evidence base for the offending-related risk factors included in OASys

Introduction

Many research studies have focused on what causes people to offend, to continue offending and desist from offending. The original development of OASys was based on the evidence available from such research on offending behaviour along with findings from research on existing risk assessment tools.

The research reported in this chapter summarises the findings from a rapid evidence assessment (REA) of the literature on the evidence underpinning the factors included in OASys. The review focused on research published since OASys was rolled out in 2002, and the findings were combined with the original evidence base used in the development of OASys.

Method

REAs are a quasi-systematic review and are often used where time and resources are not sufficient for a full systematic review which can take six months to a year to complete. The functions of an REA are to:

- search the electronic and print literature as comprehensively as possible within the constraints of a policy or practice timetable;
- collate descriptive outlines of the available evidence on a topic;
- critically appraise the evidence (including an economic appraisal);
- sift out studies of poor quality; and
- provide an overview of what the evidence is saying (Davies, 2003).

The main difference in data collation between REAs and systematic reviews is that exhaustive database searching, hand searching of journals and textbooks, or searches of 'grey' literature are not immediately undertaken (Butler *et al.*, 2004).

In this research, search terms were given to Home Office library staff, who searched the following databases for abstracts published between January 2002 and November 2007:

- National Criminal Justice Records Service (NCJRS);
- Criminal Justice Abstracts;
- MEDLINE;
- PsycINFO;
- Applied Social Sciences Index and Abstracts (ASSIA);
- Social Science Citations Index (SSCI); and
- Public Affairs Information Service (PAIS).

In total, 1,289 abstracts (including duplicates) were retrieved from the databases. Abstracts were removed from further consideration where a) they were not relevant to the research question; b) where the study was based on a sample of the population that was unlikely to translate to England and Wales; and c) if the source was a dissertation abstract and detail of the methodology was unavailable. A total of 52 papers remained which were retrieved for inclusion in the evidence assessment.

The remaining abstracts were assessed according to their methodology. To assess the quality of the evidence, a Quality Assessment Tool (QAT) was designed based upon an approach used by Butler *et al.* (2004) in an REA on the evidence for reducing gang-related violence.⁴ The QAT was used to score each study on its methodology in four areas: sample selection; bias; data collection; and data analysis. The scores for each component were then added together to provide an overall score for the study. Those studies with the highest scores were considered the least methodologically robust. Studies with a score of greater than ten were excluded from further consideration as the method was considered either too poor for the results to be reliable or was inappropriate for the project objective (19 in total).

Findings

Detailed below is a summary of the evidence base for each of the 12 sections of the OASys core assessment. This includes evidence from the original literature review that underpinned the development of OASys as well as new findings from the 2002–2007 Rapid Evidence Assessment.

Sections 1 and 2: Criminal history and offence details

Static predictor tools, including the Offender Group Reconviction Scale (OGRS) and the Sentence Planning Risk Predictor, mainly consist of offence history and current offence details. Many research studies in England and Wales and abroad have identified that criminal history is a very good predictor of future reconviction (e.g. Nuttall, 1996; Cornish and Clark, 1975; Andrews, 1983; Copas *et al.*, 1994; May, 1999).

Criminal history is often identified as the strongest factor in predicting recidivism. In a reconviction study of just under 300 male offenders, Hollin and Palmer (2006) found the criminal history subscales of the LSI-R assessment were key in predicting reconviction. Similarly, Girard and Wormith (2004) found that criminal history in the LSI-R was the best predictor of any reconviction for a violent offence, and offence severity. In a meta-analysis on predicting recidivism among mentally disordered offenders, Bonta *et al.*, (1998) found that all criminal history variables investigated and most of the personal demographic variables significantly predicted violent and general reoffending. For violent offending, the criminal history domain was the most significant predictor with personal demographic variables being the second strongest predictor.

⁴ A copy of the QAT and accompanying guidance can be found in Appendix 1.

More specifically, research has shown the following.

- Previous convictions have been found to be predictive of future offending. Lloyd *et al.* (1995) found a strong relationship between the number of court appearances and reconviction within two years. The rate of previous convictions has also been demonstrated as an important predictor in large-scale studies (e.g. Copas *et al.* 1998; Lloyd *et al.* 1995; Spicer and Glicksman, 2004).
- Individuals who have been arrested in the past are more likely to be arrested in the future (Nagin and Paternoster, 2000; Kurlychek *et al.*, 2006).
- Previous custodial sentences are also strongly linked to likelihood of reconviction, independent of the number and rate of previous convictions (e.g. Copas *et al.*, 1998; Lloyd *et al.*, 1995; Spicer and Glicksman, 2004; Spivak and Damphousse, 2006).
- The younger an offender is when first convicted, the more likely they are to be reconvicted after their current sentence. This has been demonstrated by robust meta-analyses (Andrews, 1995) and primary research.
- Male offenders have been found to be more likely to be reconvicted than female offenders (e.g. Copas *et al.*, 1998; Lloyd *et al.*, 1995; Spivak and Damphousse, 2006; Bowles and Florackis, 2007).
- The type of current offence has been found to be predictive of reconviction rates. Spicer and Glicksman (2004) found that the highest reconviction rates were associated with acquisitive crimes such as burglary and theft/handling, and the lowest rates were for sexual offences and fraud/forgery. Research across 15 American states has also shown that those convicted for using, selling or possessing illegal weapons are highly likely to reoffend (Langan and Levin, 2002). The design of OGRS reflects such findings.
- Non-compliance with orders, for instance, a previous breach of a court order, is associated with increased risk of reconviction (Taylor, 1999).
- The number of different offence types an offender has previously been convicted for has also been associated with increased risk of reconviction. Not many studies have reported this, but Hare *et al.* (1993) validated it as a factor in the Psychopathy Checklist.

Taylor (1999) concluded that although criminal history was a strong predictor of recidivism, this was because it acted as a proxy for social and behavioural problems. The research on the different social and behavioural problems associated with offending that are covered in OASys are discussed below.

Section 3: Accommodation

Accommodation is one of the seven NOMS Reducing Reoffending Pathways and getting offenders into stable accommodation is seen as the foundation for successful rehabilitation and for ensuring risk is managed efficiently:

“With regard to planning interventions to rehabilitate offenders, accommodation can provide the anchor for a previously chaotic life and act as a springboard for other crucial steps such as getting and keeping a job, and accessing health care or drug treatment.”

(National Offender Management Service, 2004)

Findings from the literature support the idea that accommodation is important to reducing the risk of reoffending with the UK prison population. The Social Exclusion Unit’s 2002 report on reducing reoffending cited figures demonstrating that prisoners released from custody had a 20% better chance of reducing their rate of reconviction compared to those who had severe accommodation problems. The following accommodation issues have been identified from meta-analysis and validation of risk assessment tools as risk factors:

- stability of accommodation, including frequent address changes (Gendreau *et al.*, 1996; Raynor *et al.*, 2000);
- being of no fixed abode or living in hostels (Raynor *et al.*, 2000); and
- living in a high-crime neighbourhood (Raynor *et al.*, 2000) or council estates with signs of drug dealing/use (Baker *et al.*, 2002).

Another potentially relevant factor was identified by Kubrin and Stewart (2006). In a sample of over 5,000 US offenders, the relationship between reoffending and the neighbourhood context that offenders live in was investigated. The results showed that whilst individual factors (e.g. criminal history and age) were strong predictors of recidivism, living in a neighbourhood characterised by poverty and socio-economic disadvantage also increased risk. It was argued that this was due to fewer resources and amenities (e.g. housing and jobs) required for successful reintegration. It is not known if these findings would similarly apply within the UK.

The literature indicates that accommodation overlaps with other risk domains, including alcohol, drugs, lifestyle and employment (Social Exclusion Unit, 2002). Webster *et al.* (2006), in a research study in Teesside, found that higher neighbourhood crime levels were associated with poverty, a lack of opportunity due to economic decline, a transient population, unsupportive social networks and lack of social capital. As a result, drug use, truancy and poor use of leisure time were also high.

Section 4: Education, Training and Employability (ETE)

The NOMS Reducing Reoffending Pathways Paper reports an association between offending, poor literacy, language and numeracy skills, and low achievement and truancy at school. Many offenders have poor experience of education and little or no experience of stable employment (Social Exclusion Unit, 2002). The literature also widely recognises education and employment as important risk factors for offending behaviour. The meta-analysis completed by Gendreau *et al.* (1996) identified education and employment as the

most important dynamic risk factor, with the highest correlations with recidivism. Similarly, Hollin and Palmer (2006) found that the education and employment subscales of the LSI-R assessment were key in predicting reconviction.

The following education and employment issues have been identified through good quality longitudinal research and validation of risk assessment tools as associated with an increased risk of reconviction:

- underachieving and dropping out of school (Farrington, 1990) or being excluded (Baker *et al.*, 2002; Raynor *et al.*, 2000);
- poor literacy and numeracy skills (Baker *et al.*, 2002);
- lack of skills and qualifications (Baker *et al.*, 2002);
- job instability (Farrington *et al.*, 1986);
- being unemployed (May, 1999; Oldfield, 1996; Raynor *et al.*, 2000); and
- poor or spasmodic employment history (Farrington, 1989; Raynor *et al.*, 2000).

The evidence on the importance of education to risk of reconviction is mixed. It is clear that a large number of offenders have poor education levels which are linked to offending in the first instance. However, improving educational skills has not been universally demonstrated to reduce subsequent offending. Harer (1995) demonstrated that that an increase in basic skills reduced the risk of reconviction for some groups, but other research has failed to demonstrate this link (e.g. Hollin and Palmer, 1995). Harper *et al.* (2005) argued that because the evidence on the impact of prison education programmes on reoffending is mixed, that level of education post-conviction is not important in predicting the likelihood of further offending. A higher level of education, however, can increase employability which is a key factor in reducing reoffending and is therefore an important rehabilitative intervention (Social Exclusion Unit, 2002).

The evidence on employment interventions to reduce offending has shown this to be successful in comparison to other interventions. Lipsey (1995) found that employment-related interventions were associated with the largest reductions in offending. Sampson and Laub also (1993) argued that offenders can modify their behaviour to be more prosocial with the help of good jobs. Unemployment will make it harder to maintain stable accommodation or to earn money legitimately, increasing the risk of reoffending.

Section 5: Financial management and income

Finance, benefit and debt is another of the NOMS seven reducing reoffending pathways:

“ensuring ex-offenders have enough lawfully-obtained money to live on is key to their rehabilitation.”

(National Offender Management Service, 2004)

Research has shown that there is a strong correlation between income and reconviction, with those who reoffend having lower incomes. However, financial circumstances are often linked to employment, and research has not clearly identified whether finances are an independent factor in predicting reconviction. May (1999) found a weak link between financial circumstances and reconviction, but Oldfield (1996) did find a correlation which existed independently of employment. The finance components of ACE and LSI-R were found to be predictive (Raynor *et al.*, 2000) and Baker *et al.* (2002) found that inadequate legitimate personal income was predictive in juvenile offending.

Nilsson (2003) found in a population of 346 Swedish offenders that an accumulation of different types of resource problems were correlated with recidivism. In particular, offenders sentenced to prison had a reduced opportunity to lead a conventional life with a legitimate income once released. The correlation with unemployment suggests that lack of money is important, but where it is independently predictive, this may be due to poor financial management, which could be linked to poor coping skills generally. Many people on low incomes do not offend and therefore factors like peer pressure, lifestyle and social-status needs, and poor financial coping skills may help to explain the link between poverty, social disadvantage and offending.

Section 6: Relationships

Children and families is one of the NOMS reducing reoffending pathways: “[they] *play a significant role in supporting an offender to make and sustain changes which reduce reoffending*”. The relationships section of OASys covers relationships with family and partners and childhood experiences. The research base is mixed on how important relationships are to predicting reoffending. The following issues have been identified as associated with recidivism from robust research and validation of risk assessment tools:

- family criminality (Andrews, 1995; Farrington, 2002; Gendreau *et al.*, 1996; Rutter *et al.*, 1998);
- poor parenting (Andrews, 1995; Farrington, 2002);
- quality of parent-child relationships (Oddone-Paolucci *et al.*, 2000; Gendreau *et al.*, 1996);
- abuse and neglect as a child (Andrews, 1995; Farrington, 2002, Oddone-Paolucci *et al.*, 2000);
- family psychopathy (Oddone-Paolucci *et al.*, 2000); and
- for juvenile offenders, not living with a birth father or mother, and not having contact with the birth father (Baker *et al.*, 2002).

OASys also includes questions on the quality of relationships with family members and with a partner. The evidence for the predictive relevance of these factors is mixed. Laub *et al.* (1998) found that persistent offenders ceased criminal activity after marrying or

having children. Gendreau *et al.* (1996) found evidence for family discord and conflict as an important risk factor. However, Raynor *et al.* (2000) only found weak evidence for their importance in the validation of LSI-R and ACE. King *et al.* (2007) found that marriage suppresses offending for males, but not for females. Wright and Wright (1992) reported that marriage and family were not associated with offending among adult offenders. They also found that offenders were as likely to be in a significant relationship as non-offenders, but they were more likely to divorce or separate, to fail to get along with their spouses and to be involved in violent relationships.

Wright and Wright (1992) also found that maintaining active relationships with the family while incarcerated and after release were associated with less subsequent reoffending. Similarly, the Resettlement Surveys Reoffending Analysis (RSRA) study (May, 1999) which surveyed prisoners in 2001, 2003, and 2004 (see also Niven and Olagundoye, 2002; Niven and Stewart, 2005) found that receiving family visits reduced the likelihood of reoffending. Quality of family relationships may therefore be the important factor. Offenders who were visited were also more likely to have accommodation and employment or training arranged on release.

A history of domestic violence also appears to be indicative of future domestic violence incidences. Hester and Westmarland (2006) analysed data on 692 perpetrators of domestic violence in the north east of England and found that 50% were involved in further domestic violence incidents within a three-year follow-up period. Eighteen per cent of offenders had reoffended against a different partner. It is likely that not all incidents of domestic violence were reported and therefore the true reoffending rate may have been even higher. This research has not been replicated across England and Wales, and it is not known to what extent the sample of perpetrators was representative of the offending population.

Section 7: Lifestyle and associates

The lifestyle and associates section covers how offenders spend their leisure time, who they spend it with (excluding family), and risk-taking behaviour. This section has been shown to have strong associations with other domains, e.g. drug abuse and attitudes (Raynor *et al.*, 2000).

Sutherland and Cressey's (1970) differential association theory argued that offenders who spend an excessive amount of time with other offenders and less time with non-offenders are encouraged to offend. This is because it provides an opportunity for anti-social modelling to occur and anti-social attitudes to be influenced. The evidence base from meta-analytic studies and validity studies of risk assessments has identified the following factors as increasing likelihood of recidivism:

- peer group mainly consisting of criminal and or anti-social associates (Andrews, 1995; Gendreau *et al.*, 1996; Baker *et al.*, 2002; Goggin *et al.*, 1998);

- peer pressure (May, 1999) and people who offended with peers (Oldfield, 1996);
- poor participation in organised activity (Raynor *et al.*, 2000, Mills *et al.*, 2004), and for juveniles non-constructive use of time (Baker *et al.*, 2002). In the U.S. a meta-analysis found that ineffective use of leisure time was one of the best predictors of juvenile recidivism (Cottle *et al.*, 2001);
- risk-taking activities (Andrews, 1995; Baker *et al.*, 2002); and
- boredom (Raynor *et al.*, 2000).

The research evidence shows that in general criminal or anti-social associates are a stronger predictor of recidivism than activities and hobbies. For instance, in validation of the LSI-R assessment, having pro-criminal associates was more influential than not having pro-criminal ones (Raynor *et al.*, 2000). In other words, associates were a stronger risk than a protective factor. Additionally, companions rather than leisure activities seemed more important. Validation of the Canadian Offender Intake Assessment (OIA) found that within the associates/social interaction domain 'associates with substance abusers', 'has many criminal acquaintances', and 'has mostly criminal friends' were strongly related to reconviction for offenders. 'Unattached to community groups', and 'resides in a criminogenic neighbourhood' only demonstrated a moderate relationship to reconviction. In terms of the leisure component, 'no hobbies' and 'does not participate in organised activities' were found to be moderate predictors of readmission (Brown and Motiuk, 2005).

There is evidence that criminal associates may not be an important risk factor for all groups of offenders. Sigmour (2004) studied a sample of long-term incarcerated offenders and found that recidivists had higher mean LSI-R scores than non-recidivists on the leisure/recreation subcomponent, but that the companions subcomponents did not significantly distinguish recidivists from non-recidivists. The sample of long-term incarcerated offenders included a large proportion of serious violent offenders with a relatively high base rate of recidivism, which may explain the lack of distinction.

More recently, research on gang membership has been linked to recidivism. Huebner, Varano and Bynum (2007) found that men who were involved in a gang or who were drug dependent before entering prison had higher reconviction rates and reoffended more quickly than men who did not report involvement in gangs or drug use. Some researchers have argued that gangs can provide the motivation and opportunity for deviance, particularly amongst younger offenders. Ebensen and Huizinga (1993) found from the Denver Youth Survey that, when controlling for association with non-gang-delinquent peers and prior delinquency, gang members self-reported two to three times more delinquency than non-gang members. Research has yet to be conducted on a large sample in the UK which would provide more conclusive findings on the importance of gang membership.

Section 8: Drug misuse

Drugs and alcohol are another of the seven NOMS reducing reoffending pathways: “*at any one time about one-third of all problematic drug users in England and Wales are in the care of NOMS*”. The majority of research has found that drug use is predictive of recidivism. In Britain, drug use has been shown to be a strong predictor of recidivism for adult offenders (e.g. Raynor *et al.*, 2000; Spohn and Holleran, 2002), and amongst juvenile offenders, both cannabis and tobacco use have been linked to reconviction (Baker *et al.*, 2002). Cottle *et al.* (2001) found drugs to be a moderately strong predictor for juveniles in the United States.

Not all research has found drug use to be an important predictor. Andrews (1995) did not list drug use as predictive. Oldfield (1996) found that drug use was only predictive of reconviction for certain offence types, namely burglary and motor theft. It may be that drug use predicts further offending behaviour where it leads to crime to finance the habit, (Wright and Decker, 1997), or it may be that drugs act to disinhibit offending behaviour (Rutter *et al.*, 1998).

Alternatively, drug use may also increase the chances of recidivism because of the consequences it can have for sustaining employment (Laub and Sampson, 2003). Kaestner (1993) found an association between drug use and a reduction in the number of hours worked per week and Zarkin *et al.* (1998) found an association with more absences from work.

Drug use also impacts on family relationships and developing a prosocial sense of self (Laub and Sampson, 2003) which may influence offending behaviour. Schroeder, Giordano and Cernkovich (2007) argued that drug use becomes embedded in social network affiliations which make desistance more difficult. They suggested that addiction to substances required access to drugs, which in turn involved maintaining relationships with criminal associates. Their longitudinal research showed positive associations between drug use and peer group deviance and partner criminality, which in turn were significantly associated with criminal offending.

Many assessment tools for predicting recidivism give a higher weighting to use of ‘harder’ drugs such as heroin and crack, than cannabis. However, some research has indicated that there is a relationship between frequent cannabis use by juveniles and crime. Research from Australia and the United States has suggested that those who begin using cannabis at an early age and who use it frequently are at risk of subsequently using cocaine or heroin (Ellickson, Hays and Bell, 1992; Kandel and Yamaguchi, 1993). Salmelainen (1995) interviewed Australian juvenile theft offenders and found that there was a relationship between self-reported use of cannabis and frequency of offending for juveniles convicted for motor vehicle theft and breaking and entering. There was also a relationship between heavy cannabis use and reporting the need to finance drug use as the main reason for committing the crime. Peersen *et al.* (2004) found in a sample of Icelandic prisoners that the frequency of cannabis use distinguished between those who reoffended and those who desisted over five years following release.

There is mixed evidence as to whether cannabis increases the risk of violent behaviour. Taylor and Hulsizer (1998) concluded that cannabis was not related to violence except in exceptional circumstances. In comparison, Sussman *et al.*, (1996) concluded that cannabis use did lead to an increased risk of violence. Moore and Stuart (2005) concluded from a review of available data that there was evidence for a relationship between cannabis use and interpersonal violence, though this was not supported for some subgroups. However, these conclusions were based on the available research which had methodological weaknesses, for instance insufficient controls and a lack of longitudinal data.

Section 9: Alcohol misuse

Alcohol is covered in the NOMS reducing reoffending pathway: drugs and alcohol. Earlier research had shown that alcohol misuse was associated with offending and in particular violent offending. However, recent studies on alcohol abuse have suggested that whilst it is linked to risk of reoffending, it is not as strong a predictor as drug use (Cooke, 1989).

Raynor *et al.* (2000) found no predictive relationship between alcohol use and reconviction in their validation of LSI-R and ACE. May (1999) only found a link for offenders sentenced to community orders in one out of six probation areas studied. Oldfield (1996) found a negative association (i.e. those with alcohol problems were less likely to be reconvicted). Oldfield attributed this to alcohol problems being predominately found amongst those committing violent offences, for whom the reconviction rates are lower than average. Alcohol problems were also more likely among older offenders, who also have lower reconviction rates than younger offenders.

Rutter *et al.* (1998) suggested that alcohol could act as a disinhibitor, therefore increasing the likelihood of violence, disorderly conduct and driving offences. Alcohol use may also form part of an anti-social lifestyle (though less so than drugs) which includes offending. Sampson and Laub's (2003) life-course theory emphasises the negative impact of alcohol use on marital and employment bonds, which would act as protective factors against offending, if in place.

Section 10: Emotional wellbeing

Some research has shown emotional problems to be moderately correlated with recidivism. Andrews (1995) listed psychopathy, weak socialisation and egocentrism as major risk factors and personal distress (including low self-esteem, anxiety and depression) as minor factors. Other research has found a lack of ability to cope with stress, depression and mental health problems to be correlated with recidivism (Gendreau *et al.*, 1996; Zamble and Quinsey, 1999). Girard and Wormith (2004) found with a sample of Canadian offenders that the Mental Health Issues subscale of the LSI-R was significantly correlated with recidivism, albeit a weaker correlation than other subscales. However, sample sizes were too small to be reliable. Research on predicting juvenile reconviction with Asset in Britain (Baker *et al.*, 2002) found that mental health problems were predictive of reconviction.

Other research has suggested that emotional wellbeing is not related to reoffending. Oldfield (1996) found no positive relationship between mental health problems and reconviction in a sample of probationers. Instead, there was some indication that people with mental health problems were less likely to indulge in drug abuse or to offend with peers. May (1999) did not find any correlation between mental health problems and reconviction, and Raynor *et al.* (2000) likewise found no such relationship in the validation of LSI-R and ACE. Identifying a link between emotional wellbeing and offending may be difficult due to the range of problem severity, its variability and difficulty in measuring over time.

It may be that for certain extreme mental health problems, emotional wellbeing is an important risk predictor. Hart, Kropp and Hare (1988) conducted a study in which 231 offenders were administered the Hare's Psychopathy Checklist-Revised (PCL-R) prior to release from jail. Within three years, 75% of those categorised as non-psychopathic were still in the community, in comparison to only 20% of the designated psychopaths. Hollin (2002) found that the PCL-R predicted reconviction; however, the items in the checklist spread across several domains of OASys and not just emotional wellbeing. The items include habitual lying, being manipulative, parasitic, callous, impulsive, irresponsible, needing stimulation, easily bored, lacking remorse, having no emotional depth, poor behaviour control, and criminal versatility. This research needs replicating with larger samples for the findings to be reliable.

In a sample of offenders discharged from a medium-secure unit in the UK, clinical diagnosis was not found to be predictive of reoffending (Phillips *et al.*, 2005). Instead, the number of previous offences was the strongest predictor, although the number of days hospitalised was also a predictor.

There is some evidence that emotional wellbeing may be predictive of only certain types of recidivism. Bonta's (1998) meta-analysis found that whilst most clinical variables were unrelated or inversely related to general recidivism, more clinical variables were significant for predicting violent recidivism.

In summary, the evidence is mixed and emotional wellbeing in general does not appear to be a strong predictor of reconviction. However, particular aspects such as psychopathy for predicting violent reconviction may be important.

Section 11: Thinking and behaviour

Cognitive and behavioural deficits were identified as important to offending as part of the 'What Works' movement (McGuire, 1995). Andrews (1995) suggested that personality factors including weak problem-solving and self-regulation skills were a major risk factor.

Thinking and behavioural factors that the evidence base has identified as important in predicting recidivism are as follows:

- impulsivity (Raynor *et al.*, 2000; Hare *et al.*, 1993; Baker *et al.*, 2002; Rutter *et al.*, 1998);
- boredom and a need for excitement (Raynor *et al.*, 2000; Baker *et al.*, 2002);
- risk-taking (Raynor *et al.*, 2000);
- lack of self-control (Raynor *et al.*, 2000) or self-regulation (Andrews, 1995); and
- not taking responsibility for behaviour (Raynor *et al.*, 2000).

There is mixed evidence on the links between aggression and temper control and recidivism. Novaco (1997) looked at violent offenders and found a strong correlation between expressive anger (as opposed to instrumental). Amongst juveniles, aggression and poor temper control have been shown to be predictive of delinquency (Rutter *et al.*, 1998) and recidivism (Baker *et al.*, 2002) However, research on ACE has not found a relationship between aggression/temper control for the general adult offending population. This could be because of the relative rarity of violent offences amongst adults.

There is also mixed evidence on the relationship between problem-solving skills and recidivism. Ross and Fabiano (1985) found that amongst Canadian offenders, persistent offenders often lacked problem-solving skills when compared to non-persistent offenders and non-offenders. Robinson (1995) showed that programmes to improve problem-solving skills amongst Canadian offenders did reduce recidivism. However, in England and Wales the research is less conclusive. LSI-R did not include problem-solving skills as a measure and ACE research did not find reasoning/thinking skills predictive (Raynor *et al.*, 2000). Amongst juveniles, Asset research has found that poor understanding of consequences was a risk factor (Baker *et al.*, 2002).

In summary, the importance of thinking and behavioural factors to predict recidivism is not yet clear. This is in part because it is a relatively new construct and therefore it is not included in all assessment tools and the evidence base does not contain many studies with large enough sample sizes or follow-up periods to clearly identify the relationship.

Section 12: Attitudes

Evidence from the literature has shown that attitudes can be an important risk factor in reoffending. Andrews (1995) placed anti-social and pro-criminal attitudes, values and beliefs as highly important to risk of offending. This has been supported in meta-analyses of criminal-behaviour predictors which have found that anti-social attitudes have a strong correlation with criminal conduct (e.g. Gendreau *et al.*, 1996).

Reconviction studies of offender risk assessment tools have also found strong correlations between attitudes and recidivism. Research on ACE found that attitude and motivation domains were strong predictors of reconviction (Raynor *et al.*, 2000). These included: pro-criminal attitudes; thinking the benefits of crime outweigh the costs; regarding future offending as inevitable; being motivated to avoid offending and to deal with the problems underlying it; and being concerned about the effects of their offending on 'close' people. LSI-R research also found that attitudes were a strong risk factor (Raynor *et al.*, 2000), and research on juveniles using Asset found attitudes and motivation were strongly correlated with reconviction (Baker *et al.*, 2002). Mills *et al.* (2004) showed that the addition of the Measure of Criminal Attitudes and Associates (MCAA) significantly improved the prediction of violent recidivism over purely actuarial measures.

Pro-criminal attitudes, even those measured by self-report questionnaires, have been shown to be predictive of reconviction, parole violation and general misconduct (Walters, 1992; Sigmourd, 1997). Palmer and Hollin (2004) found one scale of the Psychological Inventory of Criminal Thinking Styles (PICTS) developed by Walters (1996) to be a powerful predictor of reconviction for 174 male prisoners. A high score on the 'superoptimism' scale indicating a belief that the negative consequences of criminal behaviour can be avoided indefinitely had a predictive effect beyond that of age and criminal history. However, the research sample was small and needs replicating on a wider range of offenders.

Not all attitudinal measures have been found to be predictive. Neither ACE nor LSI-R found that attitudes to supervision were predictive, a factor included in the original version of OASys (Raynor *et al.*, 2000).

In summary, research has clearly shown that the strongest attitudinal predictors of offending are pro-criminal attitudes and little motivation to change.

Gender differences

Research has found that male and female offenders exhibit different types and patterns of offending, as well as having different personal circumstances (for a general review see Heidensohn, 2002). Despite this, much of the research informing risk/need assessments has been carried out with male offenders, with an underlying assumption that the criminogenic needs of women offenders are similar to those of male offenders.

Blanchette (2002) questioned this assumption and suggested that there may be two types of criminogenic need: those common to men and women; and women-specific criminogenic needs. More specifically, Gelsthorpe (1999) argued that finances, accommodation, education, employment, and substance use are relevant needs common to both males and females, whereas relationship problems, mental health issues and childcare problems are characteristic more often of female offenders. Palmer and Hollin (2007) agreed that some

needs may be gender specific. They found that female prisoners assessed with LSI-R had higher levels of need relating to family and marital issues, alcohol and drugs, and emotional and personal problems. With a sample of Canadian offenders assessed with OIA, Motiuk and Blanchette (2000) found that women were more likely to have encountered difficulties with associates and significant others, specifically relational (marital, maternal, sibling and other relative) difficulties. Male offenders were more likely to have experienced problems relating to substance abuse and attitudes. American research has shown that women offenders are also more likely to be economically disadvantaged, and be involved in the criminal justice system for drug-related crimes (e.g. Owen, 1998; BJS, 1999; Covington, 2003).

Despite most offender tools being developed from theories of crime and delinquency amongst males, most of the female-specific criminogenic needs outlined above are included in the major offender assessment tools (e.g. ACE, LSI-R, OASys). This has been used to defend the use of gender-neutral tools.

The research addressing how accurate these assessment tools are at predicting recidivism for female offenders is mixed. Some risk tools have been reported to generalise poorly to female offenders (e.g. Bonta, Pang and Wallace-Capretta, 1995). However, the LSI-R has been reported to be better. Coulson *et al.* (1996) found that recidivism for women discharged from a medium- security facility in Canada was significantly correlated with LSI scores. Raynor *et al.* (2000) found that in England and Wales, when the LSI-R was applied to women offenders only, it performs in a similar manner as when applied to a sample of both male and female offenders. Palmer and Hollin (2007) reported that total LSI-R score was predictive of reconviction and time to reconviction for a sample of 150 English female offenders. Lowenkamp, Holsinger and Latessa (2001) also reported positive findings, albeit with a small sample size (n approx 100).

Another research study found that LSI-R may be accurate in predicting recidivism for some, but not all types of female offenders. Reisig *et al.* (2006) interviewed 235 female offenders in the US and classified them based on Daly's (1994) pathways to crime framework. The classification included four gendered pathways (street women; drug-connected; battered; and harmed and harming) as well as two gender neutral pathways (economically motivated and unclassified). An LSI-R was completed for each offender and the sample was followed for 18 months to measure recidivism. The research showed that the LSI-R was accurate for predicting recidivism for women who followed the economically motivated pathway into criminality. However, for women following gendered pathways into crime, the tool misclassified marginalised women. The sample size in this study is small and would need replicating for confirmation.

Other factors associated with recidivism

Some research has identified factors associated with a risk of reconviction that are not included in the original version of OASys. These are summarised below:

- social class and race (Gendreau *et al.*, 1996; Cottle *et al.*, 2001; Rutter *et al.*, 1998);
- intelligence/IQ (Andrews, 1995; Cottle *et al.*, 2001; Rutter *et al.*, 1998) ;
- other genetic and biological factors such as serotonin levels (Rutter *et al.*, 1998);
- conduct whilst in prison: Huebner *et al.* (2007) found that among serious young offenders, 45% of a group with moderate to high misconduct whilst in prison were reconvicted. In contrast, 38% of the low-misconduct group were reconvicted, and 24% of the zero- misconduct group were reconvicted.

Different offender populations

Hollin and Palmer (2006) suggested that the predictive value of LSI-R may vary according to the population of offenders. Girard and Wormith (2004) found different subscales of the LSI-R to be predictive for different groups – specifically offenders sentenced to prison and community sentences. Hollin and Palmer (2006) also suggested that the different needs might be predictive of different offence types, though further research in this area was required.

Conclusion

A rapid evidence assessment of recent research, combined with older research evidence underlying the factors included in OASys assessment of risk and need, generally supports the content of OASys, but points to possible new factors for consideration, and highlights the need to test the factors included with the offender population to which OASys is applied. Direct primary research on the variables included in all 12 OASys sections is important to establish the reliability and validity of OASys as a tool for assessing and managing offenders. The remainder of this compendium presents the research completed on different types of reliability and validity of OASys, along with findings from more explorative qualitative study of OASys textual information, and data output that can as a result be confidently produced for management information, performance measurement, and research on offenders.

3. Measurement of OASys rates of completion

Introduction

The aim of this research was to measure the completion rates of OASys assessments. For OASys to be an effective tool for the assessment and management of offenders the rate of data completion and the quality of data recorded are of utmost importance. Poor data completion impacts on (i) management of the individual offender determined by the assessment information; (ii) data analysis, management reports and research informed by the data within the assessments; and (iii) the benefits to offender management afforded by good use of OASys. Confidence in all three of these will be higher when standards of completion are high and data quality is verifiably accurate.

A very high completion rate of OASys data will ensure that assessments are comprehensive and thorough enough for practitioners and managers to make sound and defensible decisions about offenders and how best to reduce reoffending and protect the public. High levels of completion will also help OASys realise its benefit as an essential form of management information at both local and national levels. Levels of completion also affect the findings of research on the reliability and validity of OASys as an effective assessment tool, and research on offenders that draws on OASys data.

Completion rates are only one aspect of OASys quality. This analysis tells us how much of the assessment has been done but does not provide information on the accuracy of the information recorded. Other quality assessment procedures that exist include the Quality Management Plan (QMP), Her Majesty's Chief Inspectorate of Probation and the Public Protection and Licence Release Unit (PPLRU). These are currently being considered in a review of the QMP in order to set up a performance measure of OASys quality.

Stringent targets for data completion were suggested by the OASys business team in 2005.

- The required completion rate for items within the core assessment, the Risk of Serious Harm assessment and the initial sentence plan was set at 90%.
- The required completion rate for any of the Self Assessment Questionnaire (SAQ) questions was set at 50%.⁵

Method

Management information reports on completion rates of OASys data have been produced and shared with all probation areas quarterly and with prisons every six months since April 2005. In response to receipt of this information, over 30 probation areas and prison establishments have requested further analysis down to the level of individual assessors, in

⁵ The requirement for the SAQ is lower because its completion is not mandatory and in some cases may have been completed on paper only and therefore not included in the OASys database.

order to understand and improve their completion rates of OASys assessments.

This chapter reports on the findings between April 2005 and March 2007 and specifically aims to answer the following research questions:

1. What are the data completion rates for each OASys section?
2. Are there differences in data completion between probation and prisons?
3. Has OASys data completion improved over time?
4. Are there any regional differences in data completion?

All probation assessments where the purpose was “start of community sentence” and all prison assessments where the purpose was “start of custody”, completed between April 2005 and March 2007, were selected for inclusion in the analysis. The completion rate was of primary interest at the start of an offender’s sentence when initial sentence plans were agreed. As OASys is **not required** to be completed in the community for Tier 1 offenders or Tier 2 cases in which there is a stand-alone unpaid work requirement, probation assessments were removed from the analysis for such offenders. Similarly, prison assessments were excluded for custodial sentences shorter than 12 months for offenders aged 21 and over, for whom OASys assessments are not required. In total, 198,994 probation and 25,240 prison assessments were included in the analysis.

The OASys items and components measured are listed below in Table 3.1.

In all cases, for an OASys item/component to be considered complete the codes recorded must have been a suitable response, that is, a response recognised by the OASys guidance manual –dummy codes were treated as missing data.

The percentage of assessments that had a valid response for each item in the above table was measured. Comparisons were presented between the probation and prison service, between financial years, regions and offender demographic groups (gender, age and ethnicity). Completion rates were considered ‘excellent’ at above 90%, ‘good’ when between 75 and 90% and ‘poor’ below 75%.

Table 3.1: OASys items measured for data completion

OASys item	Explanation
Overall assessment	
Valid assessment	The % of assessments that meet the minimum validity criteria of completion for inclusion in any management information or other research using OASys data. The minimum criteria are: 1. Each of the scored sections (1 to 12) within the core assessment have at least four-fifths of the scored items completed 2. In the RoSH sections, the screening must have been completed, the decision whether to complete a full risk analysis should have been consistent with the information provided, and the four ratings of RoSH in the community must have been recorded in those cases where a full analysis was required. Core assessment
Basic demographics	The % of assessments where age, gender and ethnicity are all recorded
PNC number	The % of assessments with a valid PNC number entered. Invalid or official dummy PNC numbers (33/993399H) are counted as missing
Offence code	The % of assessments with a valid offence code entered
Sentence code	The % of assessments with a valid sentence code entered
All section links to risk	The % of assessments with either yes or no recorded for links to risk of serious harm in all sections 2 to 12
All section links to offending behaviour	The % of assessments with either yes or no recorded for links to offending behaviour in all sections 3 to 12
All scored questions	The % of assessments with all 73 OASys scored questions completed
All section scores (valid only)	The % of assessments with enough scored questions completed to calculate scores in all of sections 1 to 12 (at least 80% of scored items in each section must be completed)
Risk of Serious Harm	
Any screening sections (R1.2 to R5.1)	The % of assessments with at least one RoSH screening questions R1.2 to R5.1 completed
All screening sections (R1.2 to R5.1)	The % of assessments with all RoSH screening questions R1.2 to R5.1 completed. No data can be missing i.e. for R1.2 and R1.3 "none of the above apply" must be ticked if applicable.
Community risks	The % of assessments where if the RoSH screening indicated that the full analysis was necessary, all four risks to the community (children, public, known adult, staff) are completed.
SAQ	
Any SAQ questions (1-28)	The % of assessments with at least one SAQ question 1-28 completed
All SAQ questions (1-28)	The % of assessments with all SAQ questions 1-28 completed
Sentence plan	
Any primary sentence plan items	The % of assessments where at least one item has been recorded in the sentence plan
Primary need	The % of assessments where at least one need has been recorded in the sentence plan
Primary objective	The % of assessments where at least one objective has been recorded in the sentence plan
Primary intervention	The % of assessments where at least one intervention has been recorded in the sentence plan
Arrangements for supervision	The % of assessments where the arrangements for supervision have been recorded (not applicable for prison assessments)

Limitations

1. No measure of the accuracy of the data recorded for each assessment is considered in this analysis; thus the conclusions only partially address issues about the quality of OASys data completion.
2. Only assessments completed at the start of a community or custodial sentence have been measured and findings may not apply to completion rates for subsequent assessments during a sentence.
3. Some OASys items have a default response which would be measured as a valid response. For instance questions R1.2 to R4 in the RoSH screening are defaulted to 'no' (unless overridden); therefore in these instances the true completion rate may be lower than the analysis shows.
4. Findings from further exploration of data by probation areas or prison establishments that requested specific local-level information were not obtained, therefore it is not possible to give feedback on the reasons for non- or low- completion rates.

Results

What are the data completion rates for each OASys section?

Table 3.2 shows the number of probation and prison assessments included in the analysis for each financial year. The completion rates for each OASys component are also presented. Shaded cells indicate the components that met required targets.

For probation assessments, three out of 19 OASys items met the completion targets in 2005/06, which improved to ten out of 19 in 2006/07. For prison assessments, nine out of 18 OASys items met completion targets in 2005/06, which improved to 14 out of 18 targets in 2006/07.

The percentage of assessments that met the minimum completion levels for inclusion in any O-DEAT management information or other research was below target for both services for both years. The lowest percentage was 71% for prison assessments in 2005/06 and the highest percentage was 86% for probation assessments in 2006/07.

For the core OASys assessment, the majority of items measured either met or were close to meeting the target completion rate of 90%. A notable exception was the completion rate in probation assessments for sentence code. This was 76% in 2005/06 and 72% in 2006/07.

Table 3.2: Completion rates for probation and prison assessments per year

OASys item		Probation		Prison	
		2005/06	2006/07	2005/06	2006/07
Overall	Number of assessments completed	59,858	49,136	14,036	11,204
	Valid assessments*	77%	86%	71%	83%
Core Assessment	Basic demographics	85%	84%	92%	96%
	PNC	91%	87%	96%	99%
	Offence code	96%	98%	93%	94%
	Sentence code	76%	72%	97%	100%
	All section links to risk	81%	87%	94%	94%
	All section links to offending behaviour	87%	94%	93%	94%
	All scored questions	79%	87%	69%	77%
	All section scores (valid only)	85%	93%	79%	87%
RoSH	Any screening sections	91%	98%	94%	94%
	All screening sections	75%	79%	86%	88%
	Community risks (where full analysis necessary)	87%	93%	87%	94%
SAQ	Any SAQ questions	42%	51%	84%	84%
	All SAQ questions	21%	28%	61%	64%
Sentence Plan	Any sentence plan sections	81%	94%	89%	91%
	Primary need	80%	94%	88%	91%
	Primary objective	80%	94%	89%	91%
	Primary intervention	80%	94%	88%	91%
	Arrangements for supervision	70%	84%	N/A	N/A

* For an assessment to be valid it must meet minimum validity criteria which are: 1) each of the scored sections (1 to 12) within the core assessment have at least four-fifths of the scored items completed; and 2) in the RoSH sections, the screening must have been completed, the decision whether to complete a full risk analysis should have been consistent with the information provided, and the four ratings of RoSH in the community must have been recorded in those cases where a full analysis was required.

The completion rate for all 73 scored OASys questions in prison assessments was below target at 69% in 2005/06 and 77% in 2006/07. Table 3.3 shows the percentage of prison assessments in each financial year that did not have a valid section score for each of sections one to 12. The section least likely to have enough scored questions completed in both financial years was section six, relationships (16% in 2005/06 and 10% in 2006/07).

Table 3.4 shows the ten scored questions with the poorest completion for each service for each financial year. For both probation and prisons questions from section 6 (relationships) were problematic with at least four out of the six scored questions for the section in the top ten poorest completed questions. In particular, criminal records for partners and close family members and experience of childhood had lower completion rates.

Table 3.3: Percentage of prison assessments with insufficient scored questions completed for each section of the core OASys assessment

OASys Section	Prison 2005/06	Prison 2006/07
Section 1 & 2 combined (offence history & analysis)	8%	7%
Section 3 (accommodation)	13%	8%
Section 4 (ETE)	12%	8%
Section 5 (financial management)	13%	8%
Section 6 (Relationships)	16%	10%
Section 7 (lifestyle & associates)	12%	8%
Section 8 (drug misuse)	12%	7%
Section 9 (alcohol misuse)	8%	7%
Section 10 (emotional wellbeing)	11%	7%
Section 11 (thinking and behaviour)	6%	6%
Section 12 (attitudes)	11%	7%

Table 3.4: The 10 scored OASys questions with poorest completion rates for each service in each financial year

Probation		Prison	
2005/06	2006/07	2005/06	2006/07
S6Q2 Close family member has a criminal record	S6Q2 Close family member has a criminal record	S6Q5 Current partner has a criminal record	S6Q2 Close family member has a criminal record
S6Q5 Current partner has a criminal record	S6Q5 Current partner has a criminal record	S6Q2 Close family member has a criminal record	S6Q5 Current partner has a criminal record
S6Q3 Experience of childhood	S2Q14 Current offences an established pattern of similar offending	S6Q6 Previous experience of close relationships	S8Q5 Level of use of main drug
S6Q6 Previous experience of close relationships	S6Q3 Experience of childhood	S5Q5 Over reliance on others for financial support	S6Q6 Previous experience of close relationships
S6Q4 Current relationship with partner	S6Q6 Previous experience of close relationships	S6Q3 Experience of childhood	S1Q12 Number of different categories of conviction
S5Q5 Over reliance on others for financial support	S6Q4 Current relationship with partner	S3Q6 Suitability of location of accommodation	S6Q3 Experience of childhood
S5Q4 Illegal earnings are a source of income	S6Q1 Current relationships with close family members	S3Q4 Suitability of accommodation	S3Q6 Suitability of location of accommodation
S6Q1 Current relationships with close family members	S4Q6 School attendance	S5Q4 Illegal earnings are a source of income	S5Q5 Over reliance on others for financial support
S7Q1 Community integration	S7Q1 Community integration	S7Q1 Community integration	S3Q4 Suitability of accommodation
S4Q6 School attendance	S7Q4 Manipulative/ predatory lifestyle	S7Q4 Manipulative/ predatory lifestyle	S5Q4 Illegal earnings are a source of income

Prison assessments also had poor completion rates for two out of the four scored questions in section 3 (accommodation), specifically related to the suitability and location of offenders' accommodation.

There were some data completed in the RoSH screening component for over 90% of probation and prison assessments in each financial year. However, the entirety of the RoSH screening was completed less frequently and for both services did not achieve the 90% target. Completion of community risks was higher and in the year 2006/07 was over 90% for both services.

Completion rates for the SAQ in prison assessments exceeded the 50% target in both years. In comparison, 28% of probation assessments had all SAQ questions completed in 2006/07. True completion of the SAQ may be higher as there may be some cases where the SAQ was completed on paper and not transferred onto the electronic system.

Sentence plan items were completed in over 90% of both probation and prison assessments in the year 2006/07 with the exception of arrangements for supervision which was completed in 84% of probation assessments.

Are there differences in data completion between the probation and prison services?

There are some items which are more frequently completed by the probation service. Probation had a higher percentage of assessments than prison where all the scored OASys questions were completed (by ten percentage points). A greater percentage of probation assessments had all section scores sufficiently completed (by six percentage points).

One of the largest discrepancies between probation and prisons was completion of the sentence type. The completion rate for prison assessments was over 20 percentage points higher than probation assessments which did not meet the 90% target. The completion rates for offender basic demographics and PNC numbers were also more than ten percentage points higher for prison assessments.

Similar levels of completion were observed for each service for the RoSH screening and sentence- plan components. There was a large discrepancy in completion of the SAQ. Prisons had over twice the proportion of assessments with all SAQ questions completed. Probation had a very low proportion of assessments with all questions completed (28%).

Has OASys data completion improved over time?

Table 3.5 shows the change in the percentage of OASys items completed between 2005/06 and 2006/07 for both probation and prisons. Increases were seen for the majority of items demonstrating that OASys completion rates have improved. In total, 16 out of 19 probation items measured showed improvement and 15 out of 18 prison items showed improvement.

Table 3.5: Change in completion rates between 2005/06 and 2006/07 for each service

OASys item		Change in completion from 05/06 to 06/07	
		Probation	Prisons
Overall	Valid assessments	9%	11%
Core Assessment	Basic demographics	-1%	4%
	PNC	-4%	3%
	Offence code	3%	1%
	Sentence code	-4%	3%
	All section links to risk	6%	0%
	All section links to offending behaviour	7%	1%
	All scored questions	8%	8%
	All section scores (valid only)	8%	8%
RoSH	Any screening sections	7%	0%
	All screening sections	4%	2%
	Community risks (where full analysis necessary)	6%	7%
SAQ	Any SAQ questions	9%	0%
	All SAQ questions	8%	3%
Sentence Plan	Any sentence plan sections	14%	3%
	Primary need	14%	2%
	Primary objective	14%	2%
	Primary intervention)	14%	2%
	Arrangements for supervision	14%	N/A

Both probation and prisons demonstrated an increase in the percentage of assessments that met the minimum completion criteria for inclusion in management information and research (by nine percentage points for probation and 11 percentage points for prisons). Figure 3.1 shows the quarterly trend in this improvement. There was a sharp improvement for probation assessments between quarter one and quarter two in 2005/06 which has remained reasonably constant. For prison assessments there was gradual improvement in completion across 2005/06 which then reached a plateau in 2006/07 of over 80%.

The greatest improvement in prison assessments was all 73 scored questions and all section scores in the core assessment which both improved by eight percentage points. Figure 3.2 shows the trend in completion rates for all valid section scores. Both services saw an improvement in completion rates in the year 2005/06 which was maintained in 2006/07.

Figure 3.1: Quarterly trend for percentage of assessments considered valid

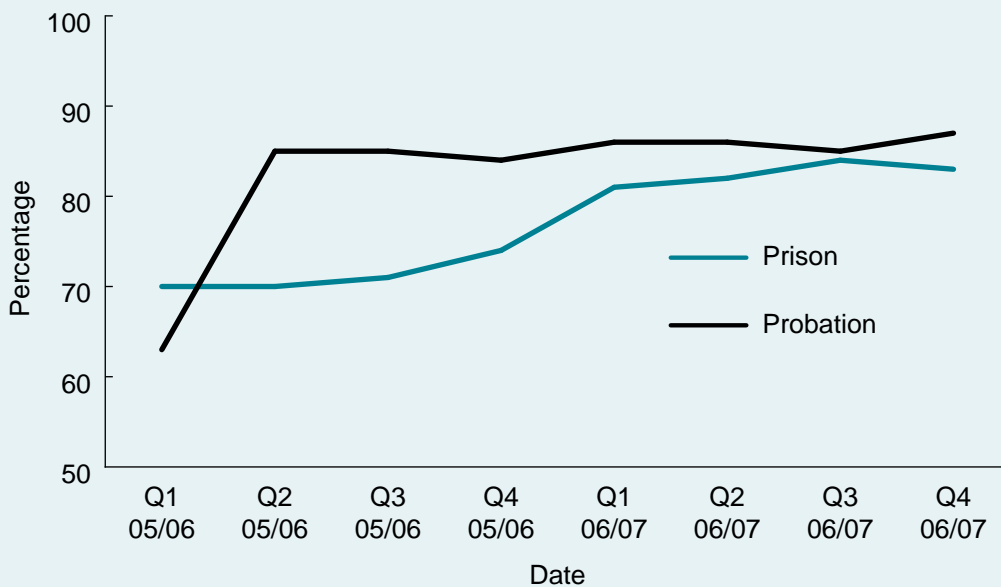
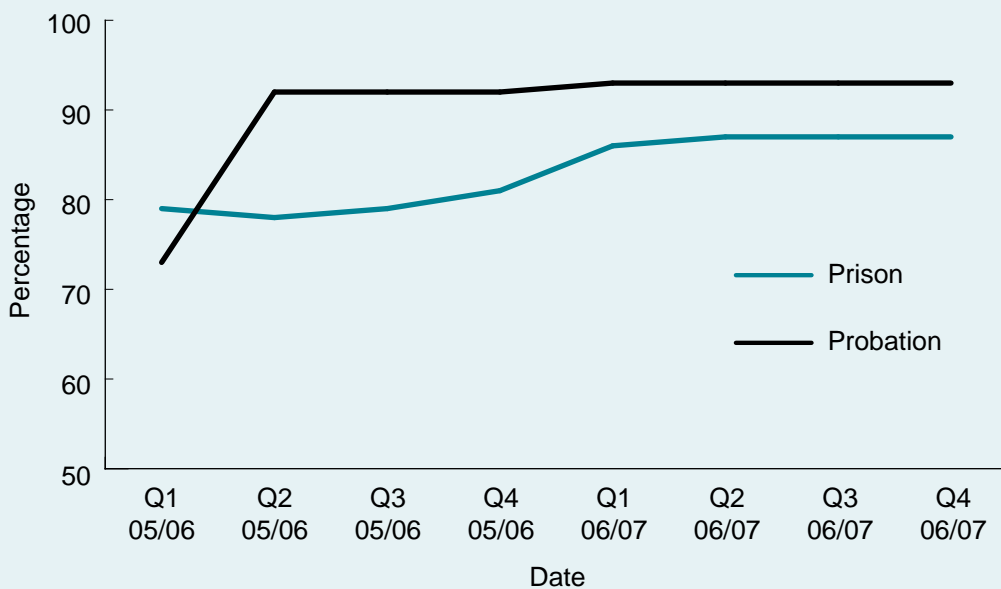


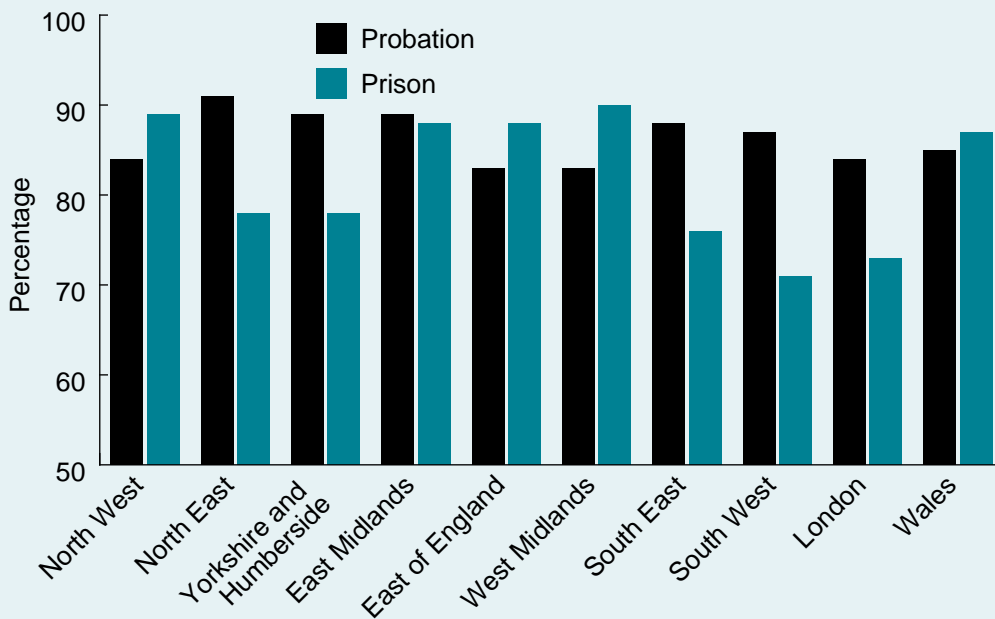
Figure 3.2: Quarterly trend for percentage of assessments with all valid section scores



Are there any regional differences in data completion?

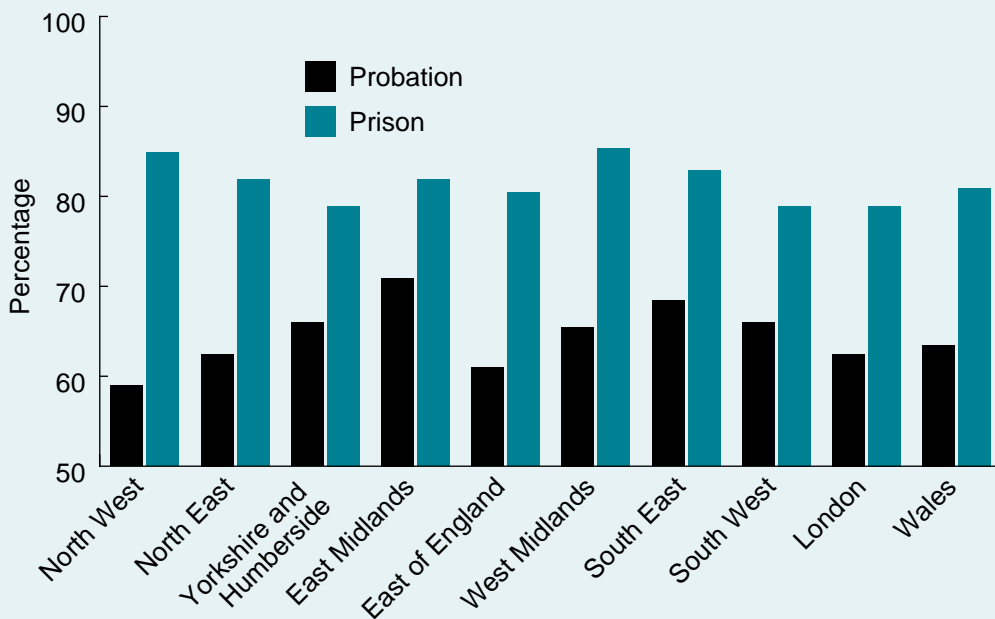
Figure 3.3 shows the percentage of probation and prison assessments in 2006/07 that were considered valid in each region. For probation assessments there was little variation between the regions with a range of between 83% and 91% of assessments deemed valid. There was wider variation between prison regions. The West Midlands had the highest percentage of valid assessments (90%) and was the only region to achieve the set target. In contrast, prisons in the South West had the lowest percentage of valid assessments (71%).

Figure 3.3: Percentage of valid probation and prison assessments by region



The SAQ had the lowest completion rate of all the OASys components measured. Figure 3.4 shows the regional differences for completion of all SAQ questions. There was little variation between the regions for prison assessments; however, probation did show regional variation. The region with the highest completion rate was the East Midlands (42%) and the lowest completion rate was the North West (18%).

Figure 3.4: Percentage of probation and prison assessments with all SAQ questions completed by region



Are there any demographic differences in data completion?

Table 3.6 shows the percentage of assessments in 2006/07 that met the minimum validity criteria broken down by offender gender, age and ethnicity. There were few differences in completion quality between different offender demographic groups. The largest difference was between the age group 18–20 where 77% of prison assessments were valid and the 21+ group where 84% of prison assessments were valid.

Table 3.6: Percentage of assessments that met minimum validity criteria in 2006/07

Offender group	Probation	Prisons
Male	86%	83%
Female	83%	79%
18-20	85%	77%
21+	87%	84%
White	87%	84%
Black	84%	80%
Asian	87%	83%
Mixed	86%	82%
Other	84%	78%

Implications

The levels of data completion were generally good for both probation and prisons, with the majority of items complete in over 85% of assessments in 2006/07. Both establishments had improved rates in the year 2006/07 compared to 2005/06. This indicates that a good proportion of start-of-community-sentence and start-of-custody assessments can be included in general research using OASys data. Additionally, in the majority of cases practitioners should have sufficient information from the OASys assessment to inform decisions on how to manage offenders.

There are, however, still further improvements that are desirable for both probation and prison assessments. Notably, the minimum completion criteria for an assessment to be considered valid was short of the set standard of 90% for both probation and prison assessments.

For probation, the components with the largest scope for improvement are the sentence code and the completion of all Risk of Serious Harm questions. In prisons, the component with the largest scope is completion of all scored questions in the core assessment.

The implications of these completion shortcomings are that the number of assessments that can be included in some specific research projects are reduced. For instance, criminogenic needs profiles for specific sentence types would be limited in probation assessments where the sentence has only been recorded in 72% of 2006/07 assessments.

The fewer scored questions that are completed, the less valid the total OASys likelihood of reconviction score will be. The questions that are most often not completed (in particular section 6) should continue to be monitored to establish if specific guidance is necessary or if there are some items that can never be universally available and should be excluded from OASys.

Implications for the future development and use of OASys are as follows.

- Local and national managers should explore with individual probation areas and prisons the reasons for OASys components with low completion rates and establish plans to drive improvement. This may include plans for further training or revising OASys guidance.
- Individuals using OASys data to inform policy decisions, research or resource allocation should be aware of sections of OASys and types of analysis that are potentially less accurate due to lower completion rates.
- Monitoring of OASys completion rates should continue to maintain data knowledge and identify where further improvements may be necessary. Where relevant this should include monitoring of individual practitioner performance to highlight any areas of difficulty.
- These findings should be used to inform the QMP review and developments on measuring the quality of OASys content.

Conclusion

Overall, completion rates for OASys were generally very good with some specific room for improvement or further investigation at local levels to understand where completion rates are not as high as they ideally should be. This evidence on OASys completion rates over time lends support to continued use of OASys in assessment and offender management, assuming that the quality of data recorded is also of similarly high standard. This research will be complemented by separate investigation of data quality in terms of accuracy, considered through audit and analysis of findings from OASys quality management processes.

4 The internal reliability and construct validity of OASys

Introduction

The analysis of OASys internal reliability and construct validity examined how well the items within each section of the core assessment measured various aspects of the same characteristic (internal reliability), and how well the assessment distinguished between discrete individual-level or social characteristics (construct validity). Both are necessary for establishing the extent to which the information captured by OASys, adequately achieves the intended assessment of offending-related factors.

The core OASys assessment

In its current format, the main body of OASys consists of 73 scored questions across 12 sections/11 scales (see Appendix 2) – the first two sections, which cover the current offence and offending history, being combined into an ‘offending information’ scale. The OASys scales and their scores are set out in Table 4.1. As can be seen, while the first scale covers offending information, the other scales focus on either individual-level factors, in terms of ‘internal’ disposition, personality, reasoning and temperament, or ‘external’ social or societal factors and their influences on offending behaviour.

Table 4.1: Scored OASys scales

OASys scales	No. scored questions	Score range	Criminogenic need cut-off
1-2: Offending information	11	0-22	7+
3: Accommodation	4	0-8	2+
4: Education, training and employability	9	0-18	5+
5: Financial management and income	5	0-10	5+
6: Relationships	6	0-12	4+
7: Lifestyle and associates	5	0-10	4+
8: Drug misuse	6	0-12	4+
9: Alcohol misuse	5	0-10	4+
10: Emotional wellbeing	6	0-12	3+
11: Thinking and behaviour	10	0-20	7+
12: Attitudes	6	0-12	4+

To guide the analysis in this research, the following three questions were set:

1. Do the scored questions within each scale measure discrete individual-level or social characteristics? (Internal reliability)
2. What are the common factors underlying the scored questions? (Construct validity)
3. What improvements could be made to the assessment of criminogenic needs? (Improving OASys)

Method

Sample

The electronic assessments are collated centrally within the Ministry of Justice in the O-DEAT database. The main sample for assessing internal reliability and construct validity was extracted from this database, selecting those assessments administered by probation assessors during the period April 2006 until March 2007 inclusive. All 73 scored questions within the core OASys assessment had to have been completed, ensuring that each individual-level and social factor was fully assessed. The sample was further restricted to the earliest valid assessment for each offender. This sampling left 230,163 assessments from all 42 probation areas.

A further sample of OASys assessments was used to look at the relationships with reoffending and to identify the most appropriate criminogenic need cut-off points. Assessments were restricted to those completed at the Pre-Sentence Report stage or at the start of either a community order or the licence period of a custodial sentence between October 2004 and March 2005 inclusive. It was ensured that these assessments included all 73 scored items from the core assessment, a sentence date and consistent risk of serious harm data. The assessments were submitted to the Police National Computer (PNC) criminal careers database managed by the Reoffending and Criminal History Team within the Ministry of Justice.⁶ Once successfully matched, the PNC records were processed to determine whether the cases could be followed up for 24 months at liberty, taking into account periods spent in custody and allowing three months for sentence and data entry to occur. This left a final sample size of 43,695 cases for use in the analysis.

Analysis

The most common **internal reliability** measure is Cronbach's alpha, and this was used to measure how well the individual questions in each OASys scale correlated with the sum of the remaining questions. Alpha scores generally increase when the correlations between the questions increase, thus indicating the extent to which each set of questions can be treated as measuring a discrete characteristic, that is, an individual-level or social problem. While a lenient cut-off of 0.6 can be used in exploratory research, many researchers require a cut-off of 0.7 for a scale to be considered 'adequate' and 0.8 for a scale to be considered 'good'. As stated by Oppenheim, '*Reliability, or self-consistency, is never perfect; it is always a matter of degree*' (1996:159).⁷

By comparing each scale's overall alpha score to the score produced when each individual question was removed, the results were used to indicate which questions were **not**

6 PNC numbers were recorded within OASys for most offenders, and an automatic matching procedure found reliable PNC numbers for most of the remaining cases. Cases in which the PNC did not record the offender's sex or recorded an unfeasible date of first or current conviction were rejected.

7 The required standard of reliability does vary, however, between subject areas. For example, cognitive tests tend to be more reliable than tests of attitudes or personality. More specifically, it is easier to construct a reliable test of a particular attitude than of a general one.

contributing to the scale's internal reliability. Item-scale correlations were also calculated to demonstrate which questions were poorly correlated with the total of scores on all other items.

Factor analysis was used to measure the overall **construct validity** of the core OASys assessment. Factor analysis assesses the interrelationships among a large number of questions and then explains these questions in terms of their common underlying dimensions (factors). The make-up of these factors gives information on the relationships between the individual questions. By comparing the factors to the established scales, the results can be used to:

- validate the established scales by demonstrating that their constituent questions load on the same factors;
- propose the construction of new scales; and/or
- propose the removal of questions which are weakly correlated with any specific factor and instead cross-load across factors (as indicated by low factor loadings, e.g. less than 0.4).

In order to assess which scales were measuring not only distinct problem areas but independent criminogenic needs, logistic regression was used to look at the associations with reoffending, taking into account the correlations between the scales themselves. Odds ratios were used to establish criminogenic need cut-off points, comparing the odds of reoffending for offenders with a particular score to the average odds of reoffending.⁸

Limitations

The main 2006/07 sample used in the analysis for assessing internal reliability and construct validity should not be seen as representative of the known offender population. Importantly, OASys is not completed with all offenders, and those offenders with an OASys were more likely to have committed a violent offence and to have a high likelihood of reconviction (see Chapter 9). The ability to validate the tool for all types of low risk offender is thus restricted. Additionally, the removal from the 2004/05 sample of those offenders who could not be followed up for 24 months at liberty due to subsequent periods in custody may also affect the findings. The inclusion of relatively fewer high likelihood of reconviction offenders is likely to have lowered the average reoffending rate, while the limited use of OASys with lower risk offenders is likely to have raised the average reoffending rate. Either of these offender shortfalls could have an impact upon the calculations of optimum mid-range criminogenic need cut-off points and the consequent targeting of interventions.

8 The PNC database, from which the reoffending data were obtained, lists offence dates and records cautions, reprimands and final warnings as well as convictions. This enables measurements of 'proven reoffending' within a given period, rather than the less complete measurement of whether an offender has been reconvicted.

Results

Reviewing the characteristics of the offenders in the main sample,⁹ 87% were male, 86% were of White ethnic classification, and their mean age was 32. The offender had committed an offence of violence against the person in 28% of the cases, while nearly a quarter (24%) had received a custodial sentence. Fewer than one in five (18%) of the offenders had a high OASys likelihood of reconviction score, while 8% were judged to present a high or very high risk of serious harm.

Internal reliability

As shown by Table 4.2, the following six OASys scales had high reliability (with Cronbach's alpha scores above 0.8), demonstrating that the questions within these scales were clearly measuring discrete characteristics:

- accommodation (0.937);
- alcohol misuse (0.881);
- offending information (0.867);
- thinking and behaviour (0.848);
- emotional wellbeing (0.827); and
- education, training and employability (0.819).

There is a limited amount of 'question routing' within OASys, in which the responses to certain questions are fixed by the responses to earlier questions. In the accommodation section, questions 3.4, 3.5 and 3.6 (see Appendix 2 for a list of all questions) are all scored 2 when the offender is recorded as being of no fixed abode or of transient accommodation (question 3.3). Similarly, within the alcohol-misuse section, questions 9.4 and 9.5 are scored 0 when questions 9.1, 9.2 and 9.3 have all been scored 0. To some extent, therefore, the high alpha scores for the accommodation and alcohol-misuse scales reflect how the scoring systems operate. It also needs to be borne in mind that alpha scores are a function of the number of scale items as well as the item correlations, facilitating high alpha scores for the longer OASys scales. These scales are offending information (eleven scored questions), thinking and behaviour (ten scored questions), and education, training and employability (nine scored questions). All three scales are listed above.

The following four OASys scales had adequate reliability (with Cronbach's alpha scores above 0.7), demonstrating that the questions within these scales were also measuring discrete individual-level or social characteristics:

- financial management and income (0.796);
- drug misuse (0.792);
- attitudes (0.740); and
- lifestyle and associates (0.709).

⁹ Valid percentages are provided. The ethnic classification was unrecorded in 16% of the cases, the offence category unrecorded in 13% of cases and the court sentence unrecorded in 40% of cases.

Table 4.2: Internal reliability of current OASys scales

Scale	Cronbach's alpha for the scale when numbered question is deleted													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Offending information	.867		.882	.851	.845	.853	.850	.853	.854	.856	.853	.841		.869
Accommodation	.937		.911	.911	.920	.927								
ETE	.819	.822	.782	.778	.794	.794	.808	.818	.809	.796				
Financial management	.796	.736	.714	.796	.784	.744								
Relationships	.580	.603	.480	.550	.574	.500								
Lifestyle & associates	.709	.597	.650	.716	.670									
Drug misuse	.792			.742	.740	.772	.808	.762	.720					
Alcohol misuse	.881	.840	.839	.881	.877									
Emotional wellbeing	.827	.777	.810	.792	.824	.811								
Thinking & behaviour	.848	.841	.839	.842	.827	.827	.832	.835	.824	.828				
Attitudes	.740	.696	.706	.694	.681	.721		.721						

Key: Bold font = increased alpha score for section when the question is removed.

Table 4.3: Item-scale correlations of current OASys scored questions

Scale	Item-scale correlation for numbered question													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Offending information			.128	.630	.702	.596	.682	.610	.610	.559	.611	.759		.420
Accommodation			.872	.868	.842	.822								
ETE		.401	.666	.692	.593	.570	.450	.344	.485	.581				
Financial management		.637	.700	.448	.486	.619								
Relationships	.433	.173	.431	.281	.210	.391								
Lifestyle & associates	.476	.609	.493	.309	.446									
Drug misuse				.618	.636	.506	.312	.551	.720					
Alcohol misuse	.785	.784	.790	.641	.641									
Emotional wellbeing	.694	.695	.539	.633	.512	.550								
Thinking & behaviour	.468	.495	.453	.463	.623	.623	.568	.535	.662	.616				
Attitudes	.502	.481	.481	.507	.566	.426		.422						

Key: Bold font = increased alpha score for section when the question is removed.

Relationships was the only scale with non-adequate reliability, with a Cronbach's alpha of 0.580. In other words, the questions within this section were failing to measure a discrete characteristic, and were instead measuring at least two factors which were weakly related to each other. While the sample used in this analysis was restricted to those assessments with all 73 scored questions completed, separate analysis of all start community/custodial sentence assessments during 2005/06 and 2006/07 revealed that the relationships section was most likely to have missing questions (see Chapter 3). The relatively poor completion of this section suggests that practitioners may have found it difficult to complete these questions reliably.

Table 4.2 also sets out the Cronbach's alpha scores for the scales when each individual question was removed (where the cell is empty there is no such scored question within the relevant section). Most scores were lower than the overall alpha, suggesting that the numbered question was contributing to the measurement of a discrete characteristic. For example, when question 9.1 was removed, the section's alpha fell from 0.881 to 0.840. But for a number of questions, the reductions were marginal, and for six of the 73 questions (indicated in bold in the table), the section's alpha score increased when it was removed, suggesting that it was not contributing to the measurement of that characteristic. These six questions were as follows:

- 1.3: Total number of separate offences for which convicted at this court appearance
- 2.14: Are current offence(s) part of an established pattern of similar offending
- 4.2: Is the person unemployed, or will be unemployed on release
- 6.2: Close family member has a criminal record
- 7.4: Manipulative/predatory lifestyle
- 8.7: Violent behaviour related to drug use.¹⁰

Turning to the item-scale correlations of the scored OASys questions, the following four questions (indicated in bold in Table 4.3) had correlations below 0.3, three of which are from the relationships section – the only section with non-adequate internal reliability:

- 1.3: Total number of separate offences for which convicted at this court appearance
- 6.2: Close family member has a criminal record
- 6.4: Current relationship with partner
- 6.5: Current partner has criminal record.

The further analysis of all start community/custodial sentence assessments during 2005/06 and 2006/07 revealed that, of all the scored questions, questions 6.2 and 6.5 were the least likely to be completed (see Chapter 3), suggesting that practitioners did not always have reliable information regarding the criminal records of partners and family members.

¹⁰ Question 4.2 regarding unemployment is used within the new OASys reoffending predictors, indicating that this question remains useful in terms of predictive validity.

Construct validity

The results of the factor analysis revealed 15 factors underlying the 73 scored questions, explaining 60 per cent of the variation in the variables (see Appendix 3).¹¹ Three of the OASys scored scales were each divided into two factors as follows:

- offending information: (i) youth/initial offending and (ii) adult/established offending;
- education, training and employability: (i) education and (ii) employment;
- relationships: (i) relationships and (ii) family offending.

Of the nine remaining factors, eight corresponded to the other scored scales within OASys, with a further factor focusing upon violence, comprising of three questions from two different OASys scales (8.7, 11.3 and 11.4). Eleven of the 73 questions (indicated in bold in Appendix 3) did not fall into factors corresponding to those OASys scales within which they currently reside. As shown by Table 4.4, five of these 11 questions were from the thinking and behaviour section.¹²

Table 4.4: Scored questions in non-corresponding factors

Scored question	Current section	Underlying factor
1.3: Total number of separate offences for which convicted at this court appearance	Offending information	Lifestyle and associates
5.4: Illegal earnings are a source of income	Financial management and income	Drug misuse
6.3: Experience of childhood	Relationships	Emotional wellbeing
7.1: Community integration	Lifestyle and associates	Employment
7.4: Manipulative/predatory lifestyle	Lifestyle and associates	Relationships
8.7: Violent behaviour related to drug use	Drug misuse	Violence
11.1: Level of interpersonal skills	Thinking and behaviour	Attitudes
11.3: Aggressive/controlling behaviour	Thinking and behaviour	Violence
11.4: Temper control	Thinking and behaviour	Violence
11.8: Achieves goals	Thinking and behaviour	Employment
11.10: Concrete/abstract thinking	Thinking and behaviour	Attitudes

11 The analysis used the principal components method, producing uncorrelated factors, as well as varimax rotation to maximise the variance of the loadings, helping to link each question to a single factor. The factors were restricted to those with an eigenvalue greater than one, recognising that further factors were contributing little to the explanation of variance in the variables.

12 The five thinking and behaviour questions remain useful in terms of predictive validity (see Chapter 6).

The loadings listed in the final column of Appendix 3 indicate the correlation between each OASys question and its respective factor – the higher the loading, the greater the contribution to the factor. For three of the 11 questions listed in Table 4.4 (6.3, 7.1 and 11.1), the factor loading was less than 0.4, indicating that the question was not a defining part of the factor but was cross-loading on several factors. Four of the 11 questions (1.3, 5.4, 7.4 and 8.7) were found above to have a negative impact upon the internal reliability of their current scales.

Revising the OASys scales

The factor analysis above revealed that the core OASys assessment had reasonable construct validity, with a fairly good match between the current OASys scales and the underlying factors, while the analysis of alpha scores revealed that all but one of the current scales had adequate internal reliability. Nevertheless, there was clear scope for improving the identification and measurement of discrete criminogenic needs – not all questions were contributing to the measurement of discrete individual-level or social problems and some loaded onto non-corresponding factors. Further analysis was thus conducted, concentrating upon the underlying criminogenic needs covered by sections 3 to 12 – the requirement for an offending-history scale having been removed through the development of the new reoffending predictors (see Chapter 6). Prior to running a further factor analysis, eight questions were removed on the following basis (three of which were removed for more than one reason):

- those questions which had a detrimental impact in terms of internal reliability (4.2, 6.2, 7.4 and 8.7);
- those questions which had low item-scale correlations (6.2, 6.4 and 6.5);
- those questions which fell within a distinct family offending factor (6.2 and 6.5); and
- those questions which fell within a distinct violence factor (8.7, 11.3 and 11.4).

The remaining questions were grouped into 11 factors corresponding to the current scales, but with questions 4.7 and 4.8 separated out into their own factor. A further five questions (5.4, 7.1, 11.1, 11.2 and 12.6) did not fall into factors corresponding to their current scales. The assessment was further streamlined by removing all seven questions. As shown by Appendix 4, the remaining 47 questions were grouped into ten factors corresponding to their current scales. All questions had a factor loading in excess of 0.4, and the ten factors were found to explain 62% of the variance in the variables.¹³ When restricting the sample to different offender subgroups – sampling on the basis of age, gender and ethnicity – nearly all questions loaded onto the correct factors, the exceptions being questions 7.5, 11.9 and 11.10 for female offenders, 11.8 and 11.10 for those offenders of Mixed Ethnic classification, and 7.5 for those offenders aged 25–40.

13 The factors were again restricted to those with an eigenvalue greater than one.

Using Cronbach's alpha cut-off points of 0.7 for adequate scores and 0.8 for high scores, Table 4.5 demonstrates that six of the ten revised scales had high internal reliability and a further two had adequate reliability, demonstrating that the questions within each of these scales were measuring a single need. As previously, relationships had non-adequate reliability, although the Cronbach's alpha had risen to above 0.6, while the alpha score for lifestyle and associates fell to below 0.7. Alpha scores are a function of the number of scale items as well as the item correlations, and both these scales had only three remaining questions.¹⁴ When three of the 47 questions (indicated in bold in Table 4.5) were removed, the relevant scale's alpha score increased, but these increases were minimal (.005, .006 and .011). As shown by Table 4.6, none of the questions had an item-scale correlation below 0.3.

Looking again at different offender subgroups – sampling on the basis of age, gender and ethnicity – the relationships scale and the lifestyle and associates scale remained the only two scales with alpha scores below the adequate cut-off point of 0.7, with the exception of section 12 for those offenders aged over 41 (score of .698) and those offenders of Asian ethnic classification (score of .645). The value to be gained from separate scales for different offender subgroups would thus appear to be limited.

Are the revised scales measuring criminogenic needs?

The results of the factor analysis demonstrated that the revised scales were measuring different individual-level or social problems. To assess whether these scales reflected criminogenic needs, the dataset linking 2004/05 OASys assessments with 24-month reoffending data was used. In the development of the risk-needs model, Andrews and Bonta (1995:176) stressed the importance of distinguishing between criminogenic needs and more general needs according to their relationship with reoffending, stating that the former are '*the dynamic attributes of an offender that, when changed, are associated with changes in the probability of recidivism*'. Criminogenic needs can thus be defined as those individual risk factors which contribute to or are supportive of offending and which are amenable to change.

14 Including the unscored question 6.7 regarding evidence of domestic violence/partner abuse (perpetration and/or victimisation) led to a very small improvement (6.32 compared to 6.23) in the internal reliability of the relationships section.

Table 4.5: Internal reliability of revised OASys scales

Scale	Cronbach's alpha	Cronbach's alpha for the scale when numbered item is deleted									
		1	2	3	4	5	6	7	8	9	10
Accommodation	.937			.911	.911	.920	.927				
ETE	.817			.774	.763	.788	.792			.822	.789
Financial management	.796		.719	.692		.803	.752				
Relationships	.623	.497		.481			.591				
Lifestyle & associates	.676		.519	.572		.648					
Drug misuse	.808				.767	.762	.798		.786	.733	
Alcohol misuse	.881	.840	.839	.838	.881	.877					
Emotional wellbeing	.827	.777	.777	.810	.792	.824	.811				
Thinking & behaviour	.843					.808	.808	.818	.831	.816	.821
Attitudes	.716	.657		.670	.644	.643			.727		

Key: Bold font = increased alpha score for section when the question is removed.

Table 4.6: Item-scale correlations of remaining scored questions

Scale	Item-scale correlation for numbered questions									
	1	2	3	4	5	6	7	8	9	10
Accommodation			.872	.868	.842	.822				
ETE			.647	.691	.595	.564			.476	.595
Financial management		.656	.707		.480	.592				
Relationships	.453		.460			.384				
Lifestyle & associates		.534	.497		.437					
Drug misuse				.619	.649	.508		.556	.730	
Alcohol misuse	.785	.784	.790	.641	.641					
Emotional wellbeing	.694	.695	.539	.633	.512	.550				
Thinking & behaviour					.670	.666	.615	.550	.626	.604
Attitudes	.501		.488	.529	.545			.365		

Recent analysis has looked at which individual questions within OASys, rather than the complete scales, are most predictive of reoffending (see Chapter 6) Logistic regression was used to account for the relationships between the OASys scored questions and static criminal history and offender demographic factors. All but one of the scored OASys questions included in the final models, and incorporated into the new OASys reoffending predictors (OGP and OVP), resided in the following six OASys sections, indicating that these domains had the strongest independent associations with reoffending:

- accommodation;
- education, training and employment;
- drug misuse (OGP);
- alcohol misuse (OVP);

- thinking and behaviour;
- attitudes.

The revised scales set out in this chapter were found to have statistically significant associations with reoffending ($p < .001$) – lower scores indicating lower probabilities of reoffending. However, in order to assess which of the revised scales were reflecting independent criminogenic needs, logistic regression was again used to take into account the correlations between the scales themselves. The OGRS 3 score (based upon static criminal history and offender demographic factors) and the six dominant scales in OGP/OVP were entered into the model. The other scales were then entered at subsequent levels. As shown by Table 4.7, the final model included seven of the ten revised scales. All seven scales had odds ratios greater than one, indicating that the odds of reoffending for those with higher scores were greater than the odds of reoffending for those with lower scores. The three scales which were not found to have independently significant associations with reoffending were: (i) the relationships scale; (ii) the lifestyle and associates scale; and (iii) the emotional wellbeing scale. In other words, while these three scales were measuring distinct problem areas, they were not measuring independently significant criminogenic needs. As noted above, the first two of these scales had only three remaining questions.

Table 4.7: Associations with 24-month reoffending

Scale	Parameter estimate	Standard error of estimate	Significance	Odds ratio
OGRS 3	0.28	.000	**	1.028
Accommodation	.042	.004	***	1.043
ETE	.029	.004	***	1.029
Financial management	.039	.006	***	1.039
Drug misuse	.077	.005	***	1.080
Alcohol misuse	.044	.003	***	1.045
Thinking and behaviour	.011	.005	*	1.011
Attitudes	.064	.008	***	1.066
Constant	-2.116	.027	***	.121

Asterisks indicate whether associations are significant (** $p < .001$; * $p < .01$; * $p < .05$).

When evaluating the factors underlying the currently scored OASys questions (see above), family offending and violence were separated out as discrete factors. Neither of these factors, when entered as separate scales, were found to have significant and independent associations with reoffending, failing to justify their inclusion as separate scales. The value of a violence scale would, in any case, be limited due to the development of the new OASys reoffending predictor for violence (OVP) (see Chapter 6).

Revising the criminogenic need cut-off points

Odds ratios can be used to compare whether the probability of a certain event is the same for two groups: an odds ratio of one indicating that the event is equally likely in both groups;

an odds ratio greater than one indicating that the event is more likely in the first group; and an odds ratio less than one indicating that the event is less likely in the first group. In this instance, odds ratios were used to set appropriate criminogenic need cut-off points for the seven remaining criminogenic need scales, comparing the odds of reoffending for offenders with a particular score to the average odds of reoffending.

Of the 43,695 offenders in the 2004/05 sample, 46.2% had reoffended within two years. As shown by Table 4.8, across all seven scales, there was a point at which the odds ratio increased to a value greater than one, i.e. where the reoffending rate surpassed 46.2%, so that the odds of reoffending for an individual with that score exceeded the average odds of reoffending (odds of 0.859). Applying the ‘need principle’ outlined by Andrews and Bonta above, this can be seen as the tipping point at which the use of interventions to address the need becomes more beneficial. There was also a point across six of the seven scales – the exception being accommodation – at which the odds ratio exceeded two (achieved by reoffending rates of at least 63.3% – odds of reoffending of 1.725). To further assist with the targeting of interventions, maximising the use of resources, a distinction was thus drawn, for these six scales, between three levels of need: (i) none/low, (ii) medium and (iii) high.

For example, in relation to the attitudes scale, the reoffending rate increased from 29.4% for a score of 0 to 74.1% for a score of 10. The odds ratio increased to a value above one for those offenders who scored 2 with an above average reoffending rate of 49.8%, while the odds ratio surpassed two for those offenders who scored 4 with a reoffending rate of 64.5%. On this basis, those offenders with scores from 0 to 1 had none/low levels of need, those with scores from 2 to 3 had a medium level of need and those with scores from 4 to 10 had a high level of need.

Table 4.8: 24-month reoffending rates by revised OASys scales

Score	24-month reoffending rate by scale						
	Accommodation	ETE	Financial management	Drugs	Alcohol	Thinking	Attitudes
0	38.4%	25.3%	34.6%	36.1%	43.9%	26.1%	29.4%
1	48.0%	36.4%	43.6%	54.9%	39.6%	33.1%	39.7%
2	52.5%	40.7%	47.2%	54.4%	41.2%	35.2%	49.8%
3	59.2%	45.5%	53.2%	57.8%	40.4%	40.4%	58.3%
4	60.3%	49.9%	56.1%	61.1%	42.0%	43.9%	64.5%
5	62.0%	53.5%	59.2%	63.6%	46.1%	49.4%	67.9%
6	55.2%	56.8%	60.7%	68.7%	48.4%	52.8%	71.0%
7	-	60.7%	63.7%	71.5%	51.5%	57.7%	71.1%
8	62.1%	63.1%	71.4%	75.3%	54.3%	57.2%	75.8%
9		65.5%		81.6%	60.7%	59.1%	82.5%
10		69.4%		83.3%	64.3%	60.9%	74.1%
11		69.0%				65.1%	
12		71.6%				66.2%	

Key: Criminogenic need level: None/low Medium High

Previous research has found that criminogenic needs and their associations with offending differ between males and females and between different age groups (e.g. Hollin and Palmer, 2006b; Raynor, 2007). Comparing the odds of reoffending for offenders with a particular score to the average odds of reoffending for different age and gender groups revealed some variation in the optimum cut-off points.¹⁵ As Table 4.9 demonstrates, the optimum none/low need/medium need cut-off points for females and males were different across four of the seven scales – the education, training and employment scale and the financial management and income scale having a higher cut-off point for females, with the alcohol misuse scale and the thinking and behaviour scale having a lower cut-off point for females.

Similarly, Table 4.10 demonstrates that the education, training and employment scale and the financial management and income scale had a higher medium need/high need cut-off point for females, with the drugs misuse scale having a lower cut-off point for females. Table 4.10 further demonstrates that, in contrast to male offenders, those female offenders with a maximum score on the accommodation scale had an odds of reoffending more than double the average odds of reoffending. For female offenders, therefore, a distinction could potentially be drawn between those with a high accommodation need and those with a medium level of need.

It would thus appear that establishing differing criminogenic need cut-off points for females and males could be beneficial. However, this finding remains tentative as the female sample was relatively small (n=6,553), resulting in some small sub-samples for particular scores across the scales. For example, Table 4.10 indicates that the optimum medium need/high need cut-off point for females on the financial management and income scale was 8, rather than the 7+ cut-off point for males. However, only 198 female offenders had a score of 7 on this scale and while the reoffending rate for these offenders fell just below the required cut-off point, the accompanying confidence interval produced a range of values for the true population both below and above the required cut-off point. Further analysis with larger samples is thus required to enable stronger conclusions to be drawn.

15 The analysis focused upon age and gender rather than ethnicity due to the relatively small sample sizes for some of the minority ethnic groups – 1,229 Black offenders, 1,147 Asian offenders and 645 offenders of Mixed Ethnic classification.

Table 4.9: None/low need / medium need cut-off points for revised OASys scales by gender and age

Offender group	n	24-month reoffending rate	None/low need / medium need cut-off point by scale						
			Accommodation	ETE	Financial management	Drugs	Alcohol	Thinking	Attitudes
All	43,695	46.2%	1+	4+	2+	1+	6+	5+	2+
Male	37,142	47.5%	1+	4+	2+	1+	6+	5+	2+
Females	6,553	38.5%	1+	5+	3+	1+	5+	4+	2+
18-20	6,929	59.4%	1+	4+	2+	1+	5+	5+	2+
21-24	8,590	52.8%	1+	4+	3+	1+	6+	5+	2+
25-40	20,809	45.9%	1+	4+	2+	1+	6+	5+	2+
41+	7,367	26.9%	1+	4+	2+	1+	5+	5+	2+
Male 18-20	6,101	61.4%	1+	4+	2+	1+	6+	5+	2+
Male 21-24	7,405	53.5%	1+	4+	2+	1+	6+	5+	2+
Male 25-40	17,462	47.1%	1+	4+	2+	1+	6+	5+	2+
Male 41+	6,174	28.0%	2+	4+	2+	1+	5+	5+	2+
Female 18-20	828	44.9%	1+	4+	2+	1+	5+	3+	2+
Female 21-24	1,185	48.4%	1+	5+	3+	1+	5+	4+	2+
Female 25-40	3,347	39.6%	2+	5+	3+	1+	5+	5+	2+
Female 41+	1,193	21.3%	1+	4+	3+	2+	7+	3+	2+

Key: Below cut-off for whole sample Above cut-off point for whole sample

Table 4.10: Medium need / high need cut-off points for revised OASys scales by gender and age

Offender group	n	24-month reoffending rate	Medium need / high need cut-off point by scale						
			Accommodation	ETE	Financial management	Drugs	Alcohol	Thinking	Attitudes
All	43,695	46.2%	-	9+	7+	5+	10	11+	4+
Male	37,142	47.5%	-	8+	7+	5+	10	11+	4+
Females	6,553	38.5%	8	9+	8	4+	10	11+	4+
18-20	6,929	59.4%	-	10+	8	5+	-	11+	4+
21-24	8,590	52.8%	-	8+	7+	6+	9+	10+	5+
25-40	20,809	45.9%	-	9+	8	6+	10	11+	4+
41+	7,367	26.9%	8	8+	8	4+	9+	11+	4+
Male 18-20	6,101	61.4%	8	10+	8	5+	-	11+	5+
Male 21-24	7,405	53.5%	-	8+	7+	5+	9+	10+	5+
Male 25-40	17,462	47.1%	-	9+	8	6+	10	11+	5+
Male 41+	6,174	28.0%	-	7+	7+	6+	9+	11+	4+
Female 18-20	828	44.9%	-	10+	5+	5+	10	10+	4+
Female 21-24	1,185	48.4%	8	9+	8	6+	10	11+	5+
Female 25-40	3,347	39.6%	8	10+	7+	6+	9+	10+	4+
Female 41+	1,193	21.3%	8	9+	-	4+	8+	11+	4+

Key: Below cut-off for whole sample Above cut-off point for whole sample

Applying the optimum cut-off points to all seven scales (without any variations between males and females), Table 4.11 compares the reoffending rates for those offenders with none/low-need levels, those with medium-need levels and those with high-need levels. The table confirms that, for the current sample, reoffending was below average for those offenders with none/low levels of need, above average for those offenders with medium levels of need and greatest for those with high levels of need.

Table 4.11: 24-month reoffending rates by revised criminogenic need cut-offs

	Offenders with none/low need		Offenders with medium need		Offenders with high need	
	% reoffended	Difference from mean	% reoffended	Difference from mean	% reoffended	Difference from mean
Accommodation	38.4%	-7.8%	55.4%	9.2%	-	-
ETE	34.1%	-12.4%	55.3%	9.1%	66.3%	20.1%
Financial management	37.1%	-6.5%	53.4%	7.2%	67.5%	21.3%
Drug misuse	36.1%	-10.3%	57.1%	10.9%	71.3%	25.1%
Alcohol misuse	43.4%	-2.3%	53.2%	7.0%	64.1%	17.9%
Thinking & behaviour	35.8%	-9.2%	54.7%	8.5%	65.7%	19.5%
Attitudes	35.0%	-11.0%	53.3%	7.1%	68.0%	21.8%

For each scale, the difference in reoffending rates between the three groups was statistically significant ($p < .001$). Table 4.12 compares the phi coefficients to those produced by the current criminogenic need scales, illustrating that five of the seven revised scales had stronger correlations with reoffending for all offenders, with no change for the accommodation scale. There was a marginally weaker correlation for thinking and behaviour (.204 compared to .216) due to the removal from the scale of those questions which were useful in terms of predictive validity but were problematic in terms of construct validity. All the correlations with reoffending were higher than those produced by the assessors' more clinical judgements – at the end of each OASys section, the assessor is asked to consider whether there is a link to offending behaviour (yes/no response).

Table 4.12 also demonstrates that the odds of reoffending for those with high levels of need were at least twice the odds of reoffending for those with none/low levels of need, with an odds ratio greater than four for the drug misuse scale and an odds ratio greater than three for the education, training and employability scale, the financial management and income scale, the thinking and behaviour scale and the attitudes scale. Distinguishing between offenders with none/low levels of need and those with high levels of need produced greater odds ratios than those produced by both the current criminogenic need yes/no scales and the assessors' more clinical judgements regarding links to offending behaviour.

Table 4.12: Correlations with reoffending and odds ratios for clinical judgements, current scales and revised scales

	Phi coefficients			Odds ratios		
	Clinical judgement	Current scales	Revised scales	Clinical judgement	Current scales	Revised scales (high need vs. none/ low need)
Accommodation	.121	.171	.171	1.93	2.10	-
ETE	.161	.256	.261	2.21	2.86	3.80
Financial management	.162	.169	.186	2.03	2.36	3.53
Drug misuse	.233	.240	.276	2.89	3.26	4.40
Alcohol misuse	.019	.077	.097	1.08	1.37	2.35
Thinking & behaviour	.131	.216	.204	1.97	2.43	3.43
Attitudes	.160	.210	.262	1.93	2.56	3.95

Returning to the sample of 2006/07 assessments used to produce the revised scales, Table 4.13 compares the criminogenic need prevalence rates produced under the current and revised systems. As can be seen, when including those offenders with medium or high levels of need under the revised system, there were increases in prevalence rates across four of the seven scales. By far the greatest change was a 31.8% increase for financial management and income; the greatest negative change was a 15.3% fall in alcohol misuse. Adjustments in the allocation of resources would thus be required to ensure that interventions were targeted at criminogenic needs. Comparing those offenders with high levels of need under the revised system to those with identified needs under the current system, the prevalence rates decreased across all scales. It would thus be possible to target interventions at more discrete groups of offenders than currently identified.

Table 4.13: Current and revised criminogenic need prevalence rates

OASys section/scale	% offenders with criminogenic need			% change from current	
	Current	Revised (medium or high)	Revised (high only)	Medium or high	High only
Accommodation	34.8%	49.0%	0%	14.2%	-34.8%
ETE	54.7%	50.2%	17.5%	-4.5%	-37.2%
Financial management	23.7%	55.5%	5.8%	31.8%	-17.9%
Drug misuse	26.1%	41.0%	18.0%	14.9%	-8.1%
Alcohol misuse	42.5%	27.2%	2.0%	-15.3%	-40.5%
Thinking & behaviour	55.9%	53.1%	5.2%	-2.8%	-50.7%
Attitudes	44.4%	50.4%	21.7%	6.0%	-22.7%

Revising the individual questions

When comparing the questions used within the revised scales (criminogenic or non-criminogenic) to those utilised within the initial version of the new OASys reoffending predictors (see Appendix 5), the following nine questions within sections 3 to 12 of OASys were left unused:

- 4.7: Has problems with reading, writing and/or numeracy
- 4.8: Has learning difficulties
- 5.4: Illegal earnings are a source of income
- 6.2: Close family member has criminal record
- 6.4: Current relationship with partner
- 6.5: Current partner has criminal record
- 7.1: Community integration
- 7.4: Manipulative/predatory lifestyle
- 8.7: Violent behaviour related to drug use

These nine questions could be removed from OASys unless: (i) they are found to be helpful in assessing risk of serious harm; (ii) they serve another specific purpose for practitioners; and/or (iii) further research reveals that they could be improved through amendments to their wording or accompanying guidance. The need for any new questions should be considered in light of research examining textual information recorded by assessors within each of the OASys sections, particularly the relationships, the lifestyle and associates and the emotional wellbeing sections (see Chapters 10 and 11).

As shown above, none of these revised scales was found to be measuring significantly independent criminogenic needs. The first two of these three scales have just three remaining questions, limiting their internal reliability. Importantly, other empirical studies have identified offenders' associates and family relationships as specific criminogenic risk factors, with Andrews, Bonta and Wormith (2006) identifying (i) anti-social associates and (ii) family and/or marital circumstances as two of the 'central eight' risk and/or need factors. Attempts should thus be made to identify alternative questions which are amenable to change and have stronger independent associations with reoffending.¹⁶

Implications

Focusing upon sections 3 to 12 of OASys, specific implications are as follows.

- While revisions would ensure that the ten scales were measuring distinct problem areas, three of the scales – the relationships scale, the lifestyle and associates scale and the emotional wellbeing scale – were not found to be measuring independently significant 'criminogenic' needs.

¹⁶ The only questions from sections 6, 7 and 10 of OASys currently used in OGP and OVP are 7.2 'Regular activities encourage offending' and the previously unscored 10.7 'Current psychiatric treatment or treatment pending'.

- Nine of the currently scored questions are not needed within the revised scales or the new violent and general reoffending predictors. These nine questions should only be retained if: (i) they are found to be helpful in assessing risk of serious harm; (ii) they serve another specific purpose for practitioners; and/or (iii) further research reveals that they could be improved through amendments to their wording or accompanying guidance.
- Those seven questions used within the new reoffending predictors but not the revised scales should remain in their current sections without contributing to the identification and measurement of criminogenic need.
- The criminogenic need cut-off points should be set in relation to the odds of reoffending. Adjustments in the allocation of resources would be required to ensure that interventions were available to address the revised criminogenic need levels.
- Offenders with 'high' levels of need should be distinguished from offenders with 'medium' levels of need to assist with the targeting of interventions, maximising the use of resources.
- The optimum criminogenic need cut-off points for different age and gender subgroups should be recalculated once larger samples are available. Any widening in the targeting of OASys would increase the validity of the calculations to the complete prison and probation caseloads.

Conclusion

The analysis of OASys internal reliability and construct validity indicates that not all questions load onto the correct factors and not all contribute to the measurement of discrete individual-level of social problems. An opportunity is thus available to streamline the assessment while improving its measurement of offending related risks and needs. The results of the factor analysis support a reduction from 73 scored questions across 11 scales to 47 scored questions across the ten individual-level or social problem scales. Additionally, new criminogenic need cut-off points could be set to be more closely related to their relationship with reoffending.

5. Inter-rater reliability of OASys

Introduction

Inter-rater reliability (IRR) is concerned with the consistency of measurement; with the degree to which an instrument measures the same way each time it is used, under similar conditions and with the same subjects.

IRR studies are a measure of reliability that assess the ability of tools such as OASys to deliver consistent assessments of offenders by asking two or more assessors to rate the same offender. The levels of IRR for OASys need to be high to ensure consistency in the decisions made by different assessors about an offender's criminogenic needs and supervision requirements.

Ensuring high IRR is also important due to the increasing use of collated OASys data to provide risk/needs profiles of offenders, measure the gap between provision and need, and measure the targeting of interventions. A reliable assessment tool will ultimately help to protect the public by improving the management of offenders to reduce reoffending.

The second and third pilot studies of OASys incorporated small-scale (n=46 and n=31 respectively) IRR tests (Howard, Clark and Garnham, 2006). Analysis from the second pilot, where assessments were reassessed and rescored by one of three experienced assessors, revealed that:

- criminogenic need was assessed very reliably in the criminal history section;
- assessment of attitudes, accommodation and relationships was also fairly reliable; and
- assessment of emotional and psychological factors¹⁷ was least reliable.

Looking at the overall likelihood of reconviction, the average difference in scores between assessors was 9.4 points (on a range from 0 to 168 points). Although exact agreement was uncommon, scores rarely differed by more than 15 points. Of the 12 sections where risk of serious harm was also assessed, only the emotional and psychological factors section achieved total agreement between assessors; criminal history, alcohol misuse and particularly the lifestyle and associates section were unreliable.

In the third pilot study, two of the three assessors used in the second pilot participated. Results revealed that while criminogenic need was assessed very reliably in the criminal history and drug misuse sections, the assessment of many factors was unreliable – it was thought that this may have been attributable to changes over time and over-lenient scoring by the principal repeat assessor. The average difference in scores for the overall likelihood of reconviction was 12.4 points. The greater number of large differences between scores

¹⁷ The criminogenic need sections have been slightly revised since the pilot.

compared to the second pilot was again attributed to individual assessment styles rather than fundamental difficulties with OASys itself. Of the ten sections where risk of serious harm was assessed, alcohol misuse was the most reliable, while the sections on emotional wellbeing and thinking and behaviour were least reliable.

At the pilot stage, the researchers concluded that OASys was not as reliable as the existing Level of Service Inventory-Revised (LSI-R) tool or the Assessment, Case Management and Evaluation tool (ACE). LSI-R has yielded correlation coefficients from 0.8 to 0.9 (Andrews and Bonta, 1995), compared to 0.71 in the second OASys pilot study, though differing research methods limited comparability. Similarly, Raynor *et al.* (2000) found both LSI-R and ACE to have acceptable levels of consistency; however, the sample sizes used in the research were fairly small. Since then, OASys training and documentation has improved greatly.

Since the piloting phases, one inter-rater study was completed by the Probation Service in 2004 (unpublished). A sample of 120 staff from 13 probation areas assessed two recorded live interviews with offenders. For the first video, analysis revealed 'excellent' agreement (90% or above) regarding the existence of a criminogenic need relating to accommodation, education and employment, financial management, relationships, alcohol misuse, thinking and behaviour and attitudes.

For the second video, there was 'excellent' agreement for education and employment, thinking and behaviour and attitudes, but 'poor' agreement (80% or below) for accommodation, relationships, alcohol misuse and emotional wellbeing. It was thought that the 'poor' agreement may have been due in part to participants not having sufficient information to inform their judgement and possibly misinterpretation of the manual guidelines. The following recommendations were made for future studies:

- providing good quality video interviews, as some videos in the study could not be used;
- increasing the pool of video offender interviews; and
- providing refresher training to reinforce previous learning.

The study of IRR reported here was conducted three years after the last study of OASys IRR, and aimed to address the methodological limitations of previous studies. Similar to the 2004 study, this research was based in the probation service and asked multiple assessors to complete an assessment on video-recorded case studies. In contrast to previous research, a larger sample of participants was recruited across a wider number of probation areas resulting in a more representative sample. The number of video case studies was also increased. The emphasis was upon the ability of assessors to reach similar judgements through interpretation of the same pieces of evidence.

The following questions were addressed by the research.

- 1 What is the extent of agreement between assessors on scored questions?
- 2 What is the extent of agreement on likelihood of reconviction profiles?
- 3 What is the extent of agreement on criminogenic need profiles?
- 4 How much variation is there in the assessment of section links to risk of serious harm and offending behaviour?
- 5 How much variation is there in the risk of serious harm ratings?
- 6 What is the inter-rater reliability of OASys as a whole?
- 7 Does assessor experience or demographic influence OASys judgements?

Method

There are two main approaches to assessing IRR. Firstly, it can be measured by two or more observers watching the same event and independently recording the variables according to a pre-determined coding system. Alternatively, it has been described as 'the ability of a group of raters, across multiple sites, to consistently apply proper interviewing techniques and scoring conventions and to reproducibly differentiate subjects based on their responses' (Endicott *et al.*, 2002). The latter approach is clearly wider in scope, encompassing interviewing techniques, in contrast to the former approach which requires observers to watch the same event.

In this study, the former methodology was chosen and DVDs of OASys interviews were created for participants to observe before completing an assessment. This method was selected because holding the interviewing skills constant allowed comparison of objective judgements through interpretation of the same information. If interviewing techniques had not been held constant, it would have been difficult to disentangle the impact of interviewing styles and the ability of OASys to produce consistent scores. This approach also enabled large numbers of assessors to rate the same case studies.

Three case studies were developed, based upon information from real assessments stored in the central OASys database. The case studies were designed to vary on age, gender, ethnicity, levels of offending, criminal history, criminogenic needs and the levels of risk posed.¹⁸ Table 5.1 shows a summary of the different case studies.

¹⁸ All key identifiers were changed to ensure data confidentiality.

Table 5.1: IRR Case studies

Case study	Basic demographics	Summary
1. Steve Doe	White male, 24	Index offence of burglary and some previous convictions. Currently unemployed and living with long-term girlfriend. Has a history of drug use and depression. Positive attitude but evidence of impulsivity.
2. Gillian Benalt	Black female, 37	Index offence of assault and no previous convictions. The assault occurred after alcohol consumption and the victim was a female having an affair with her husband. Good educational background, currently unemployed and living temporarily with a friend. History of suicide attempts.
3. Mustafa Ahmed	Asian male, 34	Index offence of supply of class A drug and previous convictions (including for drug-related offences). Currently employed as a bar manager and living alone. Family lives abroad. History of drug use and has an uncooperative attitude.

An OASys interview for each case study was professionally filmed with actors playing the offender roles (interviewers were played by one probation officer and one OASys prison colleague). A DVD for each interview was produced along with an information pack that contained instructions on how to complete the task and supporting information for each case study. The supporting information included:

- completed sections 1 and 2 of the OASys core assessment;
- a self-assessment questionnaire (which was deliberately absent for the third case study);
- pre-convictions; and
- information that practitioners might normally have been able to gather from other agencies.

Participants were randomly allocated one of the case studies. The assessors were instructed to first read the supporting information and then view the DVD of the OASys interview once only. Assessors then had to complete sections 3 to 12 of the OASys assessment, the risk of serious harm screening and the full risk of serious harm analysis should the screening have indicated this was necessary.

Participants were finally asked to complete a questionnaire to collect data on participant demographics and OASys experience.

Normative panel

A panel group of colleagues with experience and interest in offender assessment was asked to complete assessments on all three case studies before the research task was disseminated to participants. The use of the panel fulfilled two roles:

- it provided a pilot approach to remedy any initial shortcomings in the guidance and information packs; and
- the responses for each assessment were then used as normative scores against which the assessors' scores could later be compared.

The panel had the following members:

- one representative from HM Inspectorate of Probation;
- one representative involved in the design and development of OASys; and
- two representatives from the probation service.

Sample

Thirty-three out of 42 probation areas consented to take part in the research. A random sample of 296 practitioners that had completed OASys assessments during the week 25 to 31 March 2007 were identified from the central OASys database.¹⁹ The sample size from each probation area was approximately based on the proportion of OASys assessors working in the area. Therefore, London probation area had the largest target sample as it is the largest probation area. This proportionate approach was not exact because some areas were only able to offer a smaller sample than that requested.

The selected practitioners were first approached by a senior representative from their probation area, and given the opportunity to decline participation, in which case another randomly selected replacement was approached. In total 178 practitioners returned a completed OASys assessment giving a 60% response rate. The number of respondents for each case study and the average number of assessments completed per assessor each month is shown in table 5.2.

Table 5.2: Number of respondents per IRR case study

Case study	Number of respondents	Average no. of assessments completed monthly
Case study 1: Steve Doe	59	10.5
Case study 2: Gillian Benalt	57	9.4
Case study 3: Mustafa Ahmed	62	9.9
Total	178	9.9

To estimate how representative the participants were of practitioners throughout the Probation Service, a weighting was applied. This reflected the proportion of respondents from each probation area who had completed assessments for real offenders similar to the case studies in the financial year 2006/07. Likelihood of reconviction scores and risk of harm levels were compared. Table 5.3 shows the results of this analysis. There was little difference between the weighted and original sample in the likelihood of reconviction analysis; however,

¹⁹ 296 was the total number of assessors that probation areas agreed to contribute to the research.

the research participants may have had a bias in assessing high risk of serious harm more than other practitioners within the Probation Service.

Table 5.3: Likelihood of reconviction and risk of serious harm profiles for 2006/07 probation assessments and the respondent weighting

Group	Likelihood of reconviction			Risk of serious harm		
	Low	Medium	High	Low	Medium	High/Very high
2006/07 probation assessments	31.9%	49.6%	18.5%	44.2%	48.1%	7.8%
Weighted proportion	30.9%	50.4%	18.7%	40.4%	51.0%	8.6%

Analysis

To measure agreement on OASys scored question judgements for each case study, the average, mode and standard deviation statistics for raw section scores were analysed along with the percentage of assessors that agreed with the modal raw score. To identify the scored questions that were most problematic for assessors, the percentage of assessors that identified “no problem” was compared to the percentage that identified either “some problems” or “significant problems”, in their assessment.

To measure if differences in scoring the questions led to different offender profiles, the average, median, and standard deviation of the weighted total OASys score was reported for each case study. This score is grouped into three bands indicating likelihood of reconviction (low, medium or high). The percentage of assessors that scored each case study as low, medium or high was also compared.

To assess variation in criminogenic need profiles and clinical judgements on section links to serious harm and offending behaviour, the percentage of assessors that agreed with the normative panel’s assessment was compared for each case study. To measure differences in risk of serious harm ratings the percentage of assessors rating each case study’s highest risk rating as low, medium, high or very high was examined.

For all of the above analyses, the decisions from the normative panel’s assessment were also presented as a comparison. Agreement with the normative score or between assessors was considered excellent at 90% or above, good at between 80–90% and poor below 80%.

To measure the overall IRR of OASys for each case study, a two-way random-effects, single-measure absolute agreement intra-class correlation (ICC) model was used (Shrout and Fleiss, 1979 and McGraw and Wong, 1996). The scored OASys questions and RoSH ratings were included in the model. ICC correlation coefficients approach 1.0 when there is no variance within OASys questions, i.e. all raters give exactly the same ratings. Critical values

for single measure ICCs are: $ICC \geq 0.75$ = excellent; $0.60 \leq ICC < 0.75$ = good; $0.40 \leq ICC < 0.60$ = moderate; $ICC < 0.40$ = poor (Fleiss, 1986).

The F-test was also used to examine differences between assessor demographics and OASys experience on OASys subscales and total scores. For differences in categorical judgements Chi-square tests were used. Significant results are reported where $p < 0.05$.

Limitations

In practice, assessors conduct offender interviews themselves. Their interviewing styles and interactions may result in different information disclosed in interviews, which may affect the content of an OASys assessment. Because the offender interviews in this study are pre-recorded, the interviewing skills of practitioners are not measured. Therefore high IRR in this research may not translate directly to high IRR in practice.

The sample size was lower than anticipated (178 instead of 300) and not all probation areas participated. This meant that comprehensive comparisons between probation areas were not possible. Whilst assessors were randomly selected to take part, they were then able to refuse participation. The rejection rate was higher in some areas than others, particularly those with stretched resources (for instance if there had been a recent inspection). The implications of this are that the sample may be biased, i.e. those assessors that agreed to take part could differ from those who refused or from the whole population. If the assessors who refused to take part or who didn't respond are less competent or motivated than those that did take part, this would mean that this study underestimates the variation in use of OASys.

Assessors took part under different circumstances. In some areas, assessors were allowed to complete the exercise in work time. In other areas, assessors had to complete the exercise in their own time. Some were able to complete the exercise in a single session, while others reported that they could not rearrange their work to allow this and had to complete it in multiple, short sessions. This may have affected the quality of assessment completion.

There was a fair rate of missing responses. In total, 128 scored OASys questions were missing across all 178 assessments returned (39 questions for case study 1, 31 for case study 2 and 58 for case study 3).²⁰ It is unknown if the missing responses were deliberate because assessors were unsure of the answer or if questions were accidentally missed. This may have resulted in underestimating the variation if those assessors that were excluded were less competent or motivated. In practice, assessors complete OASys electronically. It would have been preferable and more reliable to use the electronic version of OASys as opposed to a paper version because the electronic version automatically completes some questions based on previous information recorded and in some instances highlights when questions have been missed.

²⁰ Each case study had 62 scored questions that should have been completed.

Fifty-five per cent of assessors indicated on the accompanying questionnaire that they found at least one OASys question difficult to complete. This was most frequently reported for case study 3 (65% of assessors reporting problems) and least frequently for case study 2 (42%). Fifty-nine per cent of assessors reported difficulty with at least one item for case study 1. Assessors were given the option to explain what they had found difficult and analysis of these comments revealed that across all three case studies some assessors had found it difficult to complete the assessment without doing the interview themselves as they preferred their own interview style. For case study 1, assessors reported more difficulty with the drug misuse and alcohol misuse sections because they would have liked to probe these issues in further detail. For case study 2, feedback reported most difficulty with mental health history in the emotional wellbeing section. For case study 3, there was no consensus on which items were most difficult.

Results

What is the extent of agreement between assessors on scored questions?

Table 5.4 shows the raw score awarded by the normative panel for each core assessment section, the mean, mode and standard deviation of the scores awarded by the assessors and the percentage of assessors that agreed with the modal score.

In case study 1, the standard deviations from the mean show that the majority of sections had good consensus, with values of less than 1.5 for accommodation, relationships, drug misuse and attitudes. In case study 2, the standard deviations from the mean again show that the majority of sections had good consensus, with values of less than 1.5 for ETE, relationships, lifestyle and associates, drug misuse and emotional wellbeing. In case study 3, the standard deviations from the mean show that four sections had good consensus, with values of less than 1.5 (accommodation, financial management, relationships and emotional wellbeing).

Section 11, thinking and behaviour, had a much higher standard deviation than any other section across all three case studies. However, section 11 also has a greater range of potential scores, which allows for a potentially higher mean and SD as a consequence. The assessor agreement with the modal score was highest across all three case studies for accommodation and drug use.

Table 5.4: Raw section scores for each IRR case study

Section (range of possible raw scores)	Case study 1					Case study 2					Case study 3					All case studies				
	Norm	Mean	mode	s.d	% agree mode	Norm	Mean	mode	s.d	% agree mode	Norm	Mean	mode	s.d	% agree mode	Norm	Mean	mode	s.d	% agree mode
S3 Accommodation (range from 0 to 8)	0	0.3	0	1.1	88%	8	6.2	8	1.8	35%	0	0.4	0	1.1	81%	0	0.4	0	1.1	69%
S4 ETE (range from 0 to 18)	11	9.9	9	1.8	29%	3	4.6	4	1.2	39%	6	4.5	5	1.8	27%	6	4.5	5	1.8	31%
S5 Financial management (range from 0 to 10)	2	3.6	3	1.6	25%	3	4.9	5	1.7	25%	1	2.2	2	1.2	47%	1	2.2	2	1.2	33%
S6 Relationships (range from 0 to 12)	1	2.6	3	1.3	32%	4	3.1	4	1.0	40%	4	4.8	4	1.4	26%	4	4.8	4	1.4	33%
S7 Lifestyle and associates (range from 0 to 10)	4	4.9	4	1.6	27%	0	1.5	2	1.1	30%	5	5.8	6	2.0	18%	5	5.8	6	2.0	25%
S8 Drug misuse (range from 0 to 12)	0	0.3	0	0.9	88%	0	0.1	0	0.5	98%	7	7.2	7	1.6	27%	7	7.2	7	1.6	70%
S9 Alcohol misuse (range from 0 to 10)	4	3.6	3	1.9	22%	7	7.1	8	1.6	30%	4	4.0	3	1.8	37%	4	4.0	3	1.8	30%
S10 Emotional wellbeing (range from 0 to 12)	2	2.8	3	1.6	25%	7	7.1	7	1.3	33%	1	0.7	0	1.0	60%	1	0.7	0	1.0	40%
S11 Thinking and behaviour (range from 0 to 20)	7	8.9	9	2.7	25%	8	7.9	8	3.3	16%	6	8.1	7	3.2	16%	6	8.1	7	3.2	19%
S12 Attitudes (range from 0 to 12)	3	2.5	2	1.4	31%	2	2.0	2	1.6	33%	2	2.0	2	1.6	26%	8	6.4	7	2.0	30%

OASys scores items as 0 (no problem), 1 (some problem) and 2 (significant problem) or 0 (no) and 2 (yes). Scoring will be least consistent if assessors fail to distinguish correctly between 'no problem' and 'a problem' or between the 'no' and 'yes' scored questions. To identify which questions had the least consistency, cross tabulations between 'yes and no' or 'score 0 and score 1 or 2' for each scored item were conducted. Table 5.5 shows the items where less than 80% of assessors were in agreement. The number of questions in each section is listed with the title of each section.

Five scored items transcending across all three case studies revealed a division of opinion. These were questions 6.1, 9.3, 11.3, 11.10 and 12.5. Sixteen other scored questions showed a division of opinion across two of the video case studies. These were questions: 4.4, 4.10, 5.5, 9.1, 9.2, 9.5, 10.2, 10.4, 11.2, 11.5, 11.7, 11.8, 11.9, 12.2, 12.6, and 12.8.

Three OASys sections were particularly problematic for all three case studies: alcohol misuse (section 9), thinking and behaviour (section 11), and attitudes (section 12). The sections with the least discrepancy were accommodation (section 3) where no questions were problematic for any case study and drug misuse (section 8).

Table 5.5: Percentage of assessors scoring no problem and a problem for each OASys question

Question	Case study 1		Case study 2		Case study 3	
	Problem		Problem		Problem	
	No	Yes	No	Yes	No	Yes
Section 3: Accommodation (4 questions)						
Section 4: ETE (9 questions)						
4.4 Work related skills			60%	40%	51%	49%
4.5 Attitude to employment	43%	57%				
4.6 School attendance					72%	28%
4.9 Any qualifications					52%	48%
4.10 Attitude to education/training	29%	71%	77%	23%		
Section 5: Financial management (5 questions)						
5.2 Offender's financial situation	24%	76%				
5.3 Financial management	44%	56%				
5.4 Illegal earnings source of income	56%	44%				
5.5 Over reliance on others for financial support			27%	73%	74%	26%
5.6 Budgeting impediment			46%	54%		
Section 6: Relationships (6 questions)						
6.1 Current relationship with close family	32%	68%	56%	44%	39%	61%
6.4 Current relationship with partner	69%	31%				
6.6 Previous experience of close relationships			66%	34%		
Section 7: Lifestyle & associates (5 questions)						
7.1 Community integration			30%	70%		
7.2 Activities encourage offending	31%	69%				
7.4 Manipulative/predatory lifestyle					49%	51%
7.5 Recklessness and risk-taking behaviour			79%	21%		
Section 8: Drug misuse (6 questions)						
8.4 Current drug use					26%	74%
Section 9: Alcohol misuse (5 questions)						
9.1 Current alcohol use a problem	34%	66%			66%	34%
9.2 Binge drinking in last six months	59%	41%			78%	22%
9.3 Level of alcohol use in the past	71%	29%	67%	33%	24%	76%
9.5 Motivation to tackle alcohol misuse	54%	46%			52%	48%
Section 10: Emotional wellbeing (6 questions)						
10.1 Difficulties coping	22%	78%				
10.2 Current psychological problems	39%	61%	52%	48%		
10.3 Social isolation	56%	44%				
10.4 Offender's attitude to themselves	39%	61%			69%	31%

Table 5.5: Percentage of assessors scoring no problem and a problem for each OASys question (continued)

Question	Case study 1		Case study 2		Case study 3	
	Problem		Problem		Problem	
	No	Yes	No	Yes	No	Yes
Section 11: Thinking and behaviour (10 questions)						
11.2 Impulsivity			46%	54%	25%	75%
11.3 Aggressive/controlling behaviour	29%	71%	26%	74%	37%	63%
11.4 Temper control					27%	73%
11.5 Ability to recognise problems	37%	63%	38%	62%		
11.7 Awareness of consequences			32%	68%	22%	77%
11.8 Achieves goals			37%	63%	52%	48%
11.9 Understands other people's views	37%	63%	23%	77%		
11.10 Concrete/abstract thinking	54%	46%	36%	64%	21%	79%
Section 12: Attitudes (6 questions)						
12.1 Pro-criminal attitudes	44%	56%				
12.2 Discriminatory attitudes			61%	39%	69%	31%
12.3 Attitude towards staff					52%	48%
12.5 Attitude to community/society	71%	29%	79%	21%	27%	73%
12.6 Understands motivation for offending	24%	76%	43%	57%		

What is the extent of agreement on likelihood of reconviction profiles?

The total score from all the items in OASys is used to calculate an overall likelihood of reconviction score which is grouped into three bands. A low likelihood of reconviction covers OASys weighted scores between 0 to 40, a medium likelihood covers scores from 41 to 99 and a high likelihood covers scores from 100 to 168. Table 5.6 details the total OASys score awarded by the normative panel for each case study and also the mean, median and standard deviation from the assessors. The percentage of assessors that judged the OASys weighted score to be within five, 10 and 15 points of the mean are also presented for each case study and the three case studies combined. The percentage of assessors that scored the offender as low, medium or high are also indicated.

Table 5.6: Normative, mean, median, standard deviation and likelihood of reconviction for each IRR case study*

OASys score	Case study 1	Case study 2	Case study 3	All case studies
Normative score	54	44	80	
Mean total score	58.4	48.7	78.3	
Median total score	57.0	48.0	79.5	
Standard deviation	10.2	10.2	11.8	
% with 5 points of mean score	36%	42%	29%	35%
% with 10 points of mean score	66%	63%	66%	65%
% with 15 points of mean score	86%	82%	74%	81%
Likelihood of reoffending				
Low	5%	23%	0%	
Medium	95%	77%	98%	
High	0%	0%	2%	

* Total OASys scores include the score for sections 1 and 2 which were pre-determined, i.e. not completed by the assessors and therefore no variation. It has been assumed that these scores would be totally reliable and therefore have no effect on overall agreement levels.

For all case studies the mean assessor score was within five points of the score decided by the normative panel, indicating close agreement. There was also good consistency of agreement on total scores between the assessors. The assessor scores in case study 1 give a median score of 57 and a mean of 58.4, indicating a symmetrical distribution. The inter-quartile range (middle 50%) of OASys scores was 51 to 65 and the standard deviation was 10.2. Ninety-five per cent of assessors scored the offender as having a medium likelihood of reconviction and 5% as having a low likelihood.

In case study 2, the mean and median scores were again close together, indicating a symmetrical distribution. The inter-quartile range was 43.0 to 55.5 and the standard deviation was 10.2. Despite the range in scores being reasonably narrow, they were distributed over the threshold for low and medium likelihood of reconviction. Approximately three-quarters (77%) of the assessors scored the offender as having a medium likelihood and 23% as having a low likelihood. However, there was wider variation here because the mean weighted score was 48.7 which is on the boundary between low and medium likelihood.

In case study 3, the mean was 78.3 and the median 79.5. The inter-quartile range was 71 to 87 and the standard deviation was the largest of the case studies at 11.8. While the range of scores was slightly wider, the majority (98%) of assessors' scores fell into the medium band for likelihood of reconviction.

What is the extent of agreement on criminogenic need profiles?

For each section score, there is a cut-off threshold indicating whether the offender is considered to have a criminogenic need. Table 5.7 shows the normative raw score, whether the normative score was classified as a criminogenic need and the percentage of assessors

that agreed with the normative panel. Cells are highlighted in the table where less than 80% of assessors were in agreement. The threshold for the raw score is listed because in cases where the normative score is within one point of the threshold we would expect there to be less assessor agreement as to whether there is a criminogenic need or not.

In case study 1, three sections had excellent agreement, three had good agreement and four were considered poor. The normative panel scored alcohol misuse and emotional wellbeing as criminogenic needs; however, only approximately 50% of assessors were in agreement.

In case study 2, five sections had excellent agreement, one had good agreement and four were poor. The normative panel scored ETE and financial management to not be criminogenic needs (and both were 2 points below the threshold). However, there was poor assessor agreement for both these sections.

In case study 3, five sections had excellent agreement, two sections had good agreement and three were poor. For the three sections with poor agreement, the normative score was close to the threshold.

Three sections overall had the highest agreement. Accommodation and drug misuse had excellent agreement in all three case studies and lifestyle and associates had at least good agreement in all of the three case studies.

Good agreement was observed for the relationships section across two of the case studies (1 and 3); however, poor agreement was found in case study 2 with 56% of assessors judging there to be no need and 44% scoring a need.

Excellent agreement was observed for the emotional wellbeing section across two of the case studies (2 and 3); however, poor agreement was observed in case study 1 with 44% of assessors judging there to be no need and 56% scoring a need.

The sections where at least two case studies had poor agreement were as follows.

1. Education, training and employability (ETE): across case studies 2 and 3, over 50% of assessors scored a criminogenic need and fewer than 50% scored no need.
2. Financial management: in case study 1, 24% of assessors judged that there was a need and in case study 2, 59% of assessors judged there to be a need.
3. Alcohol misuse: across case studies 1 and 3, approximately 50% of assessors scored a criminogenic need and 50% no need.
4. Thinking and behaviour: across case studies 2 and 3, approximately 70% of assessors judged there to be a need and 30% no need.

Table 5.7: Percentage of assessors scoring criminogenic needs for each IRR case study

	Case study 1			Case study 2			Case study 3			All case studies % agree with norm
	Norm score	Norm need	% agree norm	Norm score	Norm need	% agree norm	Norm score	Norm need	% agree norm	
S3 Accommodation (threshold = 2)	0	No	93%	8	Yes	97%	0	No	93%	94%
S4 ETE (threshold = 5)	11	Yes	98%	3	No	49%	6	Yes	53%	66%
S5 Financial management (threshold = 5)	2	No	76%	3	No	41%	1	No	97%	72%
S6 Relationships (threshold = 4)	1	No	83%	4	Yes	45%	4	Yes	86%	72%
S7 Lifestyle & associates (threshold = 4)	4	Yes	85%	0	No	96%	5	Yes	87%	89%
S8 Drug misuse (threshold = 4)	0	No	98%	0	No	98%	7	Yes	98%	98%
S9 Alcohol misuse (threshold = 4)	4	Yes	53%	7	Yes	100%	4	Yes	49%	67%
S10 Emotional wellbeing (threshold = 3)	2	No	44%	7	Yes	100%	1	No	91%	78%
S11 Thinking & behaviour (threshold = 7)	7	Yes	83%	8	Yes	70%	6	No	29%	60%
S12 Attitudes (threshold = 4)	3	No	71%	2	No	88%	8	Yes	92%	84%
Total number of needs	N/A	4	4.4	N/A	5	5.4	N/A	6	5.6	

Key:

Assessor agreement with normative panel less than 80%.

How much variation is there in the assessment of section links to risk of serious harm and offending behaviour?

As well as the scored questions, OASys asks assessors to make a clinical judgement (yes/no response) about whether each section is linked to serious harm and/or offending behaviour. Although these do not contribute to the scoring system they remain an integral part of the OASys assessment, in particular for sentence planning and risk management.

Table 5.8 shows the normative judgement for each clinical judgement and the percentage of assessors' that agreed with it. Highlighted cells indicate where there was poor consensus between assessors. In case study 1, there were five links to offending behaviour and two links to serious harm where agreement levels were poor (less than 80%). In case study 2, there was better consensus with one link to serious harm and one link to offending behaviour with agreement levels of less than 80%. In case study 3 there were three links to offending behaviour and five links to serious harm where agreement levels were less than 80%.

The sections with the least agreement for these clinical judgements were as follows.

1. Alcohol misuse – across two case studies there was poor agreement for the link to serious harm and in one case study poor agreement for the link to offending behaviour.
2. Thinking and behaviour – across all three case studies there was poor agreement for the link to serious harm.
3. Attitudes – in one case study there was poor agreement for the link to serious harm and across two case studies poor agreement for the link to offending behaviour.

Across all case studies combined there were six instances of poor agreement. Offending behaviour links for financial management and attitudes were poor. Risk of serious harm links for lifestyle and associates, alcohol misuse, thinking and behaviour, and attitudes were also poor.

The accommodation section had good agreement (at least 80%) for both links across all three case studies. ETE, relationships, and lifestyle and associates had good agreement except for one link in one of the case studies.

Table 5.8: Percentage of assessors scoring no and yes for section links to serious harm and offending behaviour for each IRR case study

Section	Case study 1		Case study 2		Case study 3		All case studies
	Norm	% agree	Norm	% agree	Norm	% agree	% agree
S3 Accommodation							
Serious harm	No	97%	No	96%	No	97%	97%
Offending behaviour	No	97%	No	100%	No	84%	93%
S4 ETE							
Serious harm	No	93%	No	98%	No	84%	91%
Offending behaviour	Yes	69%	No	95%	Yes	84%	83%
S5 Financial management and income							
Serious harm	No	95%	No	93%	No	90%	92%
Offending behaviour	No	22%	No	93%	Yes	74%	62%
S6 Relationships							
Serious harm	No	98%	Yes	86%	No	97%	92%
Offending behaviour	No	95%	Yes	95%	No	78%	87%
S7 Lifestyle and associates							
Serious harm	No	83%	No	95%	No	61%	78%
Offending behaviour	Yes	98%	No	88%	Yes	97%	93%
S8 Drug misuse							
Serious harm	No	100%	No	96%	No	52%	82%
Offending behaviour	No	79%	No	100%	Yes	98%	92%
S9 Alcohol misuse							
Serious harm	No	51%	Yes	96%	No	56%	67%
Offending behaviour	Yes	92%	Yes	100%	Yes	65%	85%
S10 Emotional wellbeing							
Serious harm	No	88%	Yes	95%	No	97%	92%
Offending behaviour	No	72%	Yes	93%	No	87%	83%
S11 Thinking and behaviour							
Serious harm	No	50%	Yes	66%	No	62%	58%
Offending behaviour	Yes	98%	Yes	93%	Yes	97%	94%
S12 Attitudes							
Serious harm	No	92%	Yes	20%	No	77%	62%
Offending behaviour	Yes	44%	Yes	42%	Yes	95%	60%

Key: Assessor agreement with normative panel less than 80%.

In case studies 2 and 3, there were greater levels of assessor agreement for the clinical links to offending behaviour than the criminogenic needs calculated from the scored questions.

How much variation is there in the risk of serious harm ratings?

A risk of serious harm screening is completed by assessors to determine whether a full risk of serious harm analysis is necessary. If the full analysis is deemed necessary, then practitioners are required to make informed judgements as to the level of risk the offender

poses to children, the public, known adults and staff. These judgements are grouped into low, medium, high or very high. If a practitioner decides that the full risk analysis is unnecessary then the offender is classed as low risk.

For the purposes of this analysis, the highest risk rating given to any of the four categories was considered. Table 5.9 shows the highest rating awarded by the normative panel and the percentage of assessors that judged each case study to be low, medium, high or very high risk. There was poor consensus in the risk of serious harm rating across two of the case studies. In case study 1, 32% of assessors rated the offender as low risk and 67% as medium risk. Similarly in case study 3, 27% of assessors gave a rating of low risk and 73% of high risk. There was good consensus for case study 2, with 4% rating the offender as low risk, 88% as medium risk and 9% as high risk. The majority of assessors did not agree with the risk rating agreed by the normative panel in case studies 1 or 3.

Table 5.9: Percentage of assessors scoring each risk of serious harm level for each IRR case study

Highest risk rating		Case study 1	Case study 2	Case study 3
Normative		Low	Medium	Low
Assessor judgement	Low	32%	4%	27%
	Medium	67%	88%	73%
	High	0%	9%	0%
	Very high	0%	0%	0%

The discrepancy in risk rating for case studies 1 and 3 appears to be influenced by the decision on whether completing the full risk of harm analysis was necessary. Table 5.10 shows the decision made by the normative panel and the percentage breakdown of assessor decisions on this judgement. In case studies 1 and 3 there is poor assessor agreement as to the necessity of the full screening.

Table 5.10: Percentage of assessors that decided the full risk analysis should be completed

Full analysis necessary?	Case study 1		Case study 2		Case study 3	
	Norm	Assessor	Norm	Assessor	Norm	Assessor
No	X	27%		2%		22%
Yes		73%	X	98%	X	78%

What is the combined inter-rater reliability of OASys?

Intra-class correlations (ICC) for the scored OASys questions and the four risk of serious harm ratings were calculated, as recommended by Shrout and Fleiss (1979). Each of the three case studies were considered separately as assessors did not assess the same offenders.

Because some participants did not complete all questions, it was necessary to exclude some assessments from the analysis. Where at least three participants had failed to complete the

same question, these questions were excluded from the analysis as it was possible that there were weaknesses in the case studies developed, i.e. not enough information was provided to answer all questions.²¹

Table 5.11 illustrates the results of the ICC analysis. Correlation coefficients for case studies 1 and 2 indicate good reliability and case study 3 indicates moderate reliability.

Table 5.11: Results of intra-class correlation coefficients for each IRR case study

	Case study 1	Case study 2	Case study 3
ICC – single measure	0.60	0.65	0.56
Alpha reliability coefficient	0.99	0.99	0.99
Number of raters included	54	52	55
Number of scored questions included	64	62	59

Does assessor experience or demographics influence OASys judgements?

Analysis of the data showed that there were no significant differences in OASys completion between different assessor grades or length of experience for the total OASys score or RoSH.

In case study 1, male assessors were significantly more likely to rate the offender as presenting a low risk of serious harm and female assessors more likely to assess the offender as medium risk ($X^2 = 6.1, df = 1, p < 0.05$). There were no significant gender differences for the other two case studies. This finding is unexpected and should be investigated in a future research project.

Summary of results

In summary, the inter-rater reliability of OASys was moderate. The total OASys score used to calculate the likelihood of reconviction showed good consistency; however, there were differences in the reliability of the individual sections, with some performing better than others. The most reliable sections were:

- **Accommodation:** which demonstrated excellent consensus across all three case studies and in comparison to the normative score.
- **Lifestyle and associates:** which demonstrated at least good consensus across all three case studies for criminogenic needs – although there was one case study where the assessors’ modal score disagreed with the normative score.
- **Drug misuse:** which demonstrated excellent consensus across all three case studies and in comparison to the normative score.

Moderately reliable sections included:

²¹ If those assessors that had to be excluded were less competent or motivated than others, this may mean that the ICC results show an underestimation of the extent of variation.

- **ETE:** there was poor consensus between assessors across two of the three case studies – however, there was consensus with the normative score across all case studies. There was one case study with poor consensus for the link to offending behaviour.
- **Relationships:** there was good consensus for scored criminogenic need across two case studies. However, there was poor consensus across two case studies for at least a third of the scored questions in the section. Additionally, in one case study, there was disagreement with the normative score.
- **Emotional wellbeing:** there was excellent consensus for the scored criminogenic need across two case studies, but poor consensus in the third. However, the modal score agreed with the normative score across all case studies. In case study one, there was poor consensus for the majority of the scored questions.
- **Attitudes:** there was at least good consensus across two of the case studies for the scored criminogenic need. However, each case study had poor consensus for at least three of the scored questions in the section. There were three instances of poor consensus for the links to serious harm/offending behaviour and one disagreement with the normative score.

The least reliable sections were:

- **Financial management:** there was poor consensus between assessors for the scored criminogenic need in two case studies and for the link to offending behaviour in two case studies. There was disagreement with the normative panel in one case study.
- **Alcohol:** there was poor consensus both between assessors and in comparison to the normative score for the scored criminogenic need in two case studies, and three instances of poor consensus for the links to serious harm and offending behaviour.
- **Thinking and behaviour:** there was poor consensus between assessors for the criminogenic need in two case studies and disagreement with the normative score in one case study. Additionally, there was poor consensus in all case studies for the link to serious harm.
- **Risk of serious harm:** there was poor consensus between assessors for the highest risk of serious harm rating for two case studies and good agreement for one case study. The decision whether the full risk of harm screening required completing was mainly responsible for disagreement between assessors.

Implications

For the sections of OASys with less assessor agreement, the implications are that similar offenders may be assessed differently and as a result experience different supervision requirements. Based on the findings from this research, this is most likely to occur for financial management, alcohol misuse, and thinking and behaviour needs.

The most concerning finding is the variability in the completion of the risk of serious harm components. Two of the three case studies had poor consensus as to the highest level of risk posed by the offender, in particular whether it was necessary to complete the full risk of serious harm analysis. Assessors were more likely to overestimate the risk of harm and the need for full screening, which in practice would tend to generate “false positives” (i.e. offenders assessed at a higher risk than they pose in reality). This may result in poor use of resources with over-intervention.

Implications for the future development and use of OASys are as follows.

- Questions with poor consensus across all three case studies could be candidates for removal in a revised version of OASys and should be avoided in any new short-format OASys. Specifically the following questions are recommended for removal should other OASys reliability and validity research agree that they are not required: 6.1 (current relationship with close family), 9.3 (level of alcohol use in the past), 11.3 (aggressive/controlling behaviour), 11.10 (concrete/abstract thinking), and 12.5 (attitude to community/society).
- To improve the reliability of identifying which offenders require an intervention to address a criminogenic need, core assessment section scores should be classified into three groupings instead of the current two (need or no need). The three groupings should be high, borderline and no need. High need offenders would be referred to the appropriate intervention, and no need offenders would not. For borderline need offenders, the section link to offending behaviour (clinical judgement) would be used to determine if the need should be addressed.²²
- The OASys manual, which provides guidance to practitioners using OASys, should be reviewed for the sections of the core assessment with poor consensus for one or two case studies. Decisions should be made as to whether the guidance for any sections and/or individual questions needs updating. In particular, assessors commented that the guidance for the alcohol section was less clear given recent government campaigns. For instance, question 9.2 refers to binge drinking and assessors were unsure what the definition of this was. Similarly, four of the questions in the thinking and behaviour section had poor agreement for two case studies and this section should therefore be prioritised for revision.
- Areas and NOMS HQ should further explore the reasons for variability in the completion of the risk of serious harm component. It may be necessary for guidance and training to be refreshed or alternatively, the risk of serious harm tool may need to be revised.

²² An alternative approach for distinguishing between three levels of need is presented in Chapter 4.

- Areas and NOMS HQ should ensure that there is continual refresher training regarding the scoring of individual questions. Local areas should use the DVD recorded interview to deliver training and use the materials to discuss and resolve with practitioners any areas of ambiguity.
- It would be useful to replicate this research with prison assessors to firstly compare IRR amongst prison assessors and secondly to compare the prison and probation services.

Conclusion

IRR is concerned with the degree to which OASys is consistent and measures the same way each time it is used, under the similar conditions and with the same subjects. The levels of IRR for OASys need to be high to ensure consistency in the decisions made by different assessors about an offender's criminogenic needs and supervision requirements.

Poor IRR will impact upon both the measurement of an offender's likelihood of reoffending (the risk principle) and the measurement of discrete criminogenic needs (the criminogenic need principle), which in turn impacts on effective offender management.

This research has found the overall IRR of OASys to be moderate, with some sections providing more assurance than others in producing consistent results from different assessors. Shortfalls may be addressed by improved training and guidance on completion or by changes to the item content of OASys, informed by the findings from this IRR study and other research on reliability and validity of OASys included in this compendium.

6. Predictive validity of OASys – Improving prediction of violent and general reoffending

Introduction

Predictive validity refers to the extent to which scores on an assessment tool are able to predict some outcome measure. In this chapter the focus is upon the ability of OASys to predict further offending, based on its assessment of offending-related risks and needs. It is essential that OASys provides valid indications of future reoffending based on accurate assessment of need and risk, in order to enable finite resources for offender management and risk management to be used as effectively as possible.

A study of reoffending following the pilots of OASys (Howard, Clark and Garnham, 2006; Howard, 2006) found the scoring system to be a reasonable predictor of general reoffending, but not as accurate as the Offender Group Reconviction Scale Version 2 (OGRS 2). OGRS 2 is an actuarial predictor of reconviction based on static risk factors: age, sex and criminal history.

The predictors in widespread use for adult offenders in England and Wales are OGRS and the OASys scoring system.

Several more complex tools are in limited use in England and Wales, typically with long-term prisoners or in forensic psychiatric units. These measure psychopathy (PCL-R, PCL-SV: Hare, 2004; Hart, Cox and Hare, 1995) or assess the individual's risk of violence using actuarial (VRAG: Quinsey, Harris, Rice and Cormier, 1998) or Structured Professional Judgement (HCR-20: Webster, Douglas, Eaves and Hart, 1997) approaches. The amount of time and assessor training required to complete these tools properly renders their widespread use unviable. Therefore, only OGRS and OASys are considered further in this chapter.

OGRS has been used in the probation service for the last decade. It is included in OASys, but is also used without OASys in the preparation of Fast Delivery Reports, the shorter form of pre-sentence report. It also appears effective for predicting harmful offences, and has the advantage that it “takes ten minutes and requires very little training as no interpretation of the findings is required in determining future risk” (Maden *et al.*, 2006). The scored items include age, sex, age at first offending, experience of custody, previous burglary and breach of order and the ‘Copas rate’ of offending (a function of the number of previous convictions and the speed at which they were acquired). The system was originally manually scored, then computerised and updated as OGRS 2 in 1998, and a predictor of sexual and violent reconviction (OGRS-SV) was added (Taylor, 1999).

A further revision of OGRS was undertaken between 2004 and 2006 (Howard, Francis, Soothill and Humphreys, 2009). The new version, OGRS 3, was implemented as part of OASys and as a standalone software application in March 2008.

OGRS 2 achieved excellent predictive validity (accuracy of prediction), and OGRS 3 is better still. They can also be completed quickly and produced by administrative staff. However, they are limited in that they do not identify dynamic risk factors which can be addressed in order to reduce an offender's likelihood of reoffending.

The OASys score is produced once the offending-related factors component of OASys is completed. Scores from a number of questions on each of the 11 factors are combined and weighted to give a total score. A maximum score of 168 is available – this unlikely number being chosen deliberately to ensure that the score is not mistaken for, say, a percentile predictor, as the system was not calibrated to attach specific reoffending rates to given scores. Up to 50 weighted points are available from sections 1 and 2 of OASys, which cover criminal history and an analysis of the current offence. The remaining 118 points are based on sections 3 to 12, the ten dynamic risk factors: accommodation (12 points), education training and employability (ETE) (20), financial management and income (12), relationships (6), lifestyle and associates (15), drug misuse (15), alcohol misuse (5), emotional wellbeing (6), thinking and behaviour (12) and attitudes (15). Age and gender are not scored.

While the OASys score was found to be a reasonably good predictor of future offending, showing some predictive validity, it did not perform as well as the OGRS predictors in identifying future risk. Analysis since the pilot studies suggested that improvement was needed.

The research described in this chapter sought to:

- 1) improve the current OASys scoring system and the predictive validity of OASys;
- 2) test new tools to separately predict violent and other reoffending, known as the OASys Violence Predictor (OVP) and OASys General reoffending Predictor (OGP);
- 3) make these new predictors user friendly; and
- 4) include dynamic risk factors – items which can change over time – in the new predictors, so that they reflect changes in offenders' needs and can be used as the basis for sentence plans.

Method

Defining violent reoffending

Existing predictors of violent reoffending use a range of definitions of violence. Those of the structured clinical assessments tend to be victim-focused, grouping all offences which cause physical harm to another individual. Meanwhile, OGRS-SV uses a far narrower definition based on the Home Office coding of offences of “violence against the person” (Home Office, 2006). This chapter takes an alternate, evidence-based approach to produce a definition which groups offences which not only include interpersonal violence but tend to be committed by offenders with similar risk factors. This type of grouping should produce an accurate and practical prediction of future violent offending, for two reasons: firstly, a

history of committing one of the group's component offences should help to predict future offending across the group, and secondly, high scores on the predictor's dynamic risk factors will represent meaningful treatment targets, as addressing these factors should reduce the likelihood of committing each of the offences in the group.

OASys-based predictors of violent and non-violent reoffending were produced in the following stages:

- 1) selecting suitable OASys assessments;
- 2) matching these assessments with the Home Office Police National Computer (HOPNC) database to obtain data on offending history and reoffending;
- 3) producing a definition of violent offending;
- 4) coding violent and non-violent reoffending outcomes, and corresponding measures of offending history;
- 5) fitting statistical models of violent and non-violent reoffending, and producing scoring systems based on the model results; and
- 6) comparing the predictive validity of the new systems with that of existing predictors.

These stages are now described in more detail.

Selecting suitable OASys assessments

The initial sampling frame comprised all assessments completed between the start of 2002 and September 2004. These were filtered to ensure completion of all 73 offending-related factors questions which are included in the OASys scoring system, and basic consistency and completeness of the risk of serious harm section. The sentence date – often missing in OASys – was also needed to make matching with PNC data feasible. The remaining assessments were narrowed down to assessments completed at the start of either a community order or the licence period of a custodial sentence, or at pre-sentence report. Only one assessment per offender per sentence was included, to avoid double-counting.

Matching OASys assessments with the Police National Computer

The surviving cases were submitted to HOPNC. PNC ID numbers were available for most offenders, and the automatic matching procedures of the Ministry of Justice's Reoffending and Criminal History Team (who administer HOPNC) found reliable PNC IDs for most of the remainder. The resulting PNC records were checked to ensure that they included a conviction on a date corresponding with that recorded in OASys. Records where the PNC did not record the offender's sex or implied that the offender had been convicted aged under ten (the age of criminal responsibility) were also rejected as unreliable.

Where the assessment was associated with a custodial sentence, OASys was checked to see if a discharge date was recorded (as this would be the follow-up start date). This was

not present in the majority of cases, so the Reconviction Analysis Team matched the PNC records of such cases with RDS-NOMS prison discharge records, using sentence date as an additional check that the correct custodial sentence was identified.

Finally, the PNC records for the remaining offenders were processed to code whether proven reoffending occurred within 24 months of the start of community sentence or custodial discharge. Some offenders were removed as their follow-up was biased by imprisonment for an offence committed before the start of the 24-month period (a pseudoreconviction).

The PNC data were extracted on 8 January 2008. As the “proven reoffending” measurement allows conviction to occur outside the 24-month period provided that the offence itself occurred within 24 months, an extra 12 months was allowed to ensure that most convictions for offences committed within 24 months had occurred and had been recorded on the PNC. Table 6.1 summarises the selection, matching and follow-up process.

Table 6.1: Selection and matching of assessments

Stage of matching	Number of records
Assessment at start of community order or licence, or pre-sentence, completed by 30 September 2004	198,103
De-duplicated (only one assessment per offender per sentence)	152,358
Checked for data quality	42,609
Found on PNC (and prison discharge file if necessary), with matched sex, plausible age and conviction history, and agreement on sentence (and discharge if necessary) dates	32,396
Successfully followed up for 24 months, allowing for 12-month reoffending “buffer” period	26,619

The final group of 26,619 had the following key characteristics: 18% were aged 18–20, 21% aged 21–24, 46% aged 25–40 and 14% aged 41+; 14% were female; 16% were sentenced to custody; 7% were known to be non-White and a further 7% had no ethnicity data; 22% were convicted of violence against the person, 2% sexual offences, 6% burglary, 2% robbery, 17% theft and handling, 4% fraud and forgery, 3% criminal damage, 6% drugs offences and 38% other offences, including motoring offences.

As the selection process involved the rejection of a large number of assessments, the group at the second stage of matching with duplicates removed (n=152,238) and the finally selected group (26,619) were compared on key offender characteristics. The 26,619 remaining assessments differed from the whole 152,358 in including a greater proportion of offenders serving Community Punishment Orders (CPO) (28% against 23%, excluding those with unknown sentence) and a lesser proportion serving custodial sentences (16% against 24%). There were also some differences in probation area distributions, due to uneven data quality. The low proportion of custodial sentences, especially short sentences (offenders sentenced

to under 12 months do not routinely receive OASys assessments), and high proportion of CPO offenders among the 26,619 offenders is likely to be a factor in their relatively low rate of overall reoffending. Table 6.9 shows the final OASys sample had a 49% rate, compared with the 55.5% of the Home Office's 2004 cohort (Cunliffe and Shepherd, 2007).

A further comparison was made between the group of 26,619 and a sample of offenders assessed in the first quarter of 2007, with duplicates were removed (n=75,529). The 2004 group had a very skewed distribution of probation areas, with only 1% from London and almost none from the West Midlands, as these areas were late adopters of the electronic version of OASys. In 2007, 10% were from London and 7% from the West Midlands. This geographical bias contributed to a change in ethnic profile: 2.7% of the 2004 sample were Black, whereas 5.3% of the 2007 sample were Black. The age distribution was also different, with the proportion aged over 40 rising from 14% in 2004 to 19% in 2007.

To check that the differences between rejected and matched assessments did not invalidate the results, the results section includes checks that the predictors worked well for a range of groups of offenders. Future validation studies will use a sample later in 2004 and 2005 to check the validity of the predictors in all 42 probation areas and apply the predictors to groups who were under-represented in this sample. This will be important in order to address any concerns about the representativeness of this sample.

Producing a definition of violent offending

Several approaches were combined in order to produce a robust definition of violent offending. At the time of the research, OASys used the Home Office (2006) offence codelist, and spreadsheets listing the code, subcode and description of each offence were consulted in order to create initial offence groupings. Offences of homicide through interpersonal violence (i.e. excluding motoring-related deaths), wounding, weapon use and assault ('homicides and assaults' below), which are universally agreed to be violent, formed a reference group with which other offences would be compared. Some offences which seem similar to homicides and assaults involve weapon possession but not necessarily weapon use, and these were placed into a different initial group.

Then, OASys data were used to check violent offence content, and the criminogenic needs of those convicted of different offences were compared. Where violent offence content was frequent and needs were fairly similar to those of the homicide and assault group, simple comparisons of the associations between criminogenic needs and reconviction for each offence were made to see if each offence group could be predicted with similar factors as homicide and assault. Additionally, simple statistical models were fitted to check the association between previous sanctions for each offence group and future sanctions for homicide and assault and a subset of the most harmful of these offences – homicide and attempted murder, and wounding with intent to inflict grievous bodily harm. Only offence groups which satisfied all of these checks were added to homicides and assaults to produce a final 'violent-type' offence group.

Violent offence content

The OASys Analysis of Offences section includes the following Yes/No questions:

- 2.2A Carrying/use of a weapon
- 2.2B Any violence or threat of violence/coercion
- 2.2C Excessive/sadistic violence
- 2.3D Physical violence towards partner

Offence codes in which at least some of these questions are often ticked may be considered violent. Given that the purpose of the definition is to predict future violence rather than classify past offending, a very high percentage on this test was not necessary. Offences with moderate results on this test would, however, need to demonstrate especially strong results on the other tests to ensure that the prediction of homicide and assault was not compromised.

Three of these questions (2.2B, 2.2C and 2.3D) are certain indicators of violence. After discussions with Ministry of Justice colleagues, it remained unclear whether carrying of a weapon when it was not used in the offence (i.e. if 2.2A was ticked and the other three items were not) usually indicates genuine violent intent rather than a need for social status or protection. Offences in this group would need to demonstrate considerable similarity to homicides and assaults on the remaining tests to be classed as 'violent-type'.

Criminogenic need profiles

Criminogenic needs are defined using the OASys scoring system. As well as producing a weighted score, each offending-related factor also produces a Yes/No criminogenic need measure. A criminogenic need is said to be present when the offender scores above a certain threshold, set at between 25% and 50% of the maximum unweighted score available for the offending-related factor. For example, an accommodation need is present when the offender scores at least 2 of a maximum 8 points, while an ETE need is present when at least 5 of 18 points are scored.

An offence group's criminogenic need profile is produced by computing the percentage of those convicted of offences in that group who are scored with each criminogenic need. For example, if 50 of 200 offenders scored at least 2 on accommodation, that group's profile would include 25% accommodation need. Criminogenic need profiles are included below to compare those convicted of different types of offences which may be considered 'violent'.

It is important that the groups of offences which will be counted as 'violent-type' show some similarities in criminogenic need profile. This is because the predictors will score offenders on their levels of static and dynamic risk factors – the dynamic risk factors being these criminogenic needs. If those convicted of different types of offence have different criminogenic needs, it is unlikely that a single set of dynamic risk factors could predict reoffending involving each of these offence groups accurately.

Associations between previous sanctions and reoffending

Where violent offence content and criminogenic need profiles did not rule out the classification of an offence group as violent, associations between counts of previous sanctions for those offences and reoffending were checked using a set of logistic regression models. These models control for the effect of all offence counts simultaneously, so are more robust than just finding a correlation between a given offence count and reoffending.

To be considered violent, the group should be associated with reoffending involving homicide and assault and/or the rare but extremely serious offences of homicide and wounding. A third model checks the prediction of reoffending for the resulting group of 'violent-type' offences.

When a final definition was agreed, PNC criminal history data were re-analysed to generate measures of previous sanctions for this group of offences.

Building logistic regression models to predict proven violent and non-violent reoffending

The statistical method used was logistic regression. Standard logistic regression models are used to predict a binary (yes/no) outcome, such as reoffending within a given timescale. It allows the independent contribution of many different variables to be assessed (that is, if an OASys question is correlated with reoffending but is also correlated with other OASys questions, the relationship with reoffending after controlling for the other questions can be estimated). As well as OASys questions, the OVP model included demographic data and separate counts of previous violent-type and non-violent sanctions (convictions and cautions). For OGP, the OGRS 3 score – a known excellent predictor of overall reoffending – was used as the basis for the static part of the model.

The logistic regression models are built with a forward stepwise method. The risk factors which best predict reoffending are added first, with other factors added until no further statistically significant improvement is possible. Some discretion was allowed in the model fitting process – for example, looking at whether individual questions or subsets of the standard OASys sections could be included. Without compromising the need for statistical significance, this was used to ensure that a relatively wide range of risk factors would be scored.

The version of the model which was eventually chosen was an ordinal logistic regression model. This uses a single set of risk factors to predict 'nested' outcomes – reoffending within the first month, within the second but not the first month, and so forth. For simplicity, only one- and two-year predictors are displayed here, but it is possible to produce predictions for anything from one to 24 months.

The predictive value of the tools is summarised with the Area Under Curve (AUC) score. The key strength of AUC is that it gives a single statistic which can be used to compare different

tools trying to predict the same proven reoffending outcome. This statistic ranges from 0 to 1 in theory with scores above 0.7 being generally good in practice, although what can be achieved depends on the sample being studied. An AUC of 0.5 is equivalent to tossing a coin, and an AUC of 1 implies perfect foresight. The AUC statistic can be understood with a real-world analogy – it is equal to the probability that a randomly chosen reconvicted offender will have a higher score on the predictor than a randomly chosen non-reconvicted offender. For example, an AUC of 0.65 implies that when 100 pairs of offenders are checked, the reconvicted offender will (on average) have a higher score on the predictor than the non-reconvicted offender in 65 cases. The AUCs of OGP and OVP were compared with those of other predictors.

Creating user-friendly versions of the new predictors

The logistic regression models for OGP and OVP would be suitable for use by researchers, and could be scored for practitioners through a behind-the-scenes calculation method like that of the OGRS IT application. However, this is not sufficient for an assessment with dynamic elements, which should allow practitioners to consider what steps can be taken to reduce the likelihood of reoffending. For this, something similar to the current OASys summary sheet is needed – that is, the offender’s score should be expressed in points, and these points should be visible to the practitioner and clearly divided between the various risk factors so that areas of strength and weakness can be identified.

Offender Assessment and Management Group colleagues agreed that total scores out of 100 should be produced, as this offers maximum clarity. The logistic regression model outputs were transformed into 100-point scales as described in Appendix 6.

The relationships between the score out of 100 with proven reoffending rates at one and two years were then calculated: this allows OASys IT to calculate the pair of predicted rates for every score out of 100. These rates are different for OGP and OVP, as non-violent reoffending is more frequent – the predicted rates for an OGP score of, say, 50/100 are higher than those for an OVP score of 50/100.

Comparing predictive validity of OGP, OVP and OGRS

In order to compare the predictive validity of the new predictors with that of other predictors, simulated scores on OGRS 2, OGRS 3 and OGRS-SV were produced. These could all be produced in full using data available in OASys and the PNC. Appendix 7 gives more details on the correlation between OASys total and section scores and the OGRS predictors.

As well as violent-type reoffending, non-violent reoffending and all reoffending, the outcome of “homicide and wounding” was examined. This comprises murder and related offences (attempts, conspiracy, threats to kill), non-vehicle manslaughter and Grievous Bodily Harm with intent (sometimes termed “wounding” to distinguish it from the more frequent and less serious offence of Grievous Bodily Harm without intent; see sections 18 and 20 of the Offences Against The Person Act 1861).

Pilot

The predictors, together with OGRS 3, were piloted with users in Nottinghamshire Probation Area in June 2007 and in Devon and Cornwall, Durham and Hertfordshire Probation Areas in early 2008. Probation officers were given case studies and asked for their opinions on the content of the predictors and their likely impact on all stages of the assessment and supervision process. Their responses were used to develop user guidance, with the help of a National Reference Group primarily comprising stakeholders from NOMS HQ and probation areas.

The national roll-out of OGP and OVP took place as part of OASys IT release 4.3.1, in August 2009. A programme of training and documentation accompanied the roll-out.

Results

Producing a definition of violent offending

Violent offence content

Table 6.2 provides a comparison of the Analysis of Offences content of various offence code groups. The 'homicide and assault' group were rated as violent on at least one of the three certain Analysis of Offences items in at least 80% of cases. The Analysis of Offences for weapon possession offences indicates that such cases frequently showed 2.2A (carrying/ use of weapon) ticked but not 2.2B, 2.2C or 2.3D (violence or threat of violence; excessive violence; violence towards partner). The 'public order' and 'criminal damage' groups are considered violent in around half of their cases. The classification of all such offences as part of the same group as 'homicide and assault' offences may be justified on this basis, and bears some similarity to the findings of Soothill, Francis and Fligelstone (2002), where two of the nine clusters of offence patterns derived from individuals' PNC conviction histories include 'general violence', encompassing possession of weapons and criminal damage, and 'wounding'.

Several other offence code groups also have some claim to be defined as violent on the basis of Table 6.2. Over 80% of robberies were violent in this sense, around 70% of aggravated burglaries and 60% of rapes. Child neglect offences less frequently included violent content, at around 40%, while arson and breach of Anti-Social Behaviour Order (ASBO) only did so in around 20% of cases.

Table 6.2: Indications of violent content in a range of offences

Offence code/ subcode	Offence description	OASys Offence Analysis item: % 'Yes'					
		2.2A	2.2B	2.2C	2.3D	Any of these	
Homicide and assault							
1	Murder	62%	81%	71%	20%	92%	90%
4/not 4 or 6	Manslaughter (not driving)	44%	75%	47%	18%	82%	80%
5	Wounding etc.	63%	84%	40%	10%	95%	88%
8/most	Grievous/actual bodily harm, racially aggravated offences	25%	83%	18%	17%	89%	86%
104	Assault on constable	11%	81%	7%	20%	84%	83%
105	Common assault	9%	85%	7%	41%	90%	89%
Threats and harassment							
3	Threats to kill	40%	92%	11%	31%	94%	93%
8/12-18, 8/23-25	Using firearms to commit or threaten assault	93%	59%	2%	4%	94%	59%
8/29-31, 195/94	Harassment offences (non-sexual)	6%	56%	1%	12%	58%	57%
Offences featuring weapon possession but often without other indicators of violence							
8/11-13,26-28	Carrying but not using a firearm/knife	91%	30%	1%	1%	92%	30%
81	Firearms Act 1968 and other related Acts	79%	33%	2%	2%	81%	33%
90	Knives Act 1997 and other related Acts	89%	26%	0%	2%	90%	27%
Child neglect							
11	Cruelty to and neglect of children	7%	38%	14%	4%	41%	40%
Violent acquisitive offences							
34	Robbery and assaults with intent to rob	43%	82%	8%	1%	87%	82%
29	Aggravated burglary in a dwelling	60%	68%	12%	4%	77%	69%
Criminal damage offences							
56	Arson	8%	15%	2%	5%	20%	17%
57	Criminal damage endangering life	17%	38%	3%	4%	47%	38%
58	Other criminal damage	16%	40%	1%	6%	48%	41%
Public order offences							

Table 6.2: Indications of violent content in a range of offences (continued)

Offence code/ subcode	Offence description	OASys Offence Analysis item: % 'Yes'					
		2.2A	2.2B	2.2C	2.3D	Any of these Any of 2.2B, 2.2C, 2.2D	
Public order offences							
65	Violent disorder	27%	82%	13%	1%	85%	83%
66	Affray and other public order offences	29%	78%	6%	5%	82%	79%
Sexual offences involving contact							
19	Rape	11%	58%	17%	12%	61%	60%
20	Indecent assault on female	2%	23%	2%	4%	25%	25%
Non-violent offences within violent offence codes							
8/4 and 8/6	Causing death by dangerous driving	1%	2%	1%	0%	3%	3%
8/32	Breach of an Anti-Social Behaviour Order	7%	20%	1%	1%	22%	20%

OASys Offence Analysis item codes: 2.2A Carrying/use of weapon; 2.2B Violence, threat or coercion; 2.2C Excessive or sadistic violence; and 2.3D Physical violence towards partner

Note: other non-scarce offences in the above groups, with similar profiles to the listed offences, include:

Homicide/assault – 2 (attempted murder); 3 (threats to kill); 6 (endangering rail passengers); 7 (endangering life at sea); 35 (blackmail); 36 (kidnapping); 103 (aggravated assault). Weapon possession – 115 (summary Firearms Act 1968); 181 (unlawful possession); 195/20 (manufacture etc. of offensive weapons). Public order – 64 (riot); 125 (summary offences against public order); 162 (disorderly behaviour); 195/94 (harassment). Criminal damage – 59 (threat and possession with intent to commit criminal damage); 149 (summary criminal damage). Child neglect – 109 (summary cruelty to and neglect of children). Non-violent offences within violent offence codes – 8/39 (breach of a restraining order (Sex Offenders Act 1997)).

Criminogenic need profiles

Table 6.3 compares the percentage of offenders with each criminogenic need among various offence groups. As well as the candidate violent offence groups included in Table 6.2, some definitely non-violent groups are included for comparison.

The different characteristics of robbers are clear – in fact, these offenders were far more similar to those convicted of other acquisitive offences than those convicted of non-acquisitive assaults, so robbery has, at this stage, only a limited claim to be a ‘violent-type’ offence. The same is true of aggravated burglary. Public order offenders were extremely similar to homicide and assault offenders. Criminal damage offenders were generally very similar to homicide and assault offenders, but those convicted of arson had much higher levels of emotional wellbeing need and considerably higher levels of other needs. Weapon-possession offenders were quite similar, apart from their levels of drug and alcohol misuse.

‘Threats to kill’ offenders and, to a lesser extent, harassment offenders, were somewhat different to those convicted of homicide and assault, with higher levels of relationship (threats to kill only), emotional wellbeing and thinking and behaviour needs.

Offenders convicted of robbery and aggravated burglary were quite different to those convicted of homicide and assault offences. They had higher levels of socio-economic and drug misuse needs, and appeared more similar to those convicted of other acquisitive offences.

Arsonists had a similar profile to those convicted of threats to kill, but given the lack of violent content shown in Table 6.2, arsonists were not considered to be ‘violent-type’ offenders.

Offenders convicted of rape were similar in some ways to homicide and assault offenders, but very different in that they had double the levels of relationship and attitude needs, and much lower levels of alcohol misuse need. Those convicted of child neglect or cruelty offences were still more different to other arguably ‘violent’ offenders while those convicted of breach of ASBO had very high levels of most needs. Those convicted of death by dangerous driving actually had fewer needs than other dangerous driving offenders – perhaps the instance of death caused a custodial or community sentence to be passed for offenders who would otherwise only require a low tariff – and were certainly unlike those convicted of other types of homicide or assault. None of these offences appears to fit within a needs-based ‘violent-type’ offence definition. However, sexual offences involving content were taken forward to the next step in order to ensure a thorough treatment of this offence group.

Table 6.3: Criminogenic needs by offence type: possible violent offences and comparison offences

Offence type	% with each need by OASys section											
	3	4	5	6	7	8	9	10	11	12		
Homicide and assault	32%	47%	12%	42%	26%	14%	58%	40%	60%	29%		
Threats and harassment												
Threats to kill	46%	56%	17%	62%	37%	21%	58%	65%	79%	44%		
Using firearms	32%	42%	12%	36%	36%	19%	42%	46%	51%	26%		
Harassment	34%	46%	11%	48%	29%	10%	45%	57%	72%	39%		
Offences featuring weapon possession but often without other indicators of violence												
Weapon possession	35%	60%	17%	36%	36%	26%	43%	45%	57%	33%		
Firearms Act	30%	46%	14%	28%	34%	18%	28%	34%	47%	28%		
Knives Act	38%	70%	17%	36%	38%	33%	51%	51%	61%	36%		
Violent acquisitive compared with other acquisitive profiles												
Robbery	48%	73%	39%	44%	63%	52%	40%	42%	66%	41%		
Aggravated burglary	42%	66%	32%	40%	57%	41%	45%	41%	64%	36%		
Non-aggravated burglary	50%	79%	43%	47%	64%	54%	32%	43%	66%	45%		
Theft and handling	42%	68%	41%	44%	52%	47%	26%	46%	56%	38%		
Criminal damage offences												
Arson	51%	62%	16%	53%	41%	16%	57%	70%	69%	35%		
Criminal damage endangering life	40%	59%	20%	42%	41%	19%	60%	48%	68%	42%		
Other criminal damage	39%	63%	19%	46%	38%	21%	66%	50%	69%	28%		
Public order offences												
Violent disorder	15%	33%	5%	16%	25%	7%	47%	15%	39%	18%		
Affray & other public order	27%	48%	12%	34%	29%	13%	58%	37%	56%	31%		
Rape and indecent assault compared with other sexual offences												
Rape	47%	49%	9%	58%	50%	11%	31%	45%	65%	53%		
Indecent assault	38%	42%	8%	46%	40%	7%	28%	49%	56%	45%		
Other contact sexual offences	41%	39%	7%	50%	48%	5%	19%	53%	52%	27%		
Non-contact sexual offences	21%	21%	7%	33%	33%	2%	11%	51%	35%	42%		

Table 6.3: Criminogenic needs by offence type: possible violent offences and comparison offences (continued)

Offence type	% with each need by OASys section											
	3	4	5	6	7	8	9	10	11	12		
Dangerous driving profiles												
Death by dangerous driving	15%	28%	6%	14%	20%	8%	19%	41%	27%	17%		
Dangerous driving	24%	50%	15%	27%	36%	18%	28%	33%	51%	31%		
Other profiles												
Child neglect	38%	60%	12%	69%	29%	13%	30%	70%	64%	33%		
Breach of ASBO	65%	85%	37%	53%	82%	41%	68%	54%	91%	72%		

OASys section codes: 3 Accommodation; 4 Education, training & employability; 5 Financial management & income; 6 Relationships; 7 Lifestyle & associates; 8 Drug misuse; 9 Alcohol misuse; 10 Emotional wellbeing; 11 Thinking & behaviour; 12 Attitudes

Associations between previous sanctions and reoffending

Table 6.4 presents the results of logistic regression models which predicted proven reoffending (conviction or caution) for homicide and wounding, all homicide and assaults and the final ‘violent-type’ offence group, including threats/harassment, criminal damage, public order, violent acquisitive (robbery and aggravated burglary) and weapon possession offences in addition to homicide and assaults.

The results of the models confirm that the ‘violent-type’ offence group should be expanded to include violent acquisitive and weapon possession offences. Previous sanctions for these offences were the strongest predictors of the most serious violent reoffending. Threats and harassment, criminal damage and public order sanctions all contributed to the prediction of all homicide and assault, though threats and harassment was not a significant predictor of homicide and wounding. Previous sanctions for all six elements of the ‘violent-type’ group helped to predict ‘violent-type’ reoffending.

The models confirm that contact sexual offences should be excluded from the definition. They did not contribute to the prediction of serious violence and were negative predictors of more broadly defined violent reoffending. Sanctions for all other offences (e.g. non-violent acquisitive offences, motoring offences, drugs offences) made small but significant positive contributions to the prediction of the two broader groups of violent reoffences.

Table 6.4: Logistic regression models to predict proven reoffending based on previous sanction counts

Previous sanction count	Outcome: proven reoffending within two years (% reoffending, Area Under Curve of model)		
	Homicide & wounding (0.7%, AUC = .67)	Homicide & assault (13.7%, AUC = .66)	Homicide & assault, threats & harassment, weapon possession, violent acquisitive, criminal damage & public order (25.3%, AUC = .68)
Constant	-5.364	-2.337	-1.690
Homicide & assault	0.196	0.196	0.166
Threats & harassment	NS	0.132	0.146
Weapon possession	0.633	0.201	0.204
Violent acquisitive	0.615	0.270	0.278
Criminal damage	0.114	0.175	0.254
Public order	0.063	0.188	0.299
Contact sexual	NS	-0.319	-0.175
All other offences	NS	0.012	0.021

NS = Not significant at $p = .05$

On the basis of these findings, the ‘violent-type’ offending to be predicted by OVP comprised homicide and assault, threats and harassment, violent acquisitive offences, weapon possession, criminal damage (with the exception of arson) and public order offences. OGP is therefore strictly defined as a predictor of all other offences – although it should not be relied upon as a predictor of rare, serious offences including sexual offences, terrorist offences, dangerous driving, child neglect and arson. The prediction of sexual reoffending is a complex topic, and the potential role of OASys in such prediction is therefore not considered further in this report.

Building and presenting logistic regression models

Appendix 6 includes the results of the initial logistic regression models. It also includes tables showing how the initial results were translated into scoring systems with maxima of 100.

Table 6.5 provides the final OGP scoring table, in a form which could be used to score an individual offender. The scoring of OGP is based heavily on OGRS 3, which contributes 60% of the scoring. Drug misuse is the most influential dynamic risk factor, with five other areas also contributing. The 100-point scoring system can be summarised as follows:

Criminal history, age and gender	60 points
Socio-economic and lifestyle issues	15 points
Drug misuse	15 points
Thinking and attitudes issues	10 points

Table 6.6 provides a guide to the proven reoffending rates associated with various OGP scores. Offenders were fairly evenly distributed between the different risk bands, with between 8 and 13% in all but one of the bands from “Below 10%” to “70 to 79%” (two-year rates). Many of the medium risk bands cover just six or seven points for each 10% change in proven reoffending rate, showing that it will be possible for offenders’ estimated likelihood of reoffending to change substantially in response to improving or worsening dynamic risk factors. The creation of score categories for use in Offender Management tiering, offending behaviour programme eligibility and management information has been developed, and in some cases already rolled out, in partnership with NOMS business teams.

Table 6.5: Scoring the OASys General reoffending Predictor (OGP)

Item	OASys items	Offender's score (max possible)	To get weighted score, map weighted unweighted as follows...	Offender's weighted score (max possible)
OGRS 3	n/a	(100)	Multiply by 0.6, round to nearest whole number	(60)
Accommodation	3.3 to 3.6	(8)	0→0; 1-2→1; 3→2; 4-5→3; 6-7→4; 8→5	(5)
Employability	4.2 to 4.5	(8)	0→0; 1-2→1; 3→2; 4-5→3; 6-7→4; 8→5	(5)
Regular activities encourage offending	7.2	(2)	0→0; 1→3; 2→5	(5)
Drug misuse (excludes drug-related violence)	8.4, 8.5, 8.6, 8.8, 8.9	(10)	0→0; 1→2; 2→3; 3→5; 4→6; 5→8; 6→9; 7→11; 8→12; 9→14; 10→15	(15)
Thinking and behaviour (excludes two violence-related questions)	11.1, 11.2, 11.5 to 11.10	(16)	0-1→0; 2-4→1; 5-7→2; 8-11→3; 12-14→4; 15-16→5	(5)
Attitudes	12.1, 12.3 to 12.6, 12.8	(12)	0-1→0; 2-3→1; 4-5→2; 6-7→3; 8-9→4; 10-12→5	(5)
Total score				(100)

Table 6.6: Likelihood of proven general reoffending by OGP score

Offender's weighted score (/100)	One year general reoffending rate (average for sample: 29%)	Two year general reoffending rate (average for sample: 39%)	% of sample
0 to 13	Below 6%	Below 10%	11
14 to 26	6 to 11%	10 to 19%	18
27 to 35	12 to 18%	20 to 29%	13
36 to 43	19 to 26%	30 to 39%	13
44 to 49	27 to 34%	40 to 49%	9
50 to 56	35 to 45%	50 to 59%	11
57 to 63	46 to 55%	60 to 69%	10
64 to 71	56 to 67%	70 to 79%	8
72 to 84	68 to 82%	80 to 89%	6
85 to 100	83% and over	90% and over	0.7

Tables 6.7 and 6.8 present OVP's scoring rules and proven reoffending rates. The scoring of static factors in OVP appears to place five times as much emphasis on 'violent-type' offences as other offences. However, as 'violent-type' offences are less frequent, many offenders will score more points from other offences, while having any previous criminal history raises risk considerably in itself. The scoring also highlights the much higher risk of violence among young male offenders. The total weight of static factors, at 60/100 points, is the same as in OGP – this was a deliberate design feature to help practitioners to get used to the tools,

and has little effect on predictive validity (see Table A6.5). Two new questions were added to OASys as part of release 4.3.1 and an earlier release which introduced OGRS 3. These allow OGRS 3, OGP and OVP to be scored. Question 1.24 records the number of cautions, reprimands and final warnings the offender has received. Question 1.26 records the number of previous sanctions which involved violent-type offences. (The current offence's violent content is checked automatically.)

Table 6.7: Scoring the OASys Violence Predictor (OVP)

Item	OASys items	Offender's score (max possible)	To get weighted score, map weighted → unweighted as follows...	Offender's weighted score (max possible)
Static factors				
Number of sanctioning occasions for violent-type offences	1.26 and current offence		None → 0; 1→4; 2→7; 3→9; 4→11; 5→12; 6→13; 7→14; 8→15; 9→16; 10→17; 11→18; 12→19; 13→20; 14→21; 15→22; 16→23; 17→24; 18+ →25	(25)
Number of sanctioning occasions which did not include violent-type offences	1.5, 1.6, 1.24 and current offence, minus 1.26		None, 1, 2 → 0; 3-4 → 2; 5-10 →3; 11-20→4; 21+→5	(5)
Is this the offender's first sanction ever?	1.5, 1.6 and 1.26	No/Yes	Score 5 if no i.e. previously convicted/cautioned etc.	(5)
Age at current conviction	n/a		Age 51+ → 0; 46-50 →2 41-45→4; 36-40→6 31-35→8; 26-30→10 24-25→12; 22-23→14 20-21→17; 18-19→20	(20)
Sex of offender	n/a	Female/ Male	Score 5 if male	(5)
Total weighted score from static factors				(60)
Total score				(100)
Dynamic factors				
Recognises impact of offending?	2.6	No/Yes	Score 4 if no i.e. does not recognise impact	(4)
Accommodation	3.3 to 3.6	(8)	0→0; 1-2→1; 3-4→2; 5-6→3; 7-8→4	(4)
Employability	4.2 to 4.5	(8)	0→0; 1→1; 2→2; 3-4→3; 5→4; 6→5; 7-8→6	(6)
Alcohol misuse (current use only)	9.1, 9.2	(4)	0→0; 1→3; 2→5; 3→8; 4→10	(10)
Current psychiatric treatment, or treatment pending	10.7	No/Yes	Score 4 if yes	(4)
Temper control	11.4	(2)	0→0; 1→3; 2→6	(6)
Attitudes	12.1, 12.3-12.6, 12.8	(12)	0→0; 1-2→1; 3-4→2; 5-6→3; 7-8→4; 9-10→5; 11-12→6	(6)
Total weighted score from dynamic factors				(40)
Total score				(100)

Table 6.8: Likelihood of proven violent-type reoffending by OVP score

Offender's weighted score (/100)	One year violent-type reoffending rate (average for sample: 17%)	Two year violent-type reoffending rate (average for sample: 26%)	% of sample
0 to 22	Up to 5%	Below 10%	14%
23 to 34	5 to 11%	10 to 19%	29%
35 to 41	12 to 17%	20 to 29%	21%
42 to 47	18 to 26%	30 to 39%	15%
48 to 53	27 to 33%	40 to 49%	10%
54 to 59	34 to 43%	50 to 59%	6%
60 to 65	44 to 54%	60 to 69%	2.8%
66 to 72	55 to 66%	70 to 79%	1.1%
73 to 100	67% and over	80% and over	0.4%

Many of the dynamic risk factors scored in OVP are similar to those scored in OGP. However, alcohol misuse is an important part of OVP whereas drug misuse and lifestyle are not scored, and recognising the impact of offending is only scored in OVP. Also, OVP gives particular emphasis to question 11.4, on temper control, in section 11 (Thinking and behaviour). The 100-point scoring system can be summarised as:

Criminal history	35 points
Age and gender	25 points
Thinking, mental health and attitudes issues	20 points
Socio-economic issues	10 points
Alcohol misuse	10 points

Table 6.8 shows that the risk distribution was slightly skewed towards lower risks, with 43% of the sample having an estimated likelihood of 'violent-type' reoffending below 20% while the overall rate was 26%. Only 10% of the sample were more likely than not to be sanctioned for any 'violent-type' offence, and very few could be identified as almost certain reoffenders.

Comparing predictive validity of OGP, OVP and other risk assessment tools

Table 6.9 uses AUC statistics to compare the predictive validity of the wide range of existing and simulated assessment tools. OGP and OVP are shown to have the greatest predictive validity: OGP is the best predictor of non-OVP offences and all offences, while OVP is the best predictor of all violent offences and the most serious violent offences. There are, however, some important caveats about particular violent offences.

In the prediction of 'violent-type' offences, OVP represents a considerable improvement on all other tools, with OGRS 3 and OGRS-SV generally the next best predictors. For all OVP-type offences, the AUC of .74 is a substantial improvement upon OGRS 3's AUC of .70 and the current OASys score's AUC of .68. (Using a coin-toss or some other completely

random process to assess offenders will yield an AUC of .5, so using OVP improves on guesswork by one-third more than using the current score i.e. by .24 rather than .18.) Its margin of superiority was similar when predicting homicide and wounding, with an AUC of .74 compared with .68 for OGRS 3 or the current score. Thinking and behaviour was generally the most predictive OASys section.

Violent acquisitive and weapon possession reoffending appear to be predicted better by OGP than OVP. The margin is small for violent acquisitive offences (AUC of .80 for OGP and .79 for OVP). In fact, the sum of OGP and OVP (i.e. a score out of 200) has an AUC of .82, showing that violent- acquisitive reoffending is most likely when both scores are high. However, weapon possession offences, which are strongly correlated with lifestyle and drug misuse and weakly correlated with alcohol misuse, are definitely better predicted by OGP (.76) than OVP (.70), and a combined score offers no benefit here.

It should be reiterated that weapon possession is included in OVP because previous weapon possession offences are highly predictive of homicide and wounding reoffending. Additional analyses (not shown) suggest that homicide and wounding reoffending is more likely, controlling for OVP score, when the current offence involves weapon possession or use (OASys question 2.2A) and also when it involves excessive or sadistic violence (question 2.2C). However, these patterns do not hold for all violent-type offending: indeed, of all 'violent-type' reoffending, excessive/sadistic violence in the current offence is a negative predictor, suggesting that such offenders are prone to extreme but infrequent violence.

OGP was the best predictor of offences not covered by OVP – predominately acquisitive and motoring offences. The AUCs for these offences are generally much higher, suggesting that prediction of these offences is usually more reliable than prediction of 'violent-type' offences. OGRS 3 predicts non-OVP offences well, and OGP improves on guesswork by 2 points more than OGRS 3 and 4 points more than the current OASys score (AUCs of .80, .78 and .76 respectively). Lifestyle and associates, education, training and employability and drug misuse are the best predictors among the OASys sections. Looking at the different columns of Table 6.9, drug misuse is generally a strong predictor of OGP offences, weapons and violent acquisitive offences but a weak predictor of other violent offences, while the opposite is true of alcohol misuse.

Table 6.9: Predicting proven reoffending using existing and new measures: Area Under Curve scores

Predictor	Measure (24 month reoffending rate)									
	Homicide & assault (14.7%)	Threats & harassment (1.0%)	Weapon possession (1.9%)	Violent acquisitive (0.9%)	Criminal damage (8.5%)	Public order (10.7%)	All violent-type offences (26.5%)	Homicide & wounding (0.9%)	OGP offences (39.7%)	All (49.1%)
Existing OASys-based predictors										
Total weighted score	.66	.58	.74	.79	.66	.67	.68	.68	.76	.75
Sections 3–12 weighted score	.64	.58	.72	.78	.65	.66	.66	.65	.73	.73
Unweighted section scores										
Accommodation	.57	.52	.63	.63	.56	.58	.58	.59	.61	.61
Education, training & employability	.61	.54	.67	.75	.63	.63	.63	.62	.69	.68
Financial mgmt & income	.55	.51	.63	.70	.58	.57	.58	.53	.66	.63
Relationships	.59	.58	.59	.56	.60	.59	.60	.57	.60	.61
Lifestyle & associates	.61	.57	.70	.76	.62	.63	.64	.65	.70	.70
Drug misuse	.55	.51	.66	.67	.56	.55	.56	.59	.68	.66
Alcohol misuse	.60	.62	.55	.57	.63	.66	.62	.59	.50	.54
Emotional wellbeing	.55	.53	.56	.52	.56	.57	.55	.51	.55	.56
Thinking & behaviour	.64	.62	.66	.72	.64	.65	.65	.68	.65	.66
Attitudes	.61	.59	.65	.68	.61	.62	.62	.66	.66	.66

Table 6.9: Predicting proven reoffending using existing and new measures: Area Under Curve scores (continued)

Predictor	Measure (24 month reoffending rate)									
	Homicide & assault (14.7%)	Threats & harassment (1.0%)	Weapon possession (1.9%)	Violent acquisitive (0.9%)	Criminal damage (8.5%)	Public order (10.7%)	All violent-type offences (OVP) (26.5%)	Homicide & wounding (0.9%)	OGP offences (39.7%)	All (49.1%)
Other existing NOMS predictors										
OGRS 2	.66	.52	.73	.78	.67	.68	.68	.68	.78	.77
OGRS-SV (sexual/violent)	.69	.56	.71	.76	.69	.70	.70	.72	.72	.73
OGRS 3	.67	.53	.73	.78	.69	.70	.70	.68	.78	.78
New OASys-based predictors										
OASys General reoff. Predictor (OGP)	.67	.56	.76	.80	.68	.70	.70	.67	.80	.79
OGP dynamic component	.62	.56	.73	.77	.63	.63	.64	.69	.75	.73
OASys Violence Predictor (OVP)	.71	.65	.70	.79	.74	.76	.74	.74	.68	.72
OVP static component	.69	.59	.68	.79	.72	.73	.72	.75	.67	.71
OVP dynamic component	.65	.66	.65	.69	.65	.70	.67	.65	.62	.65

Key: Best three predictors of each outcome (there were some third-place ties). Worst three predictors of each outcome.

Figures 6.1 and 6.2 highlight the improvements brought about by OGP and OVP. Figure 6.1 compares OGP with the current OASys score and OGRS 3, which represent the status quo and the strongest available alternative respectively. Figure 6.2 compares OVP with these predictors.

In these graphs, offenders are ranked on each predictor and divided into deciles (ten equal-sized groups). For example, in Figure 6.1 the group of offenders who had the lowest 10% of OASys scores had scores of 0 to 11 and had an 11% proven reoffending rate, while another group of offenders had the lowest 10% of OGRS 3 scores (of no more than 15) and had a 7% proven reoffending rate.

An assessment with good predictive validity would have very low reoffending rates in the leftmost (lowest-scoring) deciles and very high rates on the rightmost (highest-scoring) deciles. This would show that the predictor could successfully identify offenders very unlikely and very likely to reoffend. Ideally, there would be little or no middle ground – if few deciles have an intermediate reoffending rate, this indicates that most offenders can be identified as either high or low risk rather than being “left behind” as a medium risk.

Figure 6.1 confirms that OGP is a moderately superior predictor to OGRS 3. It has a slightly lower reoffending rate for the bottom four deciles, and a higher rate for the top four deciles. Both OGRS and OGP distinguish level of risk at a fine level of detail, with clear differences in rate between every pair of successive deciles. The current OASys score does less well: the lowest decile reoffends at a similar rate to the second decile, and the difference between the third and ninth deciles is only 44% (from 22% to 66%) compared to 53% (18% to 71%) for OGP.

Figure 6.2 shows the clear superiority of OVP for ‘violent-type’ offences. In particular, the top decile has a very high reoffending rate: 61%, compared with below 50% for the other two predictors. There are almost as many reoffenders in that single OVP decile as in the first five OVP deciles combined (i.e. the one-half of the sample who score up to 37/100; their reoffending rates sum to 66%). In this way, OVP identifies a high risk group who should be the focus of considerable attention from Offender Managers and a priority for relevant offending behaviour predictors.

Despite this success, a problem area for OVP is the middle (fourth to sixth) deciles – these comprise a large group of offenders whose reoffending rate is below the 26% average but is far from negligible. Further research will examine these offenders and determine whether there are any specific risk factors which might help to separate them into lower- and higher-risk groups.

Figure 6.1: Proven reoffending for offences not covered by OVP, by OASys, OGRS 3 and OGP decile

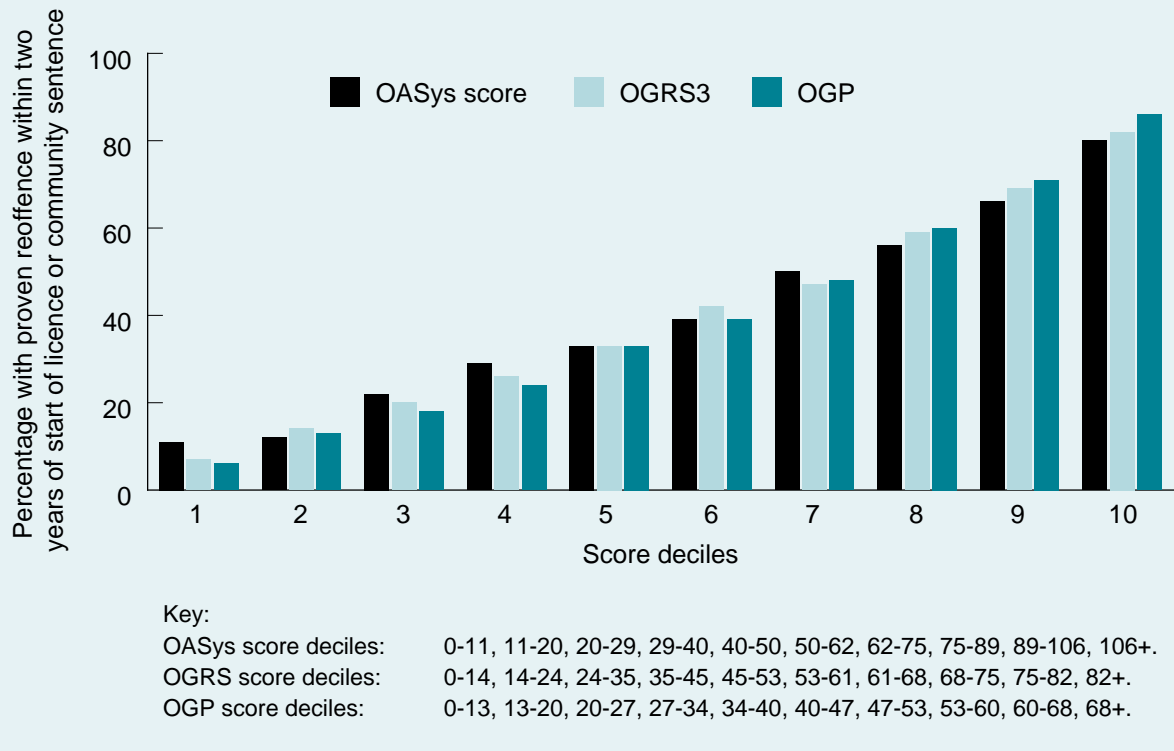
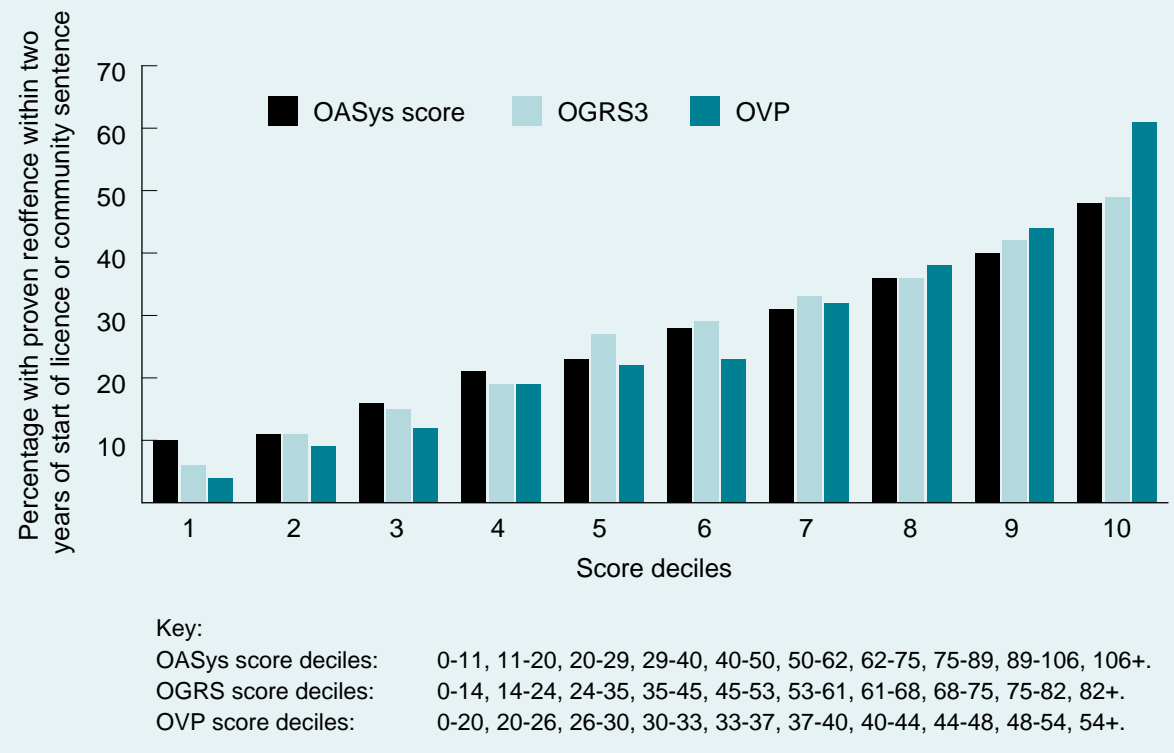


Figure 6.2: Proven reoffending for offences covered by OVP, by OASys, OGRS 3 and OVP score deciles



Diversity issues: accurate and valid prediction by age, gender and ethnicity

It is important that any risk predictors in use are accurate (do not systematically over- or under-estimate reoffending) and valid (successfully predict which individual offenders are more and less likely to reoffend) for offenders of all ages and ethnicity and for both male and female offenders. Appendix 8 presents predicted and actual proven reoffending rates (for accuracy) and AUCs (for validity) for each group of offenders for the groups of offences predicted by OGP and OVP offences. Table 6.10 provides a summary of these results. OGP and OVP appear to work very well for all groups, but additional reassurance could be gained by retesting their validity with larger samples of Minority Ethnic offenders. It may be possible to gain additional insights into the reoffending behaviour of specific groups (e.g. offenders of Black ethnicity, or female offenders) by building statistical models for those particular groups. This research will be conducted in 2009/10.

Table 6.10: Accuracy and validity of OGRS 2 and 3, OASys score and OGP/OVP by age, gender and ethnicity

Offender characteristic	Results for OGP offences	Results for OVP offences (violence against the person, threats/harassment, weapon possession, violent acquisitive, criminal damage, public order)
Age	Actual rates range from 50% at age 18–19 to 12% at age 51+. OASys score is not age-adjusted and produces large inaccuracies (it scales to predictions of 40% at 18–19 and 28% at 51+). OGRS 3 is more accurate than OGRS 2 for some groups. OGP has only slight inaccuracies: it is no more than 2% out for any age group. OGP is the most valid predictor for all age groups.	Actual rates range from 42% at age 18-19 to 8% at age 51+. All predictors except OVP underestimate reoffending by 18–19-year-olds by at least 9%. As with OGP offences, OASys scores produce large inaccuracies for the youngest and oldest offenders. OVP is the most valid predictor for all age groups.
Gender	Actual rates are 37% for women and 40% for men. OGRS 2, though not OGRS 3, underpredicts by 5% for women. All predictors produce high AUCs for women, so can distinguish well between women with lower and higher likelihoods of proven reoffending. OGP is the most valid predictor for both women and men.	Actual rates are 18% for women and 28% for men. All predictors except OVP overestimate reoffending by women by at least 4% (in practice, assessors are likely to use these predictors differently for men and women as they know women’s offending is less violent, but such adjustments are unlikely to be consistently valid). All predictors score high AUCs for women. OVP is the most valid predictor for both women and men.

Table 6.10: Accuracy and validity of OGRS 2 and 3, OASys score and OGP/OVP by age, gender and ethnicity (continued)

Offender characteristic	Results for OGP offences	Results for OVP offences (violence against the person, threats/harassment, weapon possession, violent acquisitive, criminal damage, public order)
Ethnicity	Actual rates are low for offenders of Asian origin (37%) and very low for those with unrecorded ethnicity (32%): non-recording may be more likely when the offender is a low risk and thus low priority. Reoffending by offenders of Black and 'other' ethnicity is underestimated by a large margin by all predictors – this may be random error on small samples of these offenders, but will be revisited on larger samples in the future. OGP is the most valid predictor for all ethnic groups.	Actual rates are low for offenders of Asian, Black and 'other' ethnicity (22-23%) though not offenders of Mixed Ethnicity (33%), whose rates are underestimated – again, this may be random error given the small sample available. OVP is the most valid predictor for all ethnic groups.

Implications

The results in Table 6.9 show very clearly that OGP and OVP represent a substantial improvement in the level of predictive validity which can be achieved within OASys. It was recognised that they should be implemented as soon as was practical, replacing the existing OASys scoring system, and this took place in OASys IT release 4.3.1, in August 2009.

The pilot studies gave valuable insights into practitioners' needs and preferences, and the guidance documents provided in the pilot studies have been revised. Members of the OGP/OVP National Reference Group recommended that all staff be provided with a document of two or three pages, which focuses on practical actions required when using the new tools and includes screenshots from the updated IT system. A longer document should provide full details of the mechanics of the scoring system, and explain matters such as the division of offences between OGP and OVP. Complex issues such as the relationship between weapon possession and violent and non-violent reoffending should be explained carefully, at the appropriate level of detail for each document. These recommendations were acted upon when release 4.3.1 was launched, and extensive user guidance is now available on internal staff websites.

OVP and risk of serious harm assessment

OVP, as a violence predictor, is especially relevant to NOMS Public Protection and Mental Health Group. On the basis of these results, it has been recommended to the NOMS Public Protection Board that OVP and OGRS 3 (when OASys is not used) should be used to estimate likelihood of future violence, whereas the 'V' scale of Risk Matrix 2000 was previously under consideration.

User guidance is likely to involve the use of OVP scores in risk of serious harm screening – a high score would suggest that full analysis is always necessary, a low score that it should only be necessary in the presence of other risks (e.g. sexual offending; domestic violence; risk to children; arson), and a medium score leaving the decision to the assessor’s discretion.

Exploratory O-DEAT research shows that, at present, assessed risk of serious harm levels vary widely between probation areas even for particular groups of offenders (e.g. for sexual offenders only; for violent offenders only), and suggests that wide geographic variation remains after controlling for all recorded risk factors. This inconsistency distorts decision-making in prisons (who will often hold offenders assessed at Pre-Sentence Report by a number of areas) and makes fair resource allocation more difficult to achieve. Using OVP as an objective basis for risk of serious harm assessment for non-sexual offenders should promote greater consistency and fairness. This should also promote consistency in the use of Offender Management Tiers, as an offender’s risk of serious harm level is one of the main determinants of their tier. Tiers describe the intensity of contact between the offender and their manager, using the What Works principle that ‘resources follow risk’.

The balance of static and dynamic risk factors, and measuring change

OGP and OVP each include several dynamic risk factors. Identifying these in the predictors will help in identifying key treatment targets for interventions.

While the weighting systems for both predictors (Tables 6.5 and 6.7) emphasise static factors, the remaining 40 points for dynamic factors still allow a great deal of change. For example, an offender who initially scored a maximum 15 weighted points for drug misuse as part of a total OGP score of 70, but then ceased drug misuse, would see their drug misuse and total scores drop by 13.²³ An OGP score of 70 translates to a 77% likelihood of OGP-type reoffending within two years, whereas an OGP score of 57 gives a 61% likelihood. If ceasing drug misuse also reduced dynamic scores by five points across, say, “regular activities encourage offending” and attitudes, the new OGP score of 52 would give a 53% likelihood. These changes could also happen in reverse – an offender who relapsed into drug misuse, or lost employment (especially through their own failings) or decent accommodation could gain many points and move into higher risk categories. Changes in likelihood will affect prioritisation for interventions and Offender Management tier level (influencing the extent of contact between offender and offender manager). Large changes in OVP scores will also trigger changes in risk of serious harm levels, especially where violent offending is the primary area of concern.

OGP and OVP therefore have the potential to reflect changes in dynamic risk factors while still providing robust prediction of reoffending. However, this potential will only be realised if OASys assessments are reviewed diligently. Current National Standards specify

²³ It would not drop by 15, as the offender would still score 1 unweighted point (2 weighted points) for their past drug-injecting behaviour.

that reviews, with rescoring where appropriate, should be completed no later than every four months during probation supervision and annually in custody (National Offender Management Service, 2007a, pg. 25). Separate further analysis is needed to explore whether these standards are being met.

The role of 'thinking and behaviour' in reoffending

Comparing Tables 6.5 and 6.7 with Table A7.2 shows that OASys section 11 – thinking and behaviour – scores are strongly correlated with reoffending yet have small weightings in both OGP and OVP. (In OVP, this is concentrated in the temper control question, while in OGP it is spread across eight questions.) Why is this, and what does it imply for treatment of 'thinking and behaviour' and other needs?

The answer lies in the logistic regression models behind both OGP and OVP. In both models, thinking and behaviour was highly weighted when the model included few risk factors. However, it became less important as more risk factors (e.g. employability, regular activities encourage offending, anti-social attitudes) were added. Table A7.2 shows that the levels of these risk factors are well correlated with the level of thinking and behaviour need. Put together, this implies (if one accepts that cognitive problems underlie these other problems) that the focus of the model shifts from the underlying problems to their pro-criminal effects such as poor social functioning and substance misuse.

In other words, the patterns of reoffending in our sample suggest that 'thinking and behaviour' problems mainly influence offending behaviour indirectly, through these other risk factors, rather than directly. Interventions to deal with these other factors are therefore most likely to succeed if they address underlying cognitive problems. This is the method used by the existing NOMS accredited behaviour programmes, including violence (Aggression Replacement Therapy, Controlling Anger and Learning to Manage it) and substance abuse (Addressing Substance Related Offending, Offender Substance Abuse Programme) programmes. Likewise, general offending behaviour programmes (such as Enhanced Thinking Skills) should help offenders address a range of dynamic risk factors. While OGP and OVP give only a small direct weighting to thinking and behaviour, tackling cognitive need is likely to significantly affect likelihood of reconviction through its effect on these other risk factors. Programme targeting criteria should be amended to use OVP (violence programmes) or OGP (illegal substance abuse and general programmes) as appropriate.

The role of substance misuse in reoffending

OGP and OVP both give considerable weight to substance misuse, but in different ways. OGP gives a maximum of 15 points for drug misuse, while OVP gives a maximum of ten points for alcohol misuse. (Because of the greater prevalence of alcohol misuse – see Table A7.1 – the average weighted score for alcohol misuse in OVP is actually higher than that for drug misuse in OGP.)

Not all existing tools used in violence and other criminal risk assessment make a distinction between drug and alcohol misuse. For example, item H5 of HCR-20 (Webster *et al.* 1997) is simply entitled Substance Use Problems, while LSI-R (Andrews and Bonta, 1995) scores both alcohol and drug misuse and (like OASys until now) does not distinguish their effects on different types of reoffending. However, VRAG (Quinsey *et al.* 1998) item 3 scores the offender's history of alcohol problems, which should improve its predictive validity for violence compared with tools such as HCR-20.

The different effects of drug and alcohol misuse demonstrated here are broadly supported by the literature on substance misuse and violence, as reviewed by Boles and Miotto (2003). Strong evidence links alcohol intoxication with disinhibition of fear. Alcohol intoxication also triggers aggressive acts among those who already have some propensity towards violence, when they are in the kind of situations where aggression is likely to be provoked. Chronic drinking can lead to interpersonal difficulties and personality changes which heighten the risk of conflict with others. By contrast, Boles and Miotto (2003) find that not all illegal drugs have such pharmacological effects. Given that 2002-04 OASys assessments for drug misusing offenders often lacked information on individual drugs which should be completed in the matrix part of question 8.1, distinguishing between particular drugs has not been possible in this study and would put the reliability of OVP's scoring system at risk. However, evidence that some drugs – amphetamines and methamphetamines, cocaine and PCP – can be pharmacologically linked with violence, while other drugs can be associated with violence through drug distribution networks and acquisition of money to buy drugs, can be cited in user guidance.

Content of OASys

Other research included in this compendium examines the relevance and usefulness of each part of OASys. OGP and OVP use items from most of the offending-related factors component, but the finance, relationships and emotional wellbeing sections are unused. Findings from the other studies and from textual analysis of assessor entries, were combined with OGP/OVP results to yield recommendations on OASys content.

Research implications

OGP and OVP will enable more effective research on offenders. Studies evaluating interventions designed to reduce reoffending, rely on accurate estimation of the likelihood of reconviction. Improved estimates mean that control groups can be more closely matched to experimental groups and sampling strategies should draw more representative samples of offenders, while regression-based methods such as propensity score matching should weight more relevant factors.

This is the first study to validate OGP and OVP, and as such their utility is only proven for adults in England and Wales. It is probable that the risk factors identified may be relevant for juvenile offenders or adults in other jurisdictions, but the predictors would need to be validated and recalibrated for the relevant group of offenders before being put into practice.

Variation between probation areas

The validity of the predictors may vary between probation areas. This was not examined in this study, as OASys was not fully implemented in early-mid 2004, when the current sample was drawn. Further studies will examine larger and more representative samples, and check for associations between proven reoffending and measures of police and probation performance.

Prediction of reoffending involving domestic violence

It is not clear whether OVP's strong predictive validity extends to domestic violence, where situational factors and offender characteristics may be different. This affects risk of serious harm assessment and targeting to the Integrated Domestic Abuse Programme and Community Domestic Violence Programme interventions. O-DEAT is investigating possible sources of data which would make it possible to study domestic violence reoffending further.

Prediction of sexual reoffending

It may be possible to improve the prediction of sexual reoffending using OASys. Studies on the predictive validity of Risk Matrix 2000, OGRS 3 and OASys are now underway. It will be possible to include OVP in this work.

Conclusion

The predictive validity of OASys scores has been improved considerably by the development of the OASys Violence Predictor and OASys General reoffending Predictor, which improve on the predictive value of the Offender Group Reconviction Scale and the existing OASys score. The new scores make better predictions while continuing to include dynamic risk factors which are amenable to intervention and may change over time. OGP and OVP provide valid indications of future reoffending, which adds further confidence to the use of OASys as a valuable assessment of offenders' risks and needs, and a useful tool for supporting effective offender management.

7. Measuring changes in risk and need over time using OASys

Introduction

Within NOMS, supervision requirements, programmes and interventions are often applied with the intent of effecting some change in the offending-related needs identified by OASys. Some changes in offending-related needs would therefore be expected when offenders have engaged with such rehabilitative requirements and interventions. It is also possible that offenders' needs will change autonomously as a result of life events outside the direct control of the correctional services. Both types of change should be reflected in these offenders' OASys dynamic need scores. O-DEAT's dynamic validity research examines repeated assessments of the same offender, in order to identify the extent of changes in scores, and the relationship between changes in scores and reoffending. The validity of OASys as a measure of change is of value both to practitioners working with individual offenders and researchers involved in evaluating the outcomes of interventions.

OASys was designed to be administered repeatedly. Offenders should be assessed after conviction but before sentence if they present sufficient indications of risk to receive a Standard Delivery Pre-Sentence Report (PSR). Post-sentence, a full assessment should be completed in the community for all those cases designated at Offender Management Tier 2 and above, with the exception of those Tier 2 cases in which there is a stand-alone unpaid work requirement.²⁴ In the prison establishments, all offenders aged 18–20 and older offenders serving a custodial sentence of at least 12 months should be assessed (National Offender Management Service, 2007a). If the offender was assessed at the PSR stage, then the assessment should be reviewed; otherwise a new assessment is completed. This assessment should then be reviewed at least annually in custody, and within 16-week periods in the community. A termination assessment should be completed at the end of the order or licence.

At each stage, judgements made at the previous assessment should be thoroughly reviewed, updating the 73 scored questions used to measure criminogenic need and likelihood of reconviction, several unscored questions on issues such as domestic violence and psychiatric problems, and indicators and levels of risk of serious harm. Such revisions are required in order to ensure that the assessment continues to assist practitioners in deciding upon the level and type of supervision required and interventions which should be undertaken. The OASys IT system allows previous assessments to be 'cloned', enabling review assessments to be created quickly. The system does not impose an absolute requirement to reassess any offending-related factors or the risk of serious harm.

24 Probation Circular 08/2008 sets out the Offender Management Tiers and how they are to be applied (National Offender Management Service, 2008a). The four tiers represent levels of intervention, with the approach increasing in scale and complexity as the risks and needs of the offender, and the demands of the sentence, increase. An offender's tier level is thus dependent upon his/her likelihood of reconviction, risk of serious harm and various other factors, e.g. classification as a Prolific and other Priority Offender (PPO).

Valid dynamic questions are important as they maintain the predictive validity of risk scores over time and allow the success or failure of offender management, interventions and the offender's own efforts to change to be demonstrated. Review assessments which indicate no change can legitimately occur. However, with 73 scored items (62 excluding criminal history), with most offenders having multiple criminogenic needs, and a period of several months between reviews, it would be surprising if no change occurred in a high proportion of cases. The continued use of those questions whose scores seldom change will need to be considered carefully, taking into account other research findings on their reliability and validity.²⁵

This study sought to address the following research questions:

- 1) How often are assessments being reviewed?
- 2) Do the reviewed assessments demonstrate change?
- 3) Which areas of the assessments change most?
- 4) Are the changes predictive of reoffending?

As well as assessing the extent of change, the study analysed the relationship between changes in OASys score and reconviction. The full potential of the new predictors of reoffending (see Chapter 6) will only be realised if they are sufficiently dynamic that changes in the predictor score are correlated with reoffending.

Method

Sample

The database held by O-DEAT contains around three million completed OASys assessments (as of February 2009). Well over half a million individuals have been assessed at least once. Given the greater volume of assessments in the community, and the longer intervals between assessments in custody, this study was restricted to assessments completed by probation areas.

To allow a sequence of review assessments to occur, the sample for a 'change' analysis needed to allow a substantial period of time to elapse from the start of the order or licence. The sample also needed to be as representative of the probation caseload as possible. The electronic version of OASys was rolled out across the National Probation Service during 2003 and 2004. As several larger probation areas were among the late adopters, a representative sample could not be taken until the second half of 2004. This still allowed offenders to be tracked for two years following the initial assessment.

The following 'initial' assessments completed between July and December 2004 were extracted:

- PSR assessments resulting in a community order;
- start of community order assessments (where a PSR had not been completed); and

²⁵ See other chapters in this compendium.

- start of post-custodial licence assessments.²⁶

The assessments were then filtered to ensure that they satisfied O-DEAT's quality standards. These standards require that each dynamic risk factor is well assessed, with at least four-fifths of scored questions answered, and that there is basic consistency and completeness of the risk of serious harm sections. Sentence details were also required when the assessment was completed at the PSR stage.

The 17,824 offenders included in the final sample had the following characteristics:

- 16% were aged 18–20; 19% were aged 21–24; 49% were aged 25–40; and 15% aged 41+.
- 15% were female;
- 8% were non-White and a further 8% had no ethnicity data;
- 22% were convicted of violence against the person; 18% theft and handling; 9% burglary; 8% drugs offences; 4% robbery; 3% sexual offences; 3% fraud and forgery; 3% criminal damage; and 30% other offences (of which about two-thirds were motoring-related); and
- 24% were sentenced to custody.

Counting OASys reviews

A **set** of assessments was created for each offender, comprising their initial assessment and all subsequent assessments within 24 months of the initial assessment. The number of assessments in each set was counted, separating offenders who were reconvicted and those who were not reconvicted (as reconviction would often cause a further PSR). Reconviction was measured by a proxy method, looking at data in the second and subsequent assessments. A reconviction was assumed to have occurred if the sentence date or principal offence changed, or if the previous conviction count in questions 1.5 and 1.6 of the core OASys assessment had increased.²⁷

Ideally, each offender would have been studied over the interval from their initial assessment to their termination assessment. According to National Standards, such an approach should have been possible, with every assessed offender requiring a termination assessment at the end of their sentence. In practice, however, many periods of supervision concluded with no termination assessment. Others were curtailed when reconviction led to a new sentence,

²⁶ It was necessary to include both PSR and 'start' assessments for those receiving community orders. Many offenders who were assessed at PSR stage had no 'start' assessment even though they had some later assessments in the same set/sequence. Excluding PSR assessments would therefore have removed a substantial proportion of offenders from the sample and potentially introduced bias into the results. Assessments for offenders serving Community Punishment Orders, Suspended Sentence Orders and other orders where the use of OASys was optional were excluded (see NOMS [2007a] for the current standards on OASys use, which now require OASys for many offenders serving the revised Suspended Sentence Orders introduced under the Criminal Justice Act 2003).

²⁷ While it is possible that an offender could be removed from the OASys process by reconviction, it is unlikely. Even if the offender was sentenced to a short custodial sentence, where OASys is not used, they should have a PSR before receiving that sentence.

without a termination assessment marking the end of the previous sentence. Having studied assessment patterns, rules were devised which would allow meaningful comparisons to be drawn. It was determined that a **sequence** of assessments should run from the initial assessment to whichever of the following occurred first:

- the first assessment in which reconviction was indicated; or
- the first termination assessment; or
- the final assessment in the set (i.e. the last within the two-year follow-up).

The number of sequences where change could be studied was 16,222.

Measuring changes in OASys

Comparisons between the initial and final assessments in each sequence included analysis of the following:

- changes in individual items;
- changes in section scores;
- changes in the criminogenic need prevalence rates and the number of needs;
- changes in the dynamic component of the OASys likelihood of reconviction score;
- changes in the dynamic components of the new OASys General reoffending Predictor (OGP) and OASys Violence Predictor (OVP) (see Chapter 6);²⁸ and
- changes in the risk of serious harm ratings.

Net and absolute changes were both studied, the latter enabling components where change seldom occurred to be distinguished from more dynamic components where increased scores for some offenders were balanced by decreases for others.

Bearing in mind that some reviews could have been completed in a perfunctory manner where a more thorough review might have revealed some change, a zero-inflated Poisson regression model was fitted. When the number of scored item changes was zero, this model estimated the probability that the observation came from an 'always zero' group, that is, a fully automated review, rather than being a considered review which happened to have zero score changes. The estimate was then used to weight zero-change assessments in the remaining results.

²⁸ Total OGP and OVP scores are used to produce percentile predictions of reoffending.

Predicting reoffending using OASys score changes

While those reconvicted and not reconvicted at the second assessment in each sequence could have been compared, any score changes might have been a result of the reconviction (e.g. if analysis of the new offence revealed problems which the assessor had previously under-estimated) rather than predictive of that reconviction. Therefore, changes between first and second assessments were used to predict whether or not reoffending had occurred at the third assessment. In adopting this approach, the sample was limited as follows.

- There had to be at least three assessments within a sequence.
- The offender must not have been reconvicted at the second assessment.
- It had to be possible to check the second assessment's review plan to see if a new criminal charge was pending.
- There must have been some change between the first and second assessment.

Fulfilling these conditions limited the available sample for this stage of the study to 1,862 sequences of assessments.²⁹

Limitations

It should be noted that the OASys caseload is not representative of the wider probation caseload. Differences between those assessed and not assessed (see Chapter 9) demonstrate that OASys tends to be completed for more dangerous offenders with higher levels of need. The sample for predicting reoffending was particularly small, making it difficult to interpret any differences between subgroups. In particular, very few reconvicted offenders were from Minority Ethnic groups, and analyses by gender and sentence were also compromised by small samples.³⁰ The offenders covered by the 1,862 sequences of assessments are not likely to be entirely representative of the whole sample, although their average OGRS 2 score³¹ of 48 compared with an average for the whole sample of 47.

The method for measuring sequences of assessments meant that those offenders with reconvictions or termination assessments were followed for shorter time periods than those experiencing neither event. The latter group, therefore, had a longer time period over which their risks and needs could change. However, measuring change over a fixed time period proved impractical, as the irregularity of actual assessment intervals (despite the regular process set down in National Standards) rendered the interpretation of such results difficult.

29 Over half of the 3,982 sequences of assessments which met the first three conditions failed the final condition.

30 The small number of reconvicted offenders and incomplete information on the types of new offences committed made it impractical to test whether changes in OVP predicted future violence; therefore, the standard OASys score is reported but not OGP or OVP. It would be possible to repeat the analysis at a later date in order to obtain a large enough sample to test variation in OVP scores between those whose reoffence at third assessment was violent, those whose reoffence was non-violent and those who were not reconvicted at third assessment.

31 The OGRS score represents the likelihood of reconviction within two years of community sentence or release on licence (Taylor, 1999). It is a very reliable predictor of reoffending (confirmed by Howard, Clark and Garnham, 2006), but is based only on static, actuarial characteristics (age, gender and current and past offending) and does not incorporate information on dynamic criminogenic needs.

In estimating the proportion of reviews which had been completed in a perfunctory manner, it is recognised that a zero-inflated negative binomial model would have been more appropriate than the zero-inflated Poisson model due to the skewed distribution of changes. However, it was not possible to fit this model using available statistical software in the time available. The estimates reported should therefore be considered approximate and only used in the present context of weighting results, rather than as an estimate of the true proportion of cases which were 'always zero'.

The study does not attempt to separate genuine changes from disclosure effects. 'Disclosure' describes the situation where item scores change because the assessor's understanding of the offender's condition improves, usually through the offender admitting to behaviours or problems which were previously concealed. Proper handling of disclosure data would considerably increase the complexity of data processing and analysis.³² Interpretation of the results should allow for the possibility of differential levels of disclosure. It is more likely when the assessor is reliant on the offender's account of events and cannot make an immediate independent judgement based, for example, on the facts of the current offence(s).

The method of measuring reconviction adopted in this study is unusual. In OASys studies where reconviction is the principal outcome of interest, searches of the Police National Computer (PNC) are undertaken in order to determine reconviction. PNC matching is clearly the most robust way of detecting reconviction, but was forsaken for two reasons. Firstly, PNC matching reduces sample size, as many offenders are not found and others cannot be studied as key details differ between the PNC record and OASys. Secondly, this paper is more concerned with the on-going assessment process than reconviction as such. If a reoffence is so minor that it is not identified through OASys, it should also have little effect on an offender's set of assessments.³³

Results

Table 7.1 summarises the numbers of assessments per offender, broken down by sentence, the type of initial assessment and reconviction status. As shown, for the whole sample, there was an average of 4.4 assessments per offender with an average interval between the first and second assessment of 118 days. The average number of assessments was higher for those offenders who had been reconvicted (38% of the sample) compared to those who had not been reconvicted; 5.9 compared to 3.5. Around half of the reconvicted offenders had four, five or six assessments.

32 The extent of disclosure is not always clear. For example, if an item score increases from 0 to 2 and the disclosure box is newly ticked, is the full increase due to disclosure or would an increase to 1 or 2 have happened anyway due to genuinely new information? If the score subsequently falls to 1 and the disclosure box is still ticked, has there been further, favourable disclosure or has the offender's real problem become less serious while the box has been erroneously left ticked?

33 It is important to recognise that the measure was of reconviction rather than all proven reoffending, as new cautions would not be detected. It is possible that some pseudoreconvictions – new convictions for offences committed before the original sentence – and corrections to existing data will be detected. Assuming that these events would cause little change to the rest of the assessment, the results below will therefore under-represent the degree of change noted when genuine reconviction occurs.

Grouping the offenders by sentence, those subject to a custodial sentence had the lowest average number of assessments (3.6) and the greatest average interval between the first and second assessment (159 days). One-third of these offenders had been reconvicted, which was a lower rate than for those subject to the various community sentences. For these latter offenders, the average difference of 0.9 assessments between those with an initial PSR assessment and those with an initial start community assessment can be attributed in part to some PSR offenders also receiving an assessment at the start of an order. National Standards require that the assessment should be reviewed within 15 working days of the start of an order, but this was not universal in practice and does not explain all of the average difference. The average interval between first and second assessment was 88 days (over 12 weeks), and while the median was 42 days, there was a very long tail and 21% exceeded the average of 136 days for sequences where the first assessment was at start of sentence. In other words, not all of the second assessments were start of order assessments.

Table 7.1: Numbers of assessments by sentence, type of initial assessment and reconviction status

	n	Mean assessments in set	Mean interval between 1st and 2nd assessment (days)	% reconvicted
All	17,824	4.4	118	38
Sentence				
CPRO	2,505	4.8	98	42
CRO	10,230	4.6	105	38
DTTO	801	4.9	139	60
Custody/YOI	4,288	3.6	159	33
Community sentences: type of initial assessment				
PSR	8,508	5.0	88	42
Start community	5,028	4.1	136	36
Reconviction status				
Not reconvicted	11,001	3.5	117	N/A
Reconvicted	6,823	5.9	119	N/A

Table 7.2 presents a breakdown of the reasons for assessment sequences coming to an end, comparing offenders of different age, sex, ethnicity, sentence type and OGRS 2 score band.³⁴ Overall, 8% of assessments were not reviewed (i.e. there was only one assessment in the sequence), while almost a fifth (19%) of sequences did not end in a termination or reconviction. Terminations were somewhat more frequent than reconviction. There was little variation by gender. However, Asian and ‘other’ ethnicity offenders were less likely to be reconvicted. There was a pronounced age pattern, with older offenders being more likely to have only one assessment and no termination or reconviction. Offenders with higher OGRS 2 scores were,

³⁴ Around 2% of the offenders in Table 7.2 have been removed due to missing gender, age or OGRS data. The high proportion with missing ethnicity data, and its variation between probation areas, means that it was imprudent to remove those with missing ethnicity data.

as expected, far more likely to be reconvicted. Sex offenders were the most likely to have multiple assessments without termination or reconviction. The proportion of sequences ending in reconviction ranged from approximately one-fifth (19%) for sex offenders to approximately two-fifths (41%) for offenders convicted of theft and handling.

Table 7.2 also indicates that, where there was more than one assessment in the sequence, 69% demonstrated change in at least one scored question. The proportion with any change between initial and final assessment was 60% for those without a termination or reconviction assessment, 62% for those terminated, 72% for the small group with simultaneous termination and reconviction, and 84% for those reconvicted. It is thus apparent that the proportion with no change was disproportionately high for all groups, but especially the two groups with no evidence of reconviction. The zero-inflated Poisson model results suggest that 31% of sequences with no change were not completed well. Subsequent tables which look at changes in section scores, item scores or risk of serious harm therefore give sequences with zero changes a weight of 0.69. This means that the sequences of assessments for reconvicted offenders have a strong influence in these tables, as these assessments are more likely to be weighted at 1 rather than 0.69.

Tables 7.3 and 7.4 present several measures which summarise changes between the initial and final assessments in each sequence. Table 7.3 sets out the number of changes in the 73 scored items and the number of changes in the ten criminogenic needs, while Table 7.4 sets out changes in the dynamic part of the weighted OASys score³⁵ (which ranges from 0 to 118 when the 50 points for static items are omitted) and changes in the dynamic elements of OGP and OVP (both of which range from 0 to 40). As noted above, sequences with changes in none of the 73 scored items were weighted to reduce the proportion with no change, thus adjusting for poor quality review assessments.

As shown by Table 7.3, for the whole sample, an average of seven of the 73 items and 1.0 of the ten needs changed, with decreases slightly outweighing increases. Table 7.4 demonstrates that the OASys weighted score changed by an average of ± 7.3 points, with a net decrease of almost one point. This represented an average percentage change from the initial average score for the dynamic elements of $\pm 16\%$ ($7.3/46$). In comparison, OGP was less dynamic with an average change of $\pm 13\%$ ($1.6/12.8$) but OVP was more dynamic with an average change of $\pm 18\%$ ($2.2/12.3$). Average OGP and OVP scores both fell slightly. (See Chapter 6 for more information about OGP and OVP – the OASys Violence and General reoffending Predictors.)

35 Only the dynamic part of the weighted OASys score is used in order to allow fair comparison with the dynamic elements of OGP and OVP. The weighted OASys score is actually an obsolete version: in the attitudes section, question 12.2 ('Discriminatory attitudes') has now been replaced by 12.8 ('Motivation to address offending'). The old scoring is used here as completion of 12.8 was poor prior to it becoming a scored item. This also affects OGP and OVP, both of which include scores based on attitudes need.

Table 7.2: Reasons for closure of assessment sequences

Offender group	n	Reason for closure of assessment sequence				% of multi-assessment offenders with any initial-to-final change	
		Only one assessment	Multiple assessments but no termination or reconviction	Termination	Termination and reconviction at same assessment		
All	17,611	8%	19%	42%	1%	29%	69%
Gender							
Female	2,670	7%	18%	45%	1%	28%	68%
Male	14,941	8%	19%	42%	1%	30%	69%
Ethnicity							
Asian	499	15%	21%	43%	1%	19%	61%
Black	567	16%	20%	36%	1%	27%	67%
Mixed	267	12%	21%	34%	2%	32%	73%
Other	100	17%	28%	38%	1%	16%	46%
White	14,806	7%	18%	43%	1%	31%	69%
Missing/NS	1,372	8%	22%	45%	1%	23%	67%
Age group							
18–20	2,824	6%	13%	42%	2%	37%	71%
21–24	3,379	7%	17%	40%	1%	34%	70%
25–40	8,711	9%	19%	42%	1%	29%	68%
41+	2,697	9%	25%	47%	1%	18%	65%
Sentence							
CPRO	2,476	6%	21%	37%	2%	34%	69%
CRO	10,129	7%	20%	42%	1%	31%	69%
Custody/YOI	4,206	12%	17%	50%	2%	20%	68%
DTTO	800	8%	14%	29%	2%	48%	72%

Table 7.2: Reasons for closure of assessment sequences (continued)

Offender group	n	Reason for closure of assessment sequence				% of multi-assessment offenders with any initial-to-final change	
		Only one assessment	Multiple assessments but no termination or reconviction	Termination	Termination and reconviction at same assessment		Reconviction
OGRS score							
Low (0–30)	5,238	10%	25%	49%	1%	16%	65%
Medium (31–74)	7,279	8%	19%	42%	1%	30%	69%
High (75+)	5,094	6%	13%	37%	2%	43%	71%
Offence type							
Violence against the person	3,879	8%	22%	43%	1%	26%	70%
Sexual	492	11%	35%	35%	0%	19%	71%
Burglary	1,537	6%	18%	39%	2%	35%	72%
Robbery	622	11%	15%	49%	1%	24%	76%
Theft and handling	3,240	6%	14%	38%	1%	41%	69%
Fraud and forgery	509	11%	18%	44%	1%	25%	67%
Criminal damage	585	7%	17%	41%	2%	34%	73%
Drug offences	1,358	11%	18%	49%	1%	21%	65%
Other/motoring	5,389	8%	19%	45%	1%	27%	66%

Patterns by key offender characteristics were as follows.

Gender: All measures were slightly more dynamic for male offenders, but net decreases in item scores, criminogenic needs and the OASys weighted score were greater for female offenders.

Ethnicity: Positive and negative changes in criminogenic needs, the OASys weighted score and both OGP and OVP were least likely for Asian offenders and most likely for Mixed Ethnicity offenders.

Age: Initial scores and total changes were lower for the oldest offenders, who experienced far fewer item score and need increases but only slightly fewer decreases than the youngest offenders.

Sentence: Assessments for offenders on Drug Treatment and Testing Orders (DTTOs) were most dynamic on all measures except OVP, while offenders on licence had large net reductions in the OASys weighted score and OVP. Assessments were least dynamic for offenders subject to Community Rehabilitation Orders (CROs). These sentences have now been replaced by Community Orders under the Criminal Justice Act 2003.

OGRS 2 score: Assessments for offenders with low OGRS 2 scores were least dynamic but these offenders had the largest net reductions in scores.

Termination/reconviction status: Assessment sequences ending in reconviction were most dynamic. Around one in five of the reconvicted offenders changed score on at least 16 of the 73 scored items, compared with only one in ten of those with no evidence of reconviction. The reconvicted offenders were the only group to increase on the various total scores, while those assessment sequences ending in termination had very few increases and consequently showed large net reductions on all total scores.

Offence group: Assessments for those offence groups with higher levels of problems at the initial assessment, e.g. burglary and robbery, tended to be the most dynamic. Assessments for those convicted of fraud and other/motoring offenders tended to be among the most static. Offenders convicted of robbery and drugs offences had the largest net decreases in item scores, criminogenic needs, the OASys weighted score and the OGP score. Offenders convicted of violence against the person had the largest net decreases in the OVP score. Those convicted of theft offences had the poorest net results on the existing measures, and were among those with less favourable results on OGP and OVP.

Table 7.3: Changes in individual items and criminogenic need scores

Offender group	Weighted n	Item score changes			Need changes		
		All	Increase	Decrease	All	Increase	Decrease
All	14,643	7.0	3.4	3.6	0.98	0.45	0.53
Gender							
Male	12,416	7.0	3.4	3.6	0.98	0.46	0.53
Female	2,228	6.8	3.2	3.6	0.94	0.42	0.52
Ethnicity							
White	12,433	7.1	3.4	3.7	0.99	0.46	0.53
Black	427	6.9	3.5	3.3	0.93	0.48	0.45
Asian	371	5.1	2.4	2.7	0.74	0.34	0.40
Mixed	215	7.4	3.5	3.9	1.11	0.54	0.56
Other	69	4.9	2.4	2.5	0.87	0.40	0.47
Missing/not stated	1,128	6.6	3.0	3.5	0.93	0.40	0.53
Age group							
18–20	2,424	7.4	3.8	3.6	0.97	0.47	0.50
21–24	2,847	7.4	3.6	3.8	1.01	0.48	0.53
25–40	7,188	7.0	3.3	3.7	1.00	0.46	0.54
41+	2,184	5.8	2.6	3.2	0.86	0.36	0.50
Sentence							
CPRO	2,101	7.0	3.5	3.5	1.01	0.49	0.52
CRO	8,538	6.7	3.4	3.3	0.94	0.46	0.48
DTTO	673	8.8	4.4	4.4	1.19	0.56	0.63
Custody/YOI	3,331	7.3	3.1	4.2	1.01	0.40	0.61
OGRS score							
Low (0–30)	4,213	5.7	2.5	3.3	0.81	0.33	0.48
Medium (31–74)	6,059	7.0	3.5	3.6	1.01	0.47	0.54
High (75+)	4,370	8.0	4.1	4.0	1.09	0.54	0.55
Termination/reconviction							
Neither	2,915	5.9	2.7	3.2	0.86	0.39	0.47
Reconviction	4,927	9.4	5.7	3.7	1.23	0.73	0.50
Termination	6,594	5.6	1.9	3.7	0.84	0.27	0.57
Both	206	8.0	3.8	4.2	1.00	0.45	0.55
Offence group							
Violence against the person	3,145	7.0	3.2	3.7	0.93	0.42	0.51
Sexual	375	6.4	2.8	3.6	0.99	0.42	0.57
Burglary	1,284	8.5	4.2	4.3	1.18	0.56	0.62
Robbery	498	8.7	3.8	4.9	1.14	0.48	0.66
Theft and handling	2,673	7.4	3.9	3.5	1.05	0.52	0.53
Fraud and forgery	398	6.4	3.0	3.4	0.89	0.38	0.50
Criminal damage	481	7.3	3.6	3.7	0.93	0.47	0.47
Drugs offences	1,050	6.2	2.6	3.7	0.98	0.36	0.61
Other, including motoring	4,336	6.3	3.1	3.2	0.90	0.43	0.47

Table 7.4: Changes in likelihood of reoffending scores

Offender group	Weighted n	OASys weighted score			OGP dynamic score			OVP dynamic score		
		Initial	Net change	Abs change	Initial	Net change	Abs change	Initial	Net change	Abs change
All	14,643	46	-0.9	7.3	12.8	-0.1	1.6	12.3	-0.2	2.2
Gender										
Male	12,416	46	-0.9	7.4	12.7	-0.1	1.7	12.5	-0.2	2.3
Female	2,228	48	-1.1	7.0	13.3	-0.1	1.6	11.0	-0.2	1.9
Ethnicity										
White	12,433	47	-0.9	7.4	13.1	-0.1	1.7	12.6	-0.2	2.2
Black	427	41	-0.6	7.0	11.8	-0.2	1.6	10.2	-0.2	2.0
Asian	371	34	-0.9	6.6	9.4	-0.2	1.3	8.4	-0.2	1.6
Mixed	215	48	-1.0	8.0	13.8	-0.1	1.9	11.8	-0.3	2.3
Other	69	36	-0.2	7.0	9.1	0.4	1.6	8.8	-0.5	1.9
Missing/not stated	1,128	41	-1.0	7.0	10.8	-0.2	1.4	11.6	-0.3	2.2
Age group										
18–20	2,424	50	-0.5	7.3	13.5	0.0	1.7	13.9	-0.1	2.3
21–24	2,847	49	-0.8	7.5	13.9	-0.2	1.8	12.7	-0.2	2.2
25–40	7,188	46	-1.0	7.6	13.3	-0.2	1.7	12.0	-0.2	2.2
41+	2,184	36	-1.4	6.5	8.8	-0.1	1.1	11.0	-0.5	2.1
Sentence										
CPRO	2,101	40	-0.4	7.3	10.4	-0.1	1.5	12.1	-0.2	2.3
CRO	8,538	45	-0.4	6.9	12.1	0.0	1.5	12.5	-0.1	2.1
DTTO	673	66	-1.0	8.8	23.2	-0.6	2.6	11.7	0.3	2.2
Custody/YOI	3,331	49	-2.6	8.3	13.9	-0.4	1.9	12.1	-0.6	2.3
OGRS score										
Low (0–30)	4,213	32	-1.6	6.2	7.7	-0.2	1.1	9.9	-0.6	2.0
Medium (31–74)	6,059	45	-0.8	7.7	12.7	-0.1	1.7	12.4	-0.2	2.2
High (75+)	4,370	60	-0.4	8.0	17.9	-0.1	2.1	14.5	0.0	2.3

Table 7.4: Changes in likelihood of reoffending scores (continued)

Offender group	Weighted n	OASys weighted score			OGP dynamic score			OVP dynamic score		
		Initial	Net change	Abs change	Initial	Net change	Abs change	Initial	Net change	Abs change
Termination/reconviction										
Neither	2,915	42	-1.1	6.5	11.2	-0.1	1.4	11.7	-0.4	2.0
Reconviction	4,927	53	1.9	8.7	15.2	0.4	2.0	13.5	0.6	2.6
Termination	6,594	43	-2.9	6.7	11.7	-0.6	1.5	11.6	-0.8	2.0
Both	206	53	-1.1	8.0	15.1	-0.1	1.8	14.0	-0.2	2.6
Offence group										
Violence against the person	3,145	43	-1.0	7.0	10.7	0.0	1.4	15.8	-0.7	2.5
Sexual	375	40	-1.9	7.7	9.5	-0.2	1.2	10.0	-0.5	2.4
Burglary	1,284	58	-0.8	9.0	17.5	-0.2	2.3	12.5	0.1	2.4
Robbery	498	53	-2.7	8.9	15.0	-0.6	2.4	12.7	-0.6	2.5
Theft and handling	2,673	56	-0.3	7.6	17.6	-0.2	2.0	11.6	0.2	2.0
Fraud and forgery	398	40	-1.2	6.9	10.7	-0.1	1.4	7.6	0.1	1.7
Criminal damage	481	49	-0.5	7.2	12.2	0.2	1.5	17.0	-0.4	2.4
Drugs offences	1,050	43	-2.9	7.7	13.8	-0.7	1.9	8.3	-0.3	1.8
Other, including motoring	4,336	39	-0.5	6.6	10.1	0.0	1.3	11.1	-0.2	2.1

Key: Net reduction in score No net score change Net increase in score

Changes in individual items

Table 7.5 sets out the 'top ten' and 'bottom ten' individual scored items in terms of score change prevalence. Sequences with no change in any scored items contribute to the results, but were weighted at 0.69 as described above. Not all items should be expected to be fully dynamic. Some items can only increase in score over time, while others can only decrease. For example, scores on item 4.9 (any qualifications) should only fall, as offenders can move from having no qualifications to having some qualifications but not vice versa. Scores on 8.7 and 9.4 (any drug/alcohol-related violence) should only rise, as scores of 0 apply when no such act has been committed and scores of 2 apply when any such act has been committed. In practice, some exceptions to these conditions will occur, as errors are corrected and new information uncovered over time. Appendix 9 fully explains which items are, by definition, not wholly dynamic, as well as setting out the levels of change for all 62 individual scored items across sections 3 to 12.

As shown by Table 7.5, the prevalence of score changes for individual items ranged from just 2% (Q4.8) to 22% (Q3.4). The top three ranked items all reside in the accommodation section, with a further three questions in the 'top ten' from the thinking and behaviour section. At the other end of the rankings, four of the items in the 'bottom ten' were from the education, training and employability (ETE) section.

Table 7.5: Prevalence of score changes for individual items

Item	% with score change
Ten items which change most frequently	
3.4 Suitability of accommodation	22.0%
3.5 Permanence of accommodation	21.8%
3.6 Suitability of location of accommodation	19.9%
11.7 Awareness of consequences	18.6%
11.6 Problem solving	18.2%
5.2 Financial situation	17.5%
4.2 Employment status	17.3%
11.5 Problem recognition	16.9%
10.1 Coping/depression	16.8%
7.2 Activities encourage offending	16.7%
Ten items which change least frequently	
8.6 Injecting drugs	5.0%
10.6 Current psychiatric problems	4.7%
4.9 Any qualifications	4.3%
6.2 Criminal family member	4.1%
9.4 Alcohol-related violence	4.1%
4.6 School attendance	3.9%
4.7 Reading/writing/numeracy	3.8%
6.5 Criminal partner	3.4%
8.7 Drug-related violence	2.5%
4.8 Learning difficulties	2.1%

Table 7.6 sets out the ‘top ten’ and ‘bottom ten’ individual items in terms of absolute and net changes in score. As shown, mean net changes ranged from -0.09 to 0.05, while mean absolute changes ranged from 0.03 to 0.35. Four of the top ten items in terms of absolute change were from the accommodation section, while four of the bottom ten items were from the ETE section. As for net changes, three of the items in the top ten (greatest increases in level of problems) were from the relationships section, while four of the items in the bottom ten (greatest decreases in level of problems) were from the thinking and behaviour section.

Table 7.6: Mean absolute and net changes for individual items

Item	Mean absolute change	Item	Mean net change
Ten items with most absolute change		Ten items with greatest net increases	
4.2 Employment status	0.35	6.2 Criminal family member	0.05
3.4 Suitability of accommodation	0.33	9.4 Alcohol-related violence	0.04
3.5 Permanence of accom.	0.31	12.4 Attitude to supervision	0.04
3.6 Suitability of location of accom.	0.31	6.6 Previous relationship experience	0.04
3.3 Currently no fixed abode	0.29	6.3 Experience of childhood	0.03
11.7 Awareness of consequences	0.21	9.3 Previous alcohol use	0.03
5.2 Financial situation	0.21	12.2 Discriminatory attitudes	0.02
8.5 Level of use of main drug	0.21	10.5 Suicide/self-harm	0.02
11.6 Problem solving	0.20	8.7 Drug-related violence	0.02
7.2 Activities encourage offending	0.20	12.3 Attitude to staff	0.02
Ten items with least absolute change		Ten items with greatest net decreases	
4.10 Attitude to education/training	0.08	7.2 Activities encourage offending	-0.04
6.5 Criminal partner	0.07	3.3 Currently no fixed abode	-0.04
12.2 Discriminatory attitudes	0.07	5.4 Illegal earnings	-0.04
6.3 Experience of childhood	0.07	11.5 Problem recognition	-0.04
8.6 Injecting drugs	0.06	11.2 Impulsivity	-0.04
10.6 Current psychiatric problems	0.06	8.9 Drugs major part of lifestyle	-0.04
4.6 School attendance	0.05	8.5 Level of use of main drug	-0.05
8.7 Drug-related violence	0.05	11.7 Awareness of consequences	-0.06
4.7 Reading/writing/numeracy	0.04	11.6 Problem solving	-0.06
4.8 Learning difficulties	0.03	4.2 Employment status	-0.09

The following commentary focuses upon each of the ten dynamic risk factors in turn.

Accommodation was the most dynamic factor. Items 3.4, 3.5 and 3.6 were the most likely to have any change in score and all four items were among the five with greatest absolute changes in score. However, most of the changes balanced out – the large net decrease for no fixed abode (3.3) was the only accommodation item in the net change ‘top’ or ‘bottom’ ten.

ETE: The education and training elements were mostly very static. With the exception of attitude to education/training (4.10), they were all among the ten items least likely to change, while all except any qualifications (4.9) were in the ten smallest absolute changes. By contrast, employment status (4.2) was one of the most dynamic of all items: it had by far the greatest net reduction in score of any item, yet was also the most likely to worsen in offenders with initial scores below two. The other three employability items were moderately dynamic, though employment history (4.3) rarely worsened over time.

Financial management and income: Current financial situation (5.2) changed often, with balancing increases and decreases. The percentage with illegal earnings (5.4) fell overall.

Relationships: The proportion with a criminal family member (6.2) had a large net increase, as it was an item whose score seldom fell. The same was true, to a lesser extent, of experience of childhood (6.3) and previous relationships experience (6.6). These items may be susceptible to disclosure effects, with offenders becoming more comfortable discussing family problems with their probation officers over time. Question 6.5 regarding criminal partners was exceptionally static for male offenders due to its limited relevance (35% of women and 6% of men had a criminal partner at the initial assessment) but quite dynamic for females (10% of women and 2% of men had any change).

Lifestyle and associates was a generally dynamic section, though most items ranked outside the 'top tens'. The most volatile item, regular activities encourage offending (7.2), had a net decrease.

Two **drug misuse** items (8.5 and 8.9) were among those with the greatest net decreases. Drug-related violence (8.7) was in the top ten for net increases despite being one of the least dynamic of all items – the vast majority of offenders scored zero and its scores changed very infrequently, but almost all its changes were score increases. Question 8.6 was another item with a low mean whose score should, in theory, never fall to zero and, in practice, was very static.

Alcohol misuse: Alcohol-related violence (9.4) had a large net increase but little overall change, in a similar manner to drug-related violence. The scores for this item and previous alcohol use (9.3) rarely decreased. By contrast, current alcohol use (9.1) and binge drinking (9.2) were among the ten items most likely to decrease in score.

Emotional wellbeing: Three of the six emotional wellbeing items feature in Tables 7.5 and 7.6. Coping (10.1) was fairly dynamic, while suicide/self harm (10.5) is another item where scores can rise but (according to the OASys manual) should not fall, placing it in the 'greatest increases' list. Scores for current psychiatric problems (10.6) were extremely static, with few increases. This item was unusual in being in the 'least likely to change' list as it does not relate to other people rather than the offender and it is not inherently historic.

Thinking and behaviour was a dynamic section, with no items in the 'least likely' list for any or mean absolute score changes. Problem recognition (11.5), problem solving (11.6) and awareness of consequences (11.7) were among the top ten items with any score change and the latter two items were in the top ten for mean absolute changes. All three items, along with impulsivity (11.2), were among the ten items with the greatest net decreases, though many score increases also occurred.

Attitudes problems demonstrated some worsening, chiefly in attitude to supervision (12.4). Worsening discriminatory attitudes (12.2) and attitude to staff (12.3) both followed very low mean initial scores. These increases may be the result of disclosure through ongoing contact between assessor and offender.³⁶

³⁶ Motivation to address offending (12.8) was not analysed due to high levels of missing data (around 10%). It is known to have a far higher mean than 12.2, which it replaced in the current scoring system in 2006. The completion rate for 12.8 has improved greatly since becoming a scored item.



Changes in criminogenic needs

Table 7.7 sets out the levels of change across the ten criminogenic need measures. As shown, accommodation was the section with the greatest absolute change, with around 8% of offenders moving each way across the criminogenic need threshold. The (i) thinking and behaviour and (ii) lifestyle and associates sections both had over 11% absolute change. Alcohol misuse and ETE changed least often (6.5% and 7.1%). These findings should be interpreted in light of the proportion ever having each need. For example, ETE seems particularly static given that it is a relatively frequent need.

The poor dynamism of the alcohol misuse section was largely due to those items which are based on past events. A further 3.8% of the weighted total number in Table 7.7 (or 8.6% of those with an initial need) would have moved from having an initial alcohol misuse need to having no need at the final assessment if the section score was restricted to the more dynamic questions (9.1, 9.2 and 9.5). However, these offenders' unchanged scores on past alcohol misuse (9.3) and ever committing alcohol-related violence (9.4), which seldom decrease, ensured that they still had a criminogenic need at the final assessment.

Table 7.7: Changes in criminogenic needs

Section	% with change			% with need		% net change
	Any	Increase	Decrease	Initial	Final	
Accommodation	16.3%	7.8%	8.5%	40.2%	39.6%	-0.6%
Education, training and employability	7.1%	2.7%	4.4%	62.2%	60.5%	-1.8%
Financial management and income	9.8%	3.9%	5.9%	26.2%	24.1%	-2.0%
Relationships	8.7%	4.9%	3.8%	44.4%	45.4%	1.0%
Lifestyle and associates	11.5%	4.9%	6.6%	44.9%	43.2%	-1.7%
Drug misuse	8.0%	3.4%	4.5%	32.5%	31.3%	-1.1%
Alcohol misuse	6.5%	3.4%	3.1%	43.8%	44.0%	0.3%
Emotional wellbeing	9.1%	4.6%	4.4%	48.1%	48.3%	0.2%
Thinking and behaviour	11.9%	4.5%	7.4%	62.6%	59.6%	-3.0%
Attitudes	10.0%	5.8%	4.2%	26.1%	27.7%	1.6%

Key:  Net decrease in % with need  Net increase in % with need

In terms of net changes, the greatest decrease was for thinking and behaviour which fell by 3.0%, with (i) ETE, (ii) lifestyle and associates and (iii) financial management and income needs all decreasing by between 1.7% and 2.0%. The greatest net increases in the proportion with a need were for attitudes and relationships (1.6% and 1.0% net increase respectively).³⁷

³⁷ The attitudes results might be different were 12.8 scored rather than 12.2 (see Footnote 35), though the worsening in attitudes to supervision (12.4) would remain.

Changes in risk of serious harm between initial and final assessment

Table 7.8 sets out the extent of change in risk of serious harm (community) levels.³⁸ As shown, there was a change in the overall risk level in almost one in five (18%) of the assessments. The initial and final risk levels and net changes are shown, as well as the percentages whose final risk levels were different in any way to their initial risk levels. The risk to all four groups (known adults; children; general public; and staff) and the overall risk shifted from low and high/very high (H/VH) risk towards medium risk.³⁹ Allowing for the low proportions with a risk to children, this was the most volatile of the four risks. The proportion whose highest community risk was H/VH fell considerably, despite smaller decreases in the four specific risks. This is because most offenders with H/VH overall risk were only rated H/VH risk towards one victim group, and therefore only one fall in specific risk was needed to lower their overall risk to medium.

Table 7.8: Changes in community risk of serious harm ratings

Who is at risk? (Weighted n=14,643)	Initial			Final			Net change			% with any change
	Low	Med.	High/ VH	Low	Med.	High/ VH	Low	Med.	High/ VH	
Highest community-based risk	55.6%	36.8%	7.6%	54.8%	38.6%	6.6%	-0.8%	1.8%	-1.0%	18%
Risk to known adults	79.9%	16.2%	3.9%	79.0%	17.4%	3.6%	-0.8%	1.2%	-0.4%	11%
Risk to children	89.1%	8.8%	2.1%	88.2%	10.0%	1.8%	-0.9%	1.2%	-0.3%	6%
Risk to general public	67.3%	29.0%	3.7%	66.3%	30.3%	3.4%	-1.0%	1.4%	-0.3%	14%
Risk to staff	95.0%	4.5%	0.6%	94.6%	4.8%	0.6%	-0.4%	1.3%	0.1%	3%

Key: Net reduction in % at this risk level Net increase in % at this risk level

Table 7.9 compares patterns of change in the highest community risk by gender, ethnicity, age, sentence group, OGRS 2 score and termination/reconviction status at final assessment. The table shows initial risk levels well below the overall average for female offenders, those on DTTOs, and those convicted of theft and handling, fraud and forgery and drugs offences. The proportion of H/VH initial risk offenders increased with age. Initial risk levels were generally higher for those serving custodial sentences and those convicted of sexual offences, robbery, violence against the person and, to a lesser degree, criminal damage offences. Those who received neither a termination assessment nor were reconvicted or who were given a termination assessment upon reconviction (the 'Both' category) had somewhat higher risk levels.⁴⁰

³⁸ All figures combine high- and very-high risk offenders, as the latter form well below 1% of most offender groups.

³⁹ It is important to interpret these results in the context of overall increases in risk of serious harm levels. PSR and start of order/licence assessments show highest community risk levels rising from 67% low, 27% medium and 7% H/VH in July–September 2004 to 45%, 47% and 8% respectively in July–September 2006.

⁴⁰ The latter may reflect a more thorough approach for higher-risk offenders: their orders would continue without early termination if they were not reconvicted, but would be correctly terminated if they were reconvicted.

Table 7.9: Changes in highest community risk of serious harm rating by offender characteristics

Offender group	Weighted n	Initial			Final			Net change			% with any change
		Low	Medium	High/VH	Low	Medium	High/VH	Low	Medium	High/VH	
All	14,643	55.6%	36.8%	7.6%	54.8%	38.6%	6.6%	-0.8%	1.8%	-1.0%	18%
Gender											
Male	12,416	53.5%	38.2%	8.4%	52.6%	40.1%	7.3%	-0.9%	1.9%	-1.1%	19%
Female	2,228	67.6%	29.0%	3.4%	67.3%	30.1%	2.6%	-0.3%	1.1%	-0.8%	16%
Ethnicity											
White	12,433	55.5%	36.8%	7.6%	54.6%	38.7%	6.7%	-0.9%	1.9%	-1.0%	18%
Black	427	54.3%	36.4%	9.3%	52.6%	39.0%	8.4%	-1.7%	2.6%	-1.0%	23%
Asian	371	59.8%	33.0%	7.2%	60.6%	34.3%	5.1%	0.8%	1.4%	-2.1%	11%
Mixed	215	50.8%	40.9%	8.3%	46.2%	46.4%	7.3%	-4.5%	5.5%	-1.0%	20%
Other	69	62.1%	27.4%	10.5%	59.1%	35.9%	5.0%	-3.0%	8.5%	-5.5%	15%
Missing/not stated	1,128	56.3%	37.2%	6.5%	57.6%	36.9%	5.5%	1.3%	-0.3%	-1.0%	20%
Age group											
18-20	2,424	56.1%	38.3%	5.6%	53.0%	41.6%	5.4%	-3.1%	3.3%	-0.2%	19%
21-24	2,847	58.2%	36.1%	5.7%	56.0%	38.3%	5.7%	-2.3%	2.3%	0.0%	18%
25-40	7,188	54.7%	37.3%	8.0%	54.7%	38.5%	6.8%	0.0%	1.2%	-1.2%	18%
41+	2,184	54.9%	34.1%	11.0%	55.9%	35.7%	8.4%	1.1%	1.6%	-2.7%	19%
Sentence											
CPRO	2,101	58.0%	37.5%	4.5%	56.2%	39.1%	4.7%	-1.8%	1.6%	0.2%	20%
CRO	8,538	58.8%	35.7%	5.5%	57.6%	37.6%	4.8%	-1.2%	1.9%	-0.7%	18%
DTTO	673	71.8%	25.1%	3.1%	68.0%	29.5%	2.6%	-3.9%	4.4%	-0.5%	19%
Custody/YOI	3,331	42.6%	41.4%	16.0%	44.2%	42.6%	13.2%	1.6%	1.1%	-2.7%	19%
OGRS score											
Low (0-30)	4,213	55.7%	35.6%	8.7%	56.7%	37.0%	6.3%	1.1%	1.3%	-2.4%	18%
Medium (31-74)	6,059	56.3%	36.2%	7.5%	55.6%	37.8%	6.5%	-0.7%	1.6%	-1.0%	18%
High (75+)	4,370	54.6%	38.7%	6.7%	51.9%	41.2%	6.9%	-2.7%	2.5%	0.2%	19%

Table 7.9: Changes in highest community risk of serious harm rating by offender characteristics (continued)

Offender group	W'ted n	Initial			Final			Net change			% with any change
		Low	Medium	High/VH	Low	Medium	High/VH	Low	Medium	High/VH	
Termination/reconviction											
Neither	2,915	51.6%	38.3%	10.1%	51.6%	40.5%	7.9%	0.0%	2.2%	-2.2%	19%
Reconviction	4,927	55.0%	37.7%	7.3%	49.2%	42.2%	8.6%	-5.8%	4.5%	1.3%	22%
Termination	6,594	57.9%	35.5%	6.6%	60.5%	35.1%	4.4%	2.6%	-0.4%	-2.2%	16%
Both	206	54.1%	35.4%	10.5%	52.3%	38.2%	9.5%	-1.8%	2.8%	-1.0%	16%
Offence group											
Violence against the person	3,145	25.8%	59.4%	14.9%	28.0%	60.0%	12.0%	2.2%	0.6%	-2.9%	25%
Sexual	375	11.1%	47.6%	41.3%	11.4%	57.5%	31.2%	0.3%	9.9%	-10.2%	27%
Burglary	1,284	62.4%	32.4%	5.2%	59.3%	35.5%	5.2%	-3.1%	3.1%	0.0%	17%
Robbery	498	23.3%	60.5%	16.1%	27.3%	58.7%	14.0%	4.0%	-1.9%	-2.1%	24%
Theft and handling	2,673	74.1%	23.3%	2.6%	71.8%	25.9%	2.3%	-2.3%	2.5%	-0.2%	13%
Fraud and forgery	398	83.6%	15.2%	1.3%	80.5%	18.0%	1.5%	-3.1%	2.8%	0.3%	12%
Criminal damage	481	44.6%	44.8%	10.6%	40.6%	49.9%	9.5%	-4.0%	5.1%	-1.1%	22%
Drugs offences	1,050	79.2%	18.5%	2.3%	77.2%	20.6%	2.2%	-1.9%	2.0%	-0.1%	11%
Other, including motoring	4,336	64.3%	31.8%	3.8%	63.2%	32.9%	3.9%	-1.1%	1.1%	0.1%	18%

Key: Net reduction in % at this risk level

No net change in % at this risk level

Net increase in % at this risk level

Changes over time by key offender characteristics were as follows:

Gender: Assessments were more dynamic for male offenders, with 19% demonstrating change compared to 16% for female offenders. However, the proportion of H/VH risk offenders fell almost as much for female offenders as it did for males, despite the far higher proportion of males who were initially high risk. The fall of 0.8% for females is equivalent to 23% of their initial H/VH risk cohort, whereas the fall of 1.1% for males is equivalent to 13% of their initial H/VH risk cohort.

Ethnicity: Assessments for Black offenders were most dynamic, with nearly one-quarter (23%) demonstrating change. The greatest net shifts between risk levels were observed for offenders of Mixed and 'other' ethnicity, moving to medium risk from low and H/VH risk respectively, but their small numbers make these results unreliable. The reduction in H/VH risk Asian offenders and the generally static nature of Asian offenders' risk levels should also be interpreted with caution given their relatively small numbers.

Age group: Risk levels showed net reductions for older offenders and increases for younger offenders, though each group was equally dynamic. The size of the initial difference in proportions of H/VH risk offenders was thus reduced by the final assessment.

Sentence: Offenders on licence were still the highest risk group at final assessment, but offenders had shifted from H/VH to medium risk, and (unlike most offender groups) from medium to low risk. While DTTO offenders remained the lowest risk group and their assessments were least dynamic, a high proportion moved from low to medium risk.

OGRS 2 score: At initial assessment, offenders with low OGRS scores were most likely to be H/VH risk. By the final assessment, this pattern was reversed, with those with high OGRS 2 scores having slightly higher risk levels. This pattern is likely to be strongly influenced by age and offence category patterns, as both of these factors contribute to OGRS scores.

Termination/reconviction status: Assessments for reconvicted offenders were most dynamic, with 22% demonstrating change. Furthermore, these offenders were far more likely to increase in risk level (14% had an increase and only 8% a decrease). In contrast, those who received a termination assessment were more likely to decrease in risk (5% increase, 10% decrease). Offenders whose orders continued without either event occurring (who, as noted above, had a fairly high initial risk profile) tended to shift from high to medium risk.

Offence group: The four groups with over 10% H/VH initial risk (violence against the person, sexual offences, robbery and criminal damage) all experienced decreases of at least 1% in their H/VH risk proportions. Consequently, these final H/VH groups were relatively smaller than the corresponding initial H/VH risk groups by between 10% and 25%. Sexual offenders

had the greatest absolute and relative decreases, although the much larger group of violent offenders accounted for more of the total reduction in H/VH risk offenders. The proportions of low risk offenders within the robbery and violence against the person groups increased, while offenders convicted of burglary, theft and handling, fraud and forgery, and criminal damage had the greatest shifts from low to medium risk.

Are the changes predictive of reoffending?

Table 7.10 shows how those offenders reconvicted at their third assessment differed from those not reconvicted at this assessment, focusing on changes in the current OASys weighted score between the first and second assessment. A breakdown is provided for two factors strongly correlated with reconviction rates – whether there were new offences ‘in the pipeline’ (as indicated in the review plan) and OGRS 2 score.

For the whole sample, the mean change in OASys score between second and third assessment was a fall of two points. The 12% who were reconvicted had a mean rise of 0.8 points, while there was a mean 2.4 decrease for those who were not reconvicted. For both groups, therefore, the overall direction of change was in the correct direction, i.e. rises in score for those reconvicted and falls for those who were not reconvicted. However, over one-third of those in each group had changes in the wrong direction.

The difference between reconvicted and non-reconvicted offenders was largest for those with medium OGRS 2 scores: possibly due to the fact that the moderate initial OASys scores provided ample room for scores to increase or decrease with few ‘ceiling’ or ‘floor’ effects. Reconvicted offenders with low OGRS 2 scores had small falls in mean score, but the falls were greater for non-reconvicted offenders. Reconviction was more frequent, but not inevitable, for those with offences ‘in the pipeline’: the relationship between eventual reconviction and changes in score was weaker though still present in such offenders’ assessments.

Table 7.10: Overall OASys score changes between first and second assessment by reconviction status at third assessment

Offender group	n	Weighted score		% assessments with		
		Initial	Mean net change	score decrease	no change	score increase
All	1,862	66	-2.0	52%	11%	36%
Reconviction status						
Not reconvicted	1,637	65	-2.4	54%	12%	34%
Reconvicted	225	73	0.8	40%	9%	51%
Offence in the pipeline						
No						
Not reconvicted	1,495	63	-2.7	56%	11%	33%
Reconvicted	178	71	0.3	43%	11%	46%
Yes						
Not reconvicted	142	82	1.0	37%	16%	46%
Reconvicted	47	83	2.6	26%	4%	70%
OGRS score						
Low (0–30)						
Not reconvicted	606	38	-1.9	55%	12%	33%
Reconvicted	37	46	-1.0	51%	14%	35%
Medium (31–74)						
Not reconvicted	667	70	-2.5	54%	12%	34%
Reconvicted	107	69	2.0	38%	4%	58%
High (75+)						
Not reconvicted	364	99	-2.9	52%	12%	37%
Reconvicted	81	92	0.0	36%	15%	49%

Key: Net reduction in score No net change in score Net increase in score

Table 7.11 looks at the score changes for each of the ten dynamic criminogenic needs, again demonstrating some difference between those reconvicted and not reconvicted. The largest differences were in accommodation, ETE, financial management, and thinking and behaviour, where the mean changes in score for reconvicted offenders were +0.4 points greater than for non-reconvicted offenders. Differences of at least +0.2 points were also seen in the means for all other sections except for relationships, which had a difference of +0.1 points.

Comparing separately those with and without new criminal charges pending, the above patterns were broadly replicated in both cases, but were stronger when there were no such charges. The small overall difference in score changes for low OGRS 2 offenders and the small number within this group who were reconvicted makes the results for these offenders inconclusive, but strong patterns were evident in most sections for offenders with medium and high OGRS 2 scores.

Table 7.11: Section score changes between first and second assessment by reconviction status at third assessment

Offender group	n	Mean change in unweighted section score											
		Accommodation	ETE	Finance	Relationships	Lifestyle & associates	Drug misuse	Alcohol misuse	Emotional wellbeing	Thinking & behaviour	Attitudes		
All	1,862	-0.1	-0.3	-0.2	0.1	-0.3	-0.1	0.0	0.0	-0.3	0.0	-0.3	0.0
Reconviction status													
Not reconvicted	1,637	-0.2	-0.4	-0.3	0.1	-0.3	-0.1	-0.1	0.0	-0.4	0.0	-0.4	-0.1
Reconvicted	225	0.2	0.0	0.1	0.2	0.0	0.1	0.1	0.2	0.0	0.2	0.0	0.2
New criminal charge pending at 2nd assessment													
No													
Not reconvicted	1,495	-0.2	-0.4	-0.3	0.1	-0.3	-0.1	-0.1	-0.1	-0.1	-0.1	-0.5	-0.1
Reconvicted	178	0.2	-0.1	0.1	0.2	-0.1	0.1	0.0	0.1	0.1	0.1	-0.1	0.2
Yes													
Not reconvicted	142	0.2	-0.2	0.2	0.3	0.0	-0.1	0.3	0.3	0.3	0.3	0.3	0.1
Reconvicted	47	-0.1	0.1	0.1	0.1	0.1	0.2	0.6	0.5	0.6	0.5	0.6	0.3
OGRS score													
Low (0–30)													
Not reconvicted	606	-0.3	-0.3	-0.2	0.1	-0.2	-0.1	-0.1	-0.1	-0.1	-0.1	-0.3	0.0
Reconvicted	37	0.3	-0.3	-0.1	0.4	0.0	-0.2	-0.2	-0.2	-0.2	-0.2	-0.6	0.1
Medium (31–74)													
Not reconvicted	667	-0.2	-0.4	-0.3	0.1	-0.3	-0.1	0.0	0.0	0.0	0.0	-0.5	-0.1
Reconvicted	107	0.1	0.1	0.0	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.3	0.4
High (75+)													
Not reconvicted	364	-0.1	-0.5	-0.3	0.1	-0.4	-0.2	0.0	0.1	0.1	0.1	-0.5	-0.2
Reconvicted	81	0.1	-0.1	0.2	0.1	-0.2	0.1	0.3	0.3	0.3	0.3	0.0	-0.1

Key: Net reduction in score No net change in score Net increase in score

Implications

Many offenders had several OASys assessments and some key OASys items were highly dynamic. However, not all offenders had patterns of assessment which met National Standards for assessment timing and frequency, and a substantial minority of repeat assessments reported no change in any of the 73 scored OASys questions. While deviation from Standards will legitimately occur in a minority of cases, and it is possible that some offenders will demonstrate reaction to neither supervision nor external events, the prevalence of these outcomes suggests sub-optimal practice in many cases. Statistical modelling suggested that around one-third of assessments demonstrating no change were not properly considered reviews.

A new NOMS Offender Management metric on the timely completion of termination assessments should have a positive effect. In order to maximise its impact on the quality as well as the quantity of assessments, it should be accompanied by checks that some item scores have changed in the majority of cases. The OASys Quality Management Plan should thus scrutinise reviews, ensuring that they reflect changes whenever these have occurred.

On average, seven of the 73 scored questions changed. Those which are based on historic or family factors (i.e. out of the offender's control) rarely changed. There are strong moral and practical grounds for only including such items when absolutely necessary, and their continued use will need to be considered carefully alongside other research findings on the reliability and validity of OASys. Through the assessment framework, offenders should have the opportunity to demonstrate that they have addressed their offending-related behaviour and the risk factors which underpin their offending. Historic and family factors may have contributed to the offender's present situation, but it should be possible to measure their effect in terms of something amenable to change – for example, the offender's ability to cope with past conflict and abuse, or overcome past personal under-achievement, rather than the fact that these adverse events occurred. The alcohol misuse section is seriously compromised in this respect, while multiple ETE items are based on historic factors or are very difficult to change.

The dynamism of scores on all three OASys-based predictors of recidivism is reassuring. It remains to be proven that changes in predictor scores sufficiently predict changes in reoffending rates, but the findings set out in this report are encouraging with the mean OASys scores rising for those who were reconvicted and falling for those who were not reconvicted. Changes in all ten dynamic section scores were also correlated with reconviction.

Conclusion

The findings from analysis of OASys as a measure of change suggest that offending-related needs can change considerably, as shown by offenders who have repeated reviews and reassessments. The value of OASys as a measure of change, however, is affected

by the proportion of assessments for which no change is indicated on any of 73 scored questions, and the possibility that this reflects an automated process rather than considered reassessment of offending-related needs. In addition, the occurrence of change and its value in offender management is impossible to determine for the proportion of cases for whom no review or reassessment is completed at all. Performance on completion of reviews with considered reassessment needs to be improved.

Further analysis is needed in order to determine the extent to which the offending-related needs measured in OASys are sufficiently sensitive to change in such a way that improved scores can be read as indicative of reduced likelihood of future reoffending.

8. Predictive validity of the OASys self-assessment questionnaire (SAQ)

Introduction

The purpose of the OASys self-assessment questionnaire (SAQ) is to provide a more complete picture of offending-related risks and needs by allowing the offender the chance to comment upon how they see their life. The SAQ may identify issues not raised in the interview or differences of opinion that can usefully be discussed.

Self-assessment forms have been included in other tools designed to assess risk and need. A self-assessment form was included in ACE (Haslewood-Pócsik, 2001), and a self-report inventory (SRI) version of the level of service inventory (LSI) was developed and tested (Motiuk, Motiuk and Bonta, 1992). The Youth Justice Board has introduced a separate risk/needs assessment tool for young offenders, known as *Asset*, which has an accompanying self-assessment form entitled 'What Do You Think?' (Baker *et al.*, 2002). A 72-item Self-Appraisal Questionnaire, which attempts to predict violent and non-violent recidivism, has also been tested with offenders from Australia, Canada, England, Singapore and the United States (Loza *et al.*, 2004).

Self-assessment tools have a number of potential benefits.

- They can assist in engaging the offender in the assessment process (Merrington, 2004).
- They tend to be easy and quick to administer and practitioners do not require extensive training for interpreting the results (Loza *et al.*, 2000).
- They can be used to highlight internal cognitive issues and identify further needs, assisting with the targeting of interventions.
- They can be used to measure change (due to the emphasis upon dynamic risk factors) and assist with the prediction of reoffending.

While concerns are sometimes raised regarding the vulnerability of self-assessment forms to 'lying, manipulation, and self-presentation biases', there is evidence to suggest that they can be 'accurate and valid' and 'equivalent to traditional methods' in predicting violent recidivism and general recidivism (Loza *et al.*, 2004:1174-5). However, analysis for *Asset* and ACE found that their self-assessments were less predictive than the core practitioner assessments (Merrington, 2000; Haslewood-Pócsik, 2001).

Previous analysis of over 100,000 fully completed OASys SAQs revealed that offenders were: (i) more likely to acknowledge problems in certain areas of their lives than view these problems as linked to their offending; and (ii) they were less likely to report problems than practitioners. They also appeared to be more optimistic regarding their future desistance

than indicated by their OASys likelihood of reconviction scores, with more than two-fifths responding that they would **definitely not** offend again (Moore, 2007).⁴¹

Previous research on the SAQ has not included assessment of its predictive validity. Assessment of predictive validity is thus important to ensure that the SAQ measures what it is intended to measure, i.e. offending-related problems and likelihood of further offending.

As noted above, there is evidence to suggest that self-assessments can be accurate and valid in predicting reoffending. Potentially, therefore, the SAQ could provide an easily obtainable indication of an offender's likelihood of further offending. However, as a full assessment needs to take into account all other available evidence, there is no expectation that the SAQ will outperform the core assessment in terms of predictive validity. The purpose of the research was to address the following three research questions.

1. How accurate are the offenders' judgements regarding their likelihood of further offending?
2. Are the SAQ individual-level and social problem questions predictive of reoffending?
3. How does the predictive validity of the SAQ compare to the current predictors?

Method

Sample

The sample was restricted to those offenders with two-year reoffending data and sufficiently complete data from the core OASys assessment, risk of serious harm assessment and the SAQ. The initial sampling frame was extracted from the O-DEAT database, selecting those 445,622 assessments for 215,941 offenders completed between January 2003 and March 2005 inclusive. All the SAQ questions had been completed in 83,711 (19%) of these assessments. These assessments were further filtered to ensure that the following standards of data completion had been satisfied.

- A sentence date had been recorded.
- All 73 scored items in the core assessment had been completed.
- In the risk of serious harm sections, the screening had been completed, the decision whether to complete a full risk analysis was consistent with the information provided, and the four ratings of risk of serious harm in the community had been recorded in those cases in which a full analysis was required.

41 Of those offenders who responded that they were **very likely** to offend again, over half said that they had a problem with drugs and approximately a quarter referred to drugs when explaining their likelihood of further offending. While female offenders were more likely than male offenders to perceive themselves as having a large number of problems, they were less likely to respond that they were **very likely** to offend again.

This left a total of 45,116 cases (10% of assessments in initial sampling frame) for submission to the Police National Computer (PNC) criminal careers database managed by the Reoffending and Criminal History Team in the Ministry of Justice.⁴² The PNC database lists offence dates and records cautions, reprimands and final warnings as well as convictions. This enables measurements of 'proven reoffending' within a given period, rather than the less complete measurement of whether an offender has been reconvicted.

Once successfully matched, the PNC records were processed to determine whether the cases could be followed up for 24 months at liberty from the date of the community sentence or discharge from custody, allowing three months for sentence and data entry to occur. The remaining 27,276 cases were then further filtered to ensure that the confirmed community sentence/custodial release date and the OASys completion date were within 90 days of each other. Any duplicates were removed by selecting the nearest assessment to each community sentence/custodial release date. This left a final sample size of 9,065 cases for 8,863 different offenders (representing 4% of the offenders in the initial sampling frame) for use in the analysis.

Analysis

To test the predictive validity of individual SAQ questions, chi-square tests were used. The final part of the SAQ enables offenders to elaborate upon their views regarding their likelihood of further offending. A linguistic-based text mining tool was used to process this qualitative data, employing advanced linguistic technologies and Natural Language Processing (NLP). Key concepts/terms were extracted automatically, representing the essential information, and closely related concepts were then grouped into higher-level categories, both manually and through further linguistic-based methods.⁴³

To test the predictive validity of the SAQ as a whole, alongside the current predictors, the full sample was divided into construction (60%; n=5,402) and validation samples (40%; n=3,663). Using the construction sample, logistic regression models were used to account for the relationships between the independent variables, identifying which were most predictive of reoffending and then calculating predicted rates.⁴⁴ The accuracy of the models was then checked using the validation sample, assessing goodness-of-fit and whether there was sufficient discrimination. The following models and current predictors were compared.

42 PNC numbers were recorded within OASys for most offenders, and an automatic matching procedure found reliable PNC numbers for most of the remaining cases. Cases in which the PNC did not record the offender's sex or recorded an unfeasible date of first or current conviction were rejected.

43 Linguistic text mining looks at how words are constructed, how words combine to make sentences and the meaning of words, phrases and sentences. The extracted key concepts can include compound words, phrases and idioms when appropriate.

44 The independent variables were entered using a forward stepwise approach, incorporating the most significant variables in turn and then removing them at a later stage if necessary.

- SAQ model: all SAQ questions were entered into the model.
- OASys weighted score: the current weighted OASys likelihood of reconviction score was entered.
- OGRS 2: the previous version of the Offender Group Reconviction Scale (OGRS) score was entered.
- OGRS 3: the most recent version of the OGRS score was entered.⁴⁵
- OASys model: all 73 scored questions from the core OASys assessment were entered.
- Combined OASys/SAQ model: all 73 scored questions from the core OASys assessment and all SAQ questions were entered.
- Combined OGRS 3/OASys model: the OGRS 3 score and all 73 scored questions from the core OASys assessment were entered.
- Combined OGRS 3/OASys/SAQ model: the OGRS 3 score, all 73 scored questions from the core OASys assessment and all SAQ questions were entered.

Model fit: actual and predicted proven reoffending rates

To assess the goodness-of-fit across risk levels, Hosmer-Lemeshow (2000) tests were used (dividing the validation sample into ten equal-sized deciles). Actual and predicted proven reoffending rates were then compared to assess whether any general over- or under-estimations had occurred. The difference between the actual and predicted rates – the residuals – should be as close to zero as possible. Model fit was also tested for different offender groups within the validation sample, e.g. low vs. high risk of serious harm offenders.

Model discrimination: the Area Under Curve statistic

Area Under Curve (AUC) statistics were used to check that higher predicted scores represented a higher likelihood of reoffending. In practical terms, the statistic is equivalent to the probability that a randomly selected proven reoffender has a higher score than a randomly selected non-proven reoffender. AUCs of 0.5 are the practical minimum as these could be obtained randomly, while AUCs of 1 represent the hypothetical situation where all proven reoffenders have higher scores than non-proven reoffenders. In reality, there are many ‘medium risk’ offenders, for whom continued offending depends upon future and/or unknown contingencies. Furthermore, there is variation in the criminal careers of even the most persistent offenders, and the difficulties become even more pronounced, ‘the more specific the group and type of risk (offence) being assessed’ (Broadhurst 2000:113). Predicting reoffending is thus difficult, and when predicting violent reoffending AUC statistics above 0.7 appear impressive (see Chapter 6).⁴⁶

45 The original version of OGRS was launched in 1996, with a revised version (OGRS 2) launched in 2000. The new version, OGRS 3, was developed between 2005 and 2007 and is the most accurate and user-friendly version of OGRS to date. Using criminal history and demographic data, it provides a percentage prediction of proven reoffending within one and two years of discharge from custody or start of community order - it achieved an impressive AUC of 0.81 when predicting two-year reoffending for a validation sample (Francis, Soothill and Humphreys, 2006, unpublished). OGRS 3 is calculated automatically within OASys, but it can also be used as a stand-alone predictor. Unlike OASys, it does not take account of dynamic risk factors, which are vital in understanding why offending occurs, in the targeting of interventions and in measuring change over time.

46 When piloted, the current OASys scoring system achieved an AUC score of 0.764 for a validation sample (Howard, Clark and Garnham, 2006).

A weakness of AUC statistics is that they derive from the relative rankings of offenders – if one added 20% to every offender’s prediction, the AUC for the sample would not change, even though the proven reoffending rate would be severely overestimated. AUC statistics were thus supplemented by comparisons of actual and predicted proven reoffending rates. A further measure of accuracy is provided by the percentages correctly predicted (see Copas, 1992, unpublished). These values are calculated by dividing the predicted values into ‘high’ and ‘low’ at a point corresponding to the proportions who actually reoffend, and then treating all ‘high’ scores as predicting reoffending and all ‘low’ scores as predicting non reoffending. High scorers who reoffend and low scorers who do not reoffend are then counted as correct predictions.

Limitations

All measures of reoffending have their limitations and those based upon official records of reoffending or reconviction are well documented (e.g. Lloyd *et al.*, 1995). Notably, official records under-record actual offending behaviour and they are affected by the activities of practitioners within the criminal justice system. For example, if the police secured no convictions, the reoffending rate would be zero per cent (Shepherd and Whiting, 2006).

As the sample was restricted to those offenders with two-year reoffending data and sufficiently complete data from the core OASys assessment, risk of serious harm assessment and the SAQ, the 9,065 cases cannot be seen as representative of the wider offender population.

There is no national standard relating to completion of the SAQ, and as the SAQ is initially completed on paper, it is likely that some of the results were not transferred to the electronic system. Analysis revealed that the SAQ sample corresponded to 148,951 valid and de-duplicated core OASys assessments completed between January 2003 and March 2005 inclusive.⁴⁷ Comparing the SAQ sample to the other offenders in the wider sample, there were clear regional differences, with the East Midlands contributing 28% of the SAQ sample but only 13% of the wider sample. Furthermore, those in the SAQ group were less likely to:

- have ten or more preconvictions (24% compared to 32%);
- have received a custodial sentence (17% compared to 31%);
- have a high likelihood of reconviction (13% compared to 19%);
- present a high/very high risk of serious harm (4% compared to 9%).⁴⁸

The ability to assess the predictive validity of the SAQ for all high risk offenders is thus also restricted. Notably, the 45% reoffending rate for the SAQ sample is below the 55.5% reoffending rate reported for those adults who were released from prison or commenced a community penalty during the first quarter of 2004 (Cunliffe and Shepherd, 2007).

⁴⁷ Assessments were held to be valid when meeting the standards of data completion set out above as the three bulleted points in the sample sub-section. The earliest assessment for each unique offender and sentence date was then selected.

⁴⁸ All differences were statistically significant ($p < .001$).

The results of the linguistic text mining, used to extract the textual information recorded within the SAQ, are dependent upon the linguistic resources used. The dictionary resources include synonyms, words to be excluded from extraction, types that group together multiple terms, and other more specialised tuning algorithms, such as words not to be confused when fixing spelling errors. Further editing of these resources through multiple iterations could improve the accuracy and value of the concepts extracted.

Results

The accuracy of the offenders’ judgements regarding their likelihood of further offending

The final question within the SAQ (Q28) asks offenders whether they think that they are likely to offend in the future, with a four-scale response ranging from **definitely not** to **very likely**. Analysis revealed that this question was associated with reoffending (p<.001). As shown by Table 8.1, about one-third (34%) of the **definitely not** responders, over half (56%) of the **unlikely** responders and three-quarters (75%) of the **quite likely** responders had reoffended. While the reoffending rate for the **quite likely** responders was higher than the reoffending rate for the **very likely** responders (67%), this difference was not statistically significant.

Table 8.1: 24-month reoffending rate by response to SAQ question 28

Likely to offend in the future?	n	% of sample	24-month reoffending rate
Definitely not	5,081	56%	34%
Unlikely	3,256	36%	56%
Quite likely	608	6.7%	75%
Very likely	120	1.3%	67%
Total	9,065	100%	45%

The one-third reoffending rate for the **definitely not** responders adheres to the previous analysis of over 100,000 SAQs which revealed that offenders tended to be more optimistic regarding their future desistance than indicated by their OASys likelihood of reconviction scores (Moore, 2007). To further assess the accuracy of the offenders’ own predictions regarding further offending, the responses to question 28 were collapsed so that **definitely not** and **unlikely** equated to a negative prediction and **quite likely** and **very likely** equated to a positive prediction. As shown by Table 8.2, the offenders’ predictions proved correct in about three-fifths (59%) of the cases, but the offender wrongly predicted that he or she would not reoffend in about two-fifths (39%) of the cases – these cases representing ‘false negatives’.

Table 8.2: Perceived likelihood of reoffending vs. actual 24-month reoffending

Likely to offend in the future?	Reoffended				Total	
	Yes		No			
Definitely not/Unlikely	3,558	39.2%	4,779	52.7%	8,337	92.0%
Quite likely/Very likely	537	5.9%	191	2.1%	728	8.0%
Total	4,095	45.2%	4,970	54.8%	9,065	100%

Key: Correct predictions

The predictive validity of the SAQ individual-level and social problem questions

Questions 1 to 27 of the SAQ address a range of ‘external’ social problems encompassing accommodation, employment and finances, relationships and lifestyle, as well as ‘internal’ individual characteristics, covering values, perceptions, reasoning, beliefs, attitudes and goals. All 27 questions are prefixed by the phrase ‘Are any of these a problem for you?’ In addition to the yes/no response, the offender is asked to consider a further tick box asking ‘Is this problem linked to your offending?’

When combining the responses to the two parts of each question, Table 8.3 demonstrates that all 27 questions were associated with reoffending ($p < .05$). For each question, with the exception of ‘being lonely’, those offenders who responded that it was a problem were statistically significantly more likely to reoffend than those who responded that it was not a problem. For eight of the questions, including ‘being lonely’, those offenders who thought that the problem was linked to their offending were statistically significantly more likely to reoffend than those who responded that it was a problem (but not linked to their offending). In contrast, the reoffending rate was statistically significantly lower for those offenders who thought that ‘getting on with my husband/wife/partner’ (question 21) was linked to their offending.

In the core OASys assessment, questions are generally scored 0, 1 or 2. Using a similar scoring for questions 1 to 27 of the SAQ where 0 = **no problem**, 1 = **problem but not linked to offending** and 2 = **linked to offending**, the question scores were added to produce a total raw SAQ problems score. Analysis revealed that this raw score was predictive of reoffending ($p < .001$). As shown by Table 8.4, the 24-month reoffending rate increased to two-thirds (67%) for those who scored at least 26. However, about one-quarter (27%) of those offenders who thought that they had no problems (score of 0) had also reoffended. Table 8.4 also demonstrates that this latter group had a mean of 1.5 criminogenic needs and that, according to OGRS 3, they had a mean prediction of proven reoffending within two years of 39%.

Table 8.3: 24-month reoffending rate by responses to SAQ questions 1 to 27

SAQ question	24-month reoffending rate						Significance		
	No problem		Problem but not linked to offending		Linked to offending		Overall	No problem vs. problem	Problem vs. linked
	%	n	%	n	%	n			
1. Finding a good place to live	42%	6,984	56%	1,360	60%	721	***	***	n/s
2. Understanding other people's feelings	44%	7,797	55%	856	54%	412	***	***	n/s
3. Keeping to my plans	41%	6,643	56%	1,695	63%	727	***	***	**
4. Dealing with people in authority	43%	7,792	58%	877	66%	396	***	***	**
5. Gambling	45%	8,735	56%	228	51%	102	**	**	n/s
6. Mixing with bad company	38%	6,648	59%	922	67%	1,495	***	***	***
7. Being bored	39%	5,660	51%	2,219	64%	1,186	***	***	***
8. Being lonely	44%	7,017	47%	1,467	53%	581	***	n/s	*
9. Going to places which cause me trouble	41%	7,155	60%	899	65%	1,011	***	***	*
10. Taking drugs	40%	7,237	63%	591	69%	1,237	***	***	**
11. Drinking too much alcohol	43%	6,440	50%	732	50%	1,893	***	**	n/s
12. Losing my temper	41%	6,563	56%	1,187	56%	1,315	***	***	n/s
13. Doing things on the spur of the moment	37%	5,188	55%	1,693	57%	2,184	***	***	n/s
14. Repeating the same mistakes	36%	5,536	57%	1,573	60%	1,956	***	***	n/s
15. Getting violent when annoyed	42%	7,278	59%	928	58%	859	***	***	n/s
16. Reading, writing, spelling and numbers	44%	7,525	51%	1,360	58%	180	***	***	n/s
17. Getting qualifications	42%	6,929	55%	1,809	60%	327	***	***	n/s
18. Getting a job	41%	6,426	54%	1,916	61%	723	***	***	**
19. Keeping a job	41%	7,070	60%	1,515	62%	480	***	***	n/s
20. Managing money, dealing with debts	42%	6,201	52%	2,038	52%	826	***	***	n/s
21. Getting on with my husband/ wife/ partner	44%	7,838	53%	740	45%	487	***	***	**
22. Looking after my children	45%	8,545	55%	373	46%	147	***	***	n/s
23. Worrying about things	43%	5,018	48%	3,020	48%	1,027	***	***	n/s
24. Making good decisions	41%	6,513	56%	1,546	54%	1,006	***	***	n/s
25. Feeling depressed	43%	5,494	48%	2,325	49%	1,246	***	***	n/s
26. Feeling stressed	42%	5,121	50%	2,663	50%	1,281	***	***	n/s
27. Not having a partner	45%	8,104	51%	733	47%	228	**	**	n/s

Asterisks indicate whether rates differ significantly (confidence levels * < .05, ** < .01, *** < .001)

Table 8.4: 24-month reoffending rate by SAQ raw problems score

Raw problems score	Sample		Mean OGRS 3 prediction of proven reoffending (two years)	Mean no. needs (OASys sections 3 to 12)	24-month reoffending rate
	n	%			
0	1,216	13%	39%	1.5	27%
1–5	2,371	26%	46%	2.4	36%
6–10	2,041	23%	51%	3.4	45%
11–15	1,495	17%	57%	4.4	54%
16–20	966	11%	60%	5.2	60%
21–25	528	5.8%	66%	5.8	63%
26+	448	4.9%	70%	6.3	67%
Total	9,065	100%	52%	3.5	45%

Question 28 also asks ‘Why do you think this is?’, enabling offenders to elaborate upon their views regarding their likelihood of further offending. Having extracted the key concepts from the offenders’ responses through a linguistic-based text mining tool, these concepts were grouped into higher-level categories. The links between these categories and 24-month reoffending were then analysed. The strongest association ($p < .001$) was found for the category of ‘drugs’, which covered concepts relating to addiction, detoxification and relapse, specific types of drugs, substance misuse programmes and drugs workers. Of those who made some such reference to drugs, two-thirds (66%) had reoffended in the following 24-month period ($n=603$), compared to 44% of those who did not mention drugs ($n=8,462$).

Table 8.5 demonstrates that, across the other most common categories ($n \geq 100$), reoffending rates were also significantly higher for those who mentioned some aspect of the criminal justice system, covering concepts relating to the police, the courts, sentences, probation and custody, and for those who referred to their finances (either in terms of their earnings or debts) or their employment/training. In contrast, reoffending rates were significantly lower for those who expressed some type of regret, remorse or embarrassment about their offending.⁴⁹

⁴⁹ Of the other categories listed in Table 8.5, the ‘family’ category covered concepts relating to parents, siblings, partners, children and other family members, while the ‘alcohol’ category covered concepts relating to alcoholism and alcohol courses.

Table 8.5: 24-month reoffending rate by SAQ question 28 response categories

SAQ question 28 response category	24-month reoffending rate for responses			
	with relevant concept		with no relevant concept	
	%	n	%	n
Family	46%	1,271	45%	7,794
Criminal justice system***	52%	949	44%	8,116
Drugs***	65%	620	44%	8,445
Employment/training***	53%	610	45%	8,445
Alcohol	47%	339	45%	8,726
Regret/remorse***	25%	168	46%	8,897
Motivation/confidence	43%	120	45%	8,945
Finances*	55%	102	45%	8,963

Asterisks indicate whether rates differ significantly (confidence levels *<.05, ** <.01, ***<.001)

How does the predictive validity of the SAQ compare to current predictors?

To test the predictive validity of the SAQ as a whole, the cases within the construction sample were selected and all of the SAQ questions entered into a logistic regression model. The two parts to each of the questions 1 to 27 were entered separately as in many instances they had differing levels of predictive validity.

As shown by Table 8.6, 13 of the SAQ questions were included in the model. These questions can be grouped as follows:

- the offender’s own predictions regarding their further offending (Q28);
- five questions relating to the offender’s attitudes, thinking and behaviour (2, 4, 12, 13 and 14);
- two questions relating to emotional wellbeing (8 and 23);
- two questions relating to the offender’s lifestyle and associates (6 and 7);
- one question relating to education, training and employment (17);
- one question relating to accommodation (1); and
- one question relating to drug misuse (10).

The SAQ questions relating to financial management and income (5 and 20), relationships and children (21, 22 and 27) and alcohol misuse (11) were all excluded from the model.

Table 8.6: An SAQ logistic regression model for predicting reoffending

Question	Parameter estimate	Standard error of estimate	Odds ratio
1. Finding a good place to live – a problem?	.334	.077	1.397
2. Understanding other people's feelings – linked to offending	-.402	.153	.669
4. Dealing with people in authority – linked to offending?	.468	.159	1.596
6. Mixed with bad company – a problem?	.385	.080	1.470
7. Being bored – linked to offending?	.322	.095	1.380
8. Being lonely – a problem?	-.329	.081	.719
10. Taking drugs – a problem?	.347	.088	1.415
12. Losing my temper – a problem?	.188	.073	1.207
13. Doing things on the spur of the moment – a problem?	.275	.070	1.317
14. Repeating the same mistakes – a problem?	.352	.072	1.423
17. Getting qualifications – a problem?	.228	.075	1.257
23. Worrying about things – a problem?	-.218	.069	.804
28. Likely to offend in the future?			
Unlikely	.613	.064	1.847
Quite likely	1.193	.141	3.296
Very likely	.876	.265	2.400
Constant	-.945	.049	.389

Odds ratios are compared with reference categories of problem=no and link to offending=not ticked (Qs 1 to 27) and likely to reoffend=definitely not (Q 28).

The odds ratios, set out in the final column of Table 8.6, are an indication of effect size, grouping the offenders by their responses to each question and comparing the odds of reoffending between the groups. In this instance, odds ratios of more than one indicated that reoffending was less likely for those offenders who responded that they would definitely not offend in the future (Q28) or that they did not have a problem or that it was not linked to their offending (1 to 27), while odds ratios of less than one indicated that reoffending was more likely for these offenders. Put simply, odds ratios of less than one demonstrated that the problems/links to offending, when combined with the other questions, were moderating the impact of the other problems/links to offending.⁵⁰

As can be seen, three of the 13 questions had odds ratios of less than one, including both questions relating to the offender's emotional wellbeing – those offenders who thought that 'being lonely' or 'worrying about things' (8 and 23) were problems were less likely to reoffend than those who thought otherwise. The other question with an odd ratios of less than one was 'understanding other people's feelings' (2).⁵¹

50 Protective factors have been defined as 'those that moderate the effects of exposure to risk' (Youth Justice Board, 2005), but they are usually framed in terms of internal assets and external strengths (McCarthy, Laing and Walker, 2004). In this instance the moderating factors remain negative problems.

51 When entering the most common textual categories from question 28 (see Table 8.5) into the model alongside the fixed response questions, the categories of employment/training, the criminal justice system and regret/remorse were included in the model, with an odds ratio of less than one for the latter.

To compare the predictive validity of items within the SAQ and the core assessment, all SAQ questions and all 73 scored OASys questions were entered into a second regression model. The OGRS 3 score was also entered to control for static criminal history and offender demographic factors. The offender and practitioner views, as recorded in the SAQ and core assessment respectively, are not wholly independent, as the SAQ highlights the offender's thought processes and can identify important areas for discussion, prior to the completion of the core assessment. While the SAQ is not structured in the same way as the practitioner-completed assessment and there has been no formal validation of correspondence, 26 of the first 27 questions in the SAQ have similar scored items within the core assessment, although there are differences in wording.

As shown by Table 8.7, the new model included the OGRS 3 score, five SAQ questions and eight core OASys questions. The OGRS 3 score was most significant and thus entered at step 1 of the model – its odds ratio within the final model indicates that, when holding all other variables constant, the odds of reoffending doubled with an increase in the OGRS 3 score of over 20 percentage points. Those SAQ questions remaining in the model were as follows.

- the offender's own predictions regarding their further offending (Q28);
- three questions relating to the offenders' thinking and behaviour (2, 13 and 14);
- one question relating to emotional wellbeing (8);

For two of the five remaining SAQ questions (2 and 8), the odds ratio was less than one, indicating that, when combined with the other questions in the model, they were moderating the impact of the other problems/links to offending – those offenders who thought that 'understanding other people's feelings' was linked to their offending or that 'being lonely' was a problem were less likely to reoffend than those who thought otherwise.

The other SAQ questions, including those relating to education, training and employment, lifestyle and associates, accommodation, drug misuse and attitudes were thus displaced by questions from the core assessment. The latter questions were spread across seven of the scored OASys sections – no questions were included from the financial management and income, emotional wellbeing, thinking and behaviour, or attitudes sections. For question 8.7 from the core assessment, the odds ratio was less than one, indicating that reoffending was less likely for those who had previous violent behaviour related to drug misuse compared to those who did not.

Table 8.7: A combined OGRS 3, core OASys and SAQ logistic regression model for predicting reoffending

Question/score	Parameter estimate	Standard error of estimate	Odds ratio
SAQ 2 Understanding other people's feelings – linked to offending	-.373	.155	.688
SAQ 8 Being lonely – a problem	-.200	.084	.818
SAQ 13 Doing things on the spur of the moment – a problem?	.206	.074	1.229
SAQ 14 Repeating the same mistakes – a problem?	.176	.077	1.193
SAQ 28 Likely to offend in the future?			
Unlikely	.196	.072	1.216
Quite likely	.580	.154	1.786
Very likely	.326	.295	1.386
OASys 1.3 Total number of separate offences for which convicted at this court appearance (0 = "0", 1 = "2-3", 2 = "4+")	.092	.047	1.097
OASys 3.6 Suitability of location of accommodation	.241	.045	1.273
OASys 4.4 Work-related skills (2 = 'No skills')	.120	.045	1.127
OASys 6.4 Current relationship with partner	.205	.048	1.227
OASys 7.2 Regular activities encourage offending	.130	.053	1.139
OASys 8.5 Level of use of main drug (0 = 'less frequently than weekly', 2 = 'at least weekly')	.082	.041	1.085
OASys 8.7 Violent behaviour related to drug use (0 = 'No', 2 = 'Yes')	-.185	.065	.831
OASys 9.1 Is current use [of alcohol] a problem?	.136	.044	1.145
OGRS 3	.036	.002	1.037
Constant	-2.869	.098	.057

Odds ratios are compared with reference categories of problem=no and link to offending=not ticked (SAQ Qs 1 to 27), likely to reoffend=definitely not (SAQ 28), score=0 (OASys) and percentage=0% (OGRS 3). Unless indicated otherwise, OASys questions were scored: 0 = no problems; 1 = some problems and 2 = significant problems.

As shown by Table 8.8, the final regression model, combining questions from the SAQ with OGRS 3 and questions from the core OASys assessment, achieved a high level of discrimination for the validation sample with an AUC score of 0.785. In other words, nearly eight out of ten randomly selected reoffenders had higher scores than randomly selected non reoffenders. Comparing this model with one combining OGRS 3 and the core OASys assessment, this AUC score represented a very small improvement of .002. A model based purely upon SAQ questions achieved the lowest level of discrimination, with an AUC score of 0.697 – indicating that seven out of ten randomly selected reoffenders had higher scores than randomly selected non- reoffenders.

Table 8.8 also sets out the percentages correctly predicted for each predictor/model. As shown, the highest percentage was achieved by the combined OGRS 3/core OASys/SAQ

model, with a correct prediction in more than seven out of ten (72%) of the cases in the validation sample.⁵² The combined OGRS 3/OASys model performed almost as well. The lowest percentage correctly predicted was achieved by the model based purely upon SAQ questions, with a correct prediction in 66% of the cases. Thus, adhering to the findings from Asset and ACE (Merrington, 2000; Haslewood-Pócsik, 2001), the self-assessment was less predictive than the core practitioner assessment.

Table 8.8: Accuracy of established predictors and logistic regression models

Predictor	AUC score	Std. Error	95% confidence intervals		Per cent correctly predicted
			Lower bound	Upper bound	
SAQ model	.697	.009	.680	.714	66.0
OASys weighted score	.744	.008	.728	.760	68.9
OGRS 2	.768	.008	.752	.783	69.5
OGRS 3	.775	.008	.760	.790	70.4
OASys model	.758	.008	.742	.773	69.1
Combined OASys/SAQ model	.762	.008	.747	.778	69.6
Combined OGRS 3/OASys model	.783	.008	.768	.798	72.1
Combined OGRS 3/ OASys/SAQ model	.785	.008	.768	.800	71.8

In addition to testing model discrimination, it was necessary to assess the goodness-of-fit of the combined OGRS 3/core OASys/SAQ model. A Hosmer-Lemeshow test was used to divide the validation sample into ten equal equal-sized deciles (see Table 8.9). The test revealed a chi-square value of 13.761 with eight degrees of freedom. There was no statistically significant difference between the observed and expected values ($p=0.088$), indicating that the model was valid across risk levels.

Table 8.9: Goodness-of-fit of combined OGRS 3, core OASys and SAQ model

Grouping	n	Reoffending rate		Residual (actual minus predicted rate)
		Actual	Predicted	
1	362	9.7%	13.0%	-3.4%
2	364	15.4%	16.1%	-0.7%
3	364	22.8%	20.5%	2.3%
4	364	31.6%	28.0%	3.6%
5	364	39.0%	37.7%	1.3%
6	364	44.5%	48.9%	-4.4%
7	364	62.1%	58.6%	3.5%
8	364	66.5%	67.8%	-1.3%
9	364	73.4%	75.5%	-2.2%
10	364	84.1%	82.9%	1.2%

52 Copas (1992, unpublished) explains that for an actual reconviction rate of 50%, the proportion correctly predicted cannot normally exceed 75%, even for an optimally effective predictor. When the reconviction rate is slightly higher or lower than 50%, as with the 45% rate in this study, the maximum proportion correctly predicted is lower still.

Table 8.10: Residual values (from combined OGRS 3, OASys and SAQ model) for various offender groups

Grouping variable	Value	n	Reoffending rate		Residual (actual minus predicted rate)
			Actual	Predicted	
Age	18–20	678	59.6%	56.4%	3.2%
	21–24	776	50.8%	50.3%	0.4%
	25–30	724	47.2%	48.1%	-0.9%
	31–40	940	39.9%	41.8%	-1.9%
	41+	543	24.5%	27.9%	-3.4%*
Gender	Male	3,126	47.1%	46.5%	0.6%
	Female	537	33.1%	39.6%	-6.5%***
Ethnicity	White	3,156	46.0%	46.4%	-0.4%
	Black	100	44.0%	40.5%	3.5%
	Asian	91	29.7%	29.8%	-0.2%
	Mixed	42	45.2%	42.3%	2.9%
	Other	24	29.2%	38.7%	-9.5%
Offence	Violence against the person	927	40.2%	40.6%	-0.4%
	Sexual offence	48	20.8%	26.4%	-5.5%
	Burglary	203	62.1%	60.8%	1.2%
	Robbery	38	50.0%	41.1%	8.9%
	Theft and handling	590	63.7%	61.4%	2.4%
	Fraud and forgery	127	29.1%	34.8%	-5.6%
	Criminal damage	101	52.5%	52.1%	0.4%
	Drug offences	199	40.2%	45.4%	-5.2%
	Other indictable offences	271	35.1%	40.7%	-5.7%*
	Summary motoring offences	877	41.6%	39.9%	1.8%
	Other summary offences	217	37.8%	43.0%	-5.2%
Sentence	CPO	866	29.8%	31.0%	-1.2%
	CRO	1,281	50.4%	49.2%	1.2%
	CPRO	352	44.6%	47.0%	-2.4%
	Custody/YOI	518	51.7%	54.3%	-2.6%
	Other	144	65.3%	63.5%	1.8%
Risk of serious harm (highest community level)	Low	2,295	41.9%	43.0%	-1.1%
	Medium	1,228	49.2%	49.2%	0.0%
	High/Very high	138	60.1%	55.4%	4.7%

Asterisks indicate whether rates differ significantly (confidence levels *<.05, ** <.01, ***<.001).

The goodness-of-fit for different offender groups, within the validation sample, is demonstrated by Table 8.10. The predicted reoffending rate was significantly different from the actual rate for female offenders, those aged at least 41, and those whose offence fell within the category of 'other indictable offences'. The goodness-of-fit of the combined OGRS 3/core OASys/SAQ model was therefore less strong for specific offender groups. However, greater residuals resulted from the use of the OGRS 3 score alone (without the inclusion of the OASys and SAQ questions), with significant differences between the predicted and actual reoffending rates for nine of the 31 offender groups set out in Table 8.10.

Implications

The results of the analysis have the following implications for practitioners.

- Attention should be paid to whether offenders have realistic perceptions of their likelihood of reoffending and the links between criminogenic problems and offending. About two-fifths of the offenders wrongly predicted that they would not reoffend and about one-quarter of those offenders who thought they had no problems went on to reoffend.
- To help ensure that relevant offending-related factors are recognised and that differences in opinion are discussed, assessors should be encouraged to pay close attention to the offender's perceptions when completing the core OASys assessment and conducting the prior interview. While the SAQ, on its own, was less predictive than the core practitioner assessment, all the SAQ questions were associated with reoffending. A combined model for predicting reoffending included the offender's own predictions regarding their further offending and four problems questions from the SAQ rather than corresponding questions in the core assessment.

Implications for the future use and development of OASys are as follows.

- Reasons for non-completion of the SAQ should be explored and completion in all eligible cases encouraged, helping to ensure that offenders' views are fully considered. Offenders' views will thus feed into the core OASys assessment where appropriate, contributing to the measurement of criminogenic needs and the prediction of further offending.
- Consideration should be given to introducing a closer alignment between the structure of the SAQ and the core OASys assessment, assisting practitioners to compare views (see Appendix 11).

Conclusion

The OASys SAQ provides offenders with the opportunity to comment upon how they see their lives. An evaluation of the predictive validity of the SAQ was undertaken to ensure that the tool measures what it is intended to measure – offending-related problems and likelihood of further offending. The correlation between the offenders' views and further offending was found to be good but the SAQ did not outperform the core assessment for predicting further offending. Where differences arise between the core assessment and similar items on the SAQ, there is scope for ensuring that offending-related factors are recognised by both practitioners and offenders.

9. The coverage and representativeness of OASys risk and need offender profiles: 2007 probation commencements and sentenced prisoner receptions

Introduction

While OASys is now in general use, it is not required to be used with all offenders. At the Pre-Sentence Report stage, all standard delivery reports must be based on a full OASys assessment, but fast delivery and oral reports can be based upon an Offender Group Reconviction Scale (OGRS) score and an OASys risk of serious harm screening (National Offender Management Service, 2007a).⁵³ Post-sentence, a full assessment should be completed in the community for all those cases designated at Offender Management (OM) Tier 2 and above, with the exception of those Tier 2 cases in which there is a stand-alone unpaid work requirement.⁵⁴ A sentence plan is required in all cases, with a formal review of each offender's assessment and sentence plan required every four months, regardless of the length of sentence. In prison establishments, all offenders aged 18–20 and older offenders serving a custodial sentence of at least 12 months should be assessed with formal reviews every 12 months and prior to release (National Offender Management Service, 2007b).

OASys data produced for management information purposes should not be read as representative of the entire offending population and care should be taken in generalising the results. If OASys is targeted at higher-risk offenders or offenders with certain offence types or sentence lengths, then the resulting risk and need profiles will reflect only the risks and needs of those offenders and not all others.

To establish the current levels of coverage and representativeness of OASys, and the considerations that need to be made when interpreting OASys data for commissioning and allocating resources in offender management, the OASys data within the O-DEAT database were merged with the 2007 offender management data held by Offender Management and Sentencing Analytical Services (OMSAS) within the Ministry of Justice. To guide the analysis, the following three research questions were set.

- What percentage of offenders commencing supervision by the probation and prison services had a completed OASys assessment? (OASys coverage)
- Did OASys completion rates vary across offender sub-groups? (OASys representativeness)

53 Probation Circular 12/2007 sets out a 'decision tool' for assessing which type of Pre-Sentence Report is most appropriate in any individual case. The OGRS predictor, used within the decision tool, is based upon static criminal history and offender demographic factors.

54 Probation Circular 08/2008 sets out the Offender Management Tiers and how they are to be applied (National Offender Management Service, 2008). The four tiers represent levels of intervention, with the approach increasing in scale and complexity as the risks and needs of the offender, and the demands of the sentence, increase. An offender's tier level is thus dependent upon his/her likelihood of reconviction, risk of serious harm and various other factors, e.g. classification as a Prolific and other Priority Offender (PPO).

- What adjustments were required to offenders' risk and need profiles to reflect all commencements/receptions? (Score adjustments)

Method

Sample

In addressing these questions, data from the following two OMSAS datasets was obtained:

1. Probation form 20 extract – January to December 2007 returns;
2. Prison receptions extract – January to December 2007 returns.

The form 20 extract (dataset 1 above) provides monthly information on probation commencements, caseloads and terminations. To establish OASys coverage and representativeness for all new probation cases, the commencements data (L1) were requested. The prison receptions extract (dataset 2) holds details of all those received into prison on remand (untried or convicted unsentenced), under sentence, as a non-criminal, or having been recalled to custody.⁵⁵ To establish OASys coverage and representativeness for all those sentenced to custody, the extract was restricted to those offenders under sentence.

The cases within the two OMSAS datasets were merged with O-DEAT's probation and prison OASys data, using the following variables: (i) date of birth; (ii) surname; and (iii) first initial.⁵⁶ It was ensured that the dates of probation commencement/custodial sentence and OASys completion were within 16 weeks of each other and that the offender was aged at least 18. Both pre- and post-sentence assessments were included in the merge. Duplicate assessments for the same cases were then removed by prioritising valid OASys assessments and selecting those cases in which the OASys completion date and the probation commencement/custodial sentence date were most closely matched.⁵⁷ For an assessment to be held valid, the following standards of data completion had to be satisfied.

- Each of the scored sections (1 to 12) within the core OASys assessment must have had at least four-fifths of their scored items completed – ensuring that each criminogenic need was assessed properly.
- In the risk of serious harm component of OASys, the screening must have been completed, the decision whether to complete a full risk analysis should have been consistent with the information provided, and the four ratings of risk of serious harm in the community must have been completed.

⁵⁵ OMSAS also maintain a prison population extract, but this is likely to under-represent those offenders with short custodial sentences who are processed more quickly.

⁵⁶ Merging on these variables only, without a common identifying number, does not eliminate the possibility of an inaccurate match.

⁵⁷ The de-duplication was of assessments rather than offenders – different probation commencement/custodial sentence dates for the same offenders were counted as separate cases. National standards require new OASys assessments for eligible offenders in each of these cases.

Analysis

When reviewing the findings regarding the coverage and representativeness of OASys, the standards set out within the directory of clinical databases were employed (Black and Payne, 2003). This directory employs a data quality checklist, including four aspects relating to the coverage of the data and six aspects relating to the accuracy of the data, providing an indication of validity and reliability.⁵⁸ In establishing the proportion of the eligible population that the database includes, the following levels are specified:

- Level 1: Unknown or few (<80%)
- Level 2: Many (80–89%)
- Level 3: Most (90–97%)
- Level 4: All or almost all (>97%)

As the directory recognises, if a significant proportion of the population that the database seeks to include are not captured, selection bias may be introduced whereby those included are systematically different from those who are not included. Importantly, selection bias reduces the generalisability of the results to the whole population.

In establishing the extent to which the eligible population can be generalised to the whole population, the directory employs the following levels.

- **Level 1: No evidence or unlikely to be representative.** The sample is unlikely to be representative if those included represent a subgroup.
- **Level 2: Some evidence that eligible population is representative.** Basic comparisons have been made with the reference population which show that, for example, the socio-demographic distribution of the eligible population and the total population are similar.
- **Level 3: Good evidence the eligible population is representative.** Comparisons between the eligible population and the reference population show similar characteristics such as demographics; and/or a sampling frame has been used that captures a representative sample.
- **Level 4: Total population included.** Every individual who has the common circumstance that determines inclusion is included in the database.

These levels of coverage and representativeness were applied when analysing how many offenders commencing probation or prison supervision had a completed OASys assessment and when comparing the OASys completion rates for different offender sub-groups (with logistic regression being used to account for the relationships between the independent variables). The analysis divided the cases into those which met the post-sentence eligibility

⁵⁸ A valid instrument successfully measures what it is supposed to measure, while a reliable instrument produces consistent measurements.

criteria for OASys and those which did not, as well as looking at the validity rates of completed assessments, bearing in mind that offender risk and need profiles are generated from valid assessments rather than all completed assessments. As set out above, valid assessments are those which meet specified minimum standards of data completion.

To assess whether adjustments were required to the offenders' risk and need profiles to reflect all commencements/receptions, classification decision tree models were employed to predict risk and need scores. These models divided the offenders into sub-groups according to those independent variables which produced the most accurate predictions. The approach is sometimes known as rule induction, with the splits in the models representing sets of decision rules.

The offender characteristic variables set out in Table 9.1 were entered into the decision tree models as independent variables, with valid risk and need scores entered as the dependent variables. Offenders' criminogenic needs were based upon the section scores only and did not take into account the assessors' more clinical judgements regarding links to offending behaviour,⁵⁹ while the risk of serious harm level was based upon the highest of the four risks (children/public/known adult/staff) in the community.

Table 9.1: Offender characteristic variables

Variables	Values	Probation form 20 extract	Prison receptions extract
Gender	Male; Female	✓	✓
Age	in years	✓	✓
Ethnicity	White; Black; Asian; Mixed; Other	✓	✓
Offence	Violence against the person; Sexual offences; Burglary; Robbery; Theft and handling; Fraud and forgery; Criminal damage; Drugs offences; Other offences	✓	✓
OGRS 2 score	0 to 100%	✓	x
OM Tier level	1 to 4	✓	x
Sentence	CJA 03 Community Sentence; Custody/YOI; Suspended Sentence; Other	✓	x
Number of previous convictions	-	x	✓
Number of previous custodial sentences	-	x	✓
Sentence length	In days	x	✓
Security classification	Cat A; Cat B; Cat C; Cat D; YOI closed; YOI open/short sentence; Uncategorised	x	✓

⁵⁹ There is a question at the end of each section which allows the assessor to make a yes/no judgment as to whether the section is linked to offending behaviour. The OASys pilot study found that these clinical judgements were inferior to the scored measures as predictors of reconviction (Howard, Clark and Garnham, 2006).

The goodness-of-fit of the decision tree models was checked by comparing the predicted risk and need levels with the actual valid risk and need levels to assess whether any general over- or under-estimations had occurred. To ensure that the differences between the actual and predicted levels – the residuals – were as close to zero as possible, prior probabilities (estimates of the overall relative frequency of each outcome) were entered into the models, and these were adjusted using misclassification costs (specifying the relative importance of different kinds of prediction errors).⁶⁰ However, small residuals do not mean that decision tree models are accurate, in the sense that the predictions can be entirely random even if the average prediction fits the actual rate. The analysis also included, therefore, an assessment of the percentages correctly predicted, with the cases again being divided into those which met the post-sentence eligibility criteria for OASys and those which did not.

National risk and need levels were then calculated for all offenders, using actual valid OASys risk and need scores where available and using the predicted risk and need scores in the remaining cases.

Limitations

For the prediction of risk and need scores, classification tree models were preferred to logistic regression models due to the greater control over the residual values. The rules derived from such models tend to have a straightforward interpretation, and the models are quite robust when dealing with missing data.⁶¹ The accuracy of the models was restricted, however, by the limited nature of the independent offender characteristic variables and their inability to explain all the variation in offenders' risk and need profiles. Unfortunately, this is an inevitable consequence of the restrictions upon the use of OASys and the limited amounts of information collected in the non-assessed cases.

The OGRS 2 predictor provides an easily obtainable likelihood of reconviction score, based upon static criminal history and offender demographic factors, but while this was available for the majority of probation commencements, it was not recorded for the prison receptions. There were also some concerns regarding the accuracy of the number of previous convictions/custodial sentences variables within the prison receptions dataset. More generally, as the data sources for the analysis are administrative IT systems, the detailed findings should be seen as subject to the inaccuracies inherent in any large-scale recording system.

60 Misclassification costs are basically weights applied to specific outcomes and can change the prediction as a way of protecting against costly mistakes. For example, it was thought to be particularly costly to classify a high/very high risk of serious harm offender as a low risk of serious harm offender.

61 The accuracy of the classification tree models was compared to the accuracy of logistic regression models, CHAID (Chi-squared Automatic Interaction Detection) models, QUEST (Quick, Unbiased, Efficient Statistical Tree) models and neural network models.

The classification tree models predicted risk and need scores for unassessed offenders on the basis of the relationship between offender characteristics and the actual risk and need scores of assessed offenders. There is an underlying assumption that the decision to assess can be treated as random for an offender with a given set of circumstances, but it is possible that non-random factors were involved. Also, the models predicted risk and need scores for all assessments nationally without taking into account any variations in assessment practices at the local/regional level. Nor was any attempt made to standardise actual scores by taking into account any such inconsistencies.

Results

OASys coverage

Analysis of the OMSAS form 20 commencements extract revealed that there were 231,143 probation commencements during the period January to December 2007. As shown by Table 9.2, OASys assessments had been completed in 175,559 or 76% of these cases. The coverage of the OASys data thus corresponded to the lowest of the four levels (less than 80%) set out within the directory of clinical databases. In 165,830 or 72% of the cases, the assessments were valid in terms of having sufficient data completion to be used for profiling the offenders' risks and needs. At the regional level, coverage varied from 59% in London to 93% in the North East. The difference between these two regions was more pronounced in relation to the completion of valid assessments; 49% in London and 92% in the North East.

As noted above, full OASys assessments should have been completed post-sentence in the community for all those cases designated at OM Tier 2 and above, with the exception of those Tier 2 cases in which there was a stand-alone unpaid work requirement. During 2007, 158,579 (67%) of the offenders were known to meet these post-sentence eligibility criteria and assessments had been completed in 85% of these cases.⁶² At the regional level, the completion rate ranged from 67% in London to 94% in the North East. For two of the regions, the completion rate for these eligible cases was at least 90%, falling within the third level of coverage set out within the directory of clinical databases, with all other regions except London falling within level two. Of the 59,020 (26%) offenders who were known not to meet the post-sentence eligibility for OASys, assessments had been completed in 54% of cases (with valid assessments in 46% of cases).⁶³ In 30% of these cases, the assessments had been completed pre-sentence. As noted previously, the use of OASys at the pre-sentence stage is linked to the type of pre-sentence report with all standard delivery reports requiring an OASys assessment.

⁶² Eligibility was unknown in 13,544 (5.9%) of the cases.

⁶³ In the North East, an assessment had been completed in 89% of the non-eligible cases.

Table 9.2: OASys completion rates by post-sentence eligibility and region (probation commencements)

Region	OASys post-sentence eligibility					
	Non-eligible (n=59,020)		Eligible (n=158,579)		All cases (n=231,143)	
	% with OASys	% with valid OASys	% with OASys	% with valid OASys	% with OASys	% with valid OASys
North West (n = 37,428)	55%	49%	89%	86%	79%	75%
North East (n = 14,410)	89%	87%	94%	94%	93%	92%
Yorkshire and Humberside (n = 26,336)	48%	45%	88%	87%	76%	75%
East Midlands (n = 18,808)	62%	59%	90%	89%	81%	80%
East of England (n = 19,468)	50%	46%	87%	84%	74%	71%
West Midlands (n = 29,061)	61%	43%	86%	81%	78%	71%
South East (n = 26,941)	47%	44%	86%	85%	73%	71%
South West (n = 14,036)	69%	58%	87%	85%	79%	75%
London (n = 30,555)	44%	23%	67%	61%	59%	49%
Wales (n = 14,100)	55%	50%	87%	85%	78%	75%
Total (n = 231,143)	54%	46%	85%	82%	76%	72%

Focusing upon those 175,599 cases in which assessments had been completed, Table 9.3 sets out the validity rates of these assessments. As can be seen, 94% of the assessments were valid in terms of their data completion. The validity rate of those assessments which met the post-sentence eligibility criteria was 97%, 12 percentage points higher than the rate for those assessments which did not meet the eligibility criteria.⁶⁴ The greatest difference was observed for London, where the validity rate fell from 91% for the eligible cases to 52% for the non-eligible cases, demonstrating a tendency to partially complete OASys in the latter cases.

⁶⁴ Eligibility was unknown in 8,470 (4.8%) of the cases in which an assessment had been completed.

Table 9.3: Validity rates of completed OASys assessments by post-sentence eligibility and region (probation commencements)

Region	n	% valid assessments by post-sentence eligibility		
		Non-eligible (n=32,079)	Eligible (n=135,050)	All cases (n=175,599)
North West	29,623	89%	97%	95%
North East	13,378	98%	99%	99%
Yorkshire and Humberside	20,121	95%	99%	98%
East Midlands	15,296	95%	99%	98%
East of England	14,397	91%	97%	95%
West Midlands	22,730	70%	95%	91%
South East	19,788	92%	98%	97%
South West	11,063	84%	98%	95%
London	18,177	52%	91%	82%
Wales	11,026	91%	98%	96%
Total	175,599	85%	97%	94%

Analysis of the prisons receptions extract revealed that there were 84,708 sentenced prisoner receptions during the period January to December 2007. As shown by Table 9.4, OASys assessments had been completed in 56,294 or two-thirds (66%) of these cases. The coverage of the OASys data thus corresponded to the lowest of the four levels (less than 80%) set out within the directory of clinical databases. In 51,325 or approximately three-fifths (61%) of the cases, the assessments were valid in terms of having sufficient data completion to be used for profiling offenders. At the regional level, coverage varied from less than half (46%) in London to approximately four-fifths (81%) in the North East. The difference between these two regions was more pronounced in relation to the completion of valid assessments; less than two-fifths (38%) of the London cases had valid assessments compared to nearly four-fifths (78%) of the North East cases.

As noted above, all those offenders aged 18 to 20 and all older offenders serving a custodial sentence of at least 12 months should have been assessed post-sentence. During 2007, 38,013 (45%) of the offenders met these post-sentence eligibility criteria and assessments had been completed in nearly four-fifths (78%) of these cases. At the regional level, the completion rate ranged from 58% in London to 89% in both the North East and Wales. For six of the regions, the completion rate for these eligible cases was at least 80%, falling within the second level of coverage set out within the directory of clinical databases. Of the remaining 46,695 (71%) offenders who did not meet the post-sentence eligibility for OASys, assessments had been completed by probation/prison assessors in 57% of cases (with valid assessments in 51% of cases). In 29% of these cases, the assessments had been completed pre-sentence.

Table 9.4: OASys completion rates by post-sentence eligibility and region (prison receptions)

Region	OASys post-sentence eligibility					
	Non-eligible (n=46,695)		Eligible (n=38,013)		All cases (n=84,708)	
	% with OASys	% with valid OASys	% with OASys	% with valid OASys	% with OASys	% with valid OASys
North West (n=15,915)	62%	56%	85%	79%	70%	64%
North East (n=4,704)	74%	70%	89%	85%	81%	78%
Yorkshire and Humberside (n=8,588)	69%	65%	82%	78%	76%	71%
East Midlands (n=6,791)	65%	61%	86%	81%	76%	71%
East of England (n=8,201)	56%	50%	79%	72%	68%	62%
West Midlands (n=5,535)	60%	52%	82%	76%	71%	64%
South East (n=11,707)	53%	48%	69%	63%	61%	55%
South West (n=6,182)	59%	54%	78%	73%	66%	61%
London (n=13,665)	37%	29%	58%	52%	46%	38%
Wales (n=3,408)	68%	62%	89%	83%	78%	72%
Total (n=84,708)	57%	51%	78%	72%	66%	61%

Focusing upon those 56,294 cases in which assessments had been completed, Table 9.5 sets out the validity rates of these assessments. As can be seen, 91% of the assessments were valid in terms of their data completion. The difference in data quality between those assessments which met the post-sentence eligibility criteria and those which did not meet the criteria was fairly small: validity rates of 93% and 90% respectively. The greatest difference was observed for London, where the validity rate fell from 89% for the eligible cases to 78% for the non-eligible cases.

Table 9.5: Validity rates of completed OASys assessments by post-sentence eligibility and region (prison receptions)

Region	n	% valid assessments by post-sentence eligibility		
		Non-eligible (n=26,825)	Eligible (n=29,469)	All cases (n=56,294)
North West	11,212	90%	93%	91%
North East	3,819	96%	96%	96%
Yorkshire and Humberside	6,495	94%	94%	94%
East Midlands	5,129	93%	94%	94%
East of England	5,586	89%	92%	91%
West Midlands	3,911	87%	92%	90%
South East	7,112	90%	91%	91%
South West	4,096	91%	93%	92%
London	6,262	78%	89%	84%
Wales	2,664	91%	93%	92%
Total	56,294	90%	93%	91%

OASys representativeness

Table 9.6 sets out the OASys completion rates for different offender sub-groups within the 2007 probation commencements.⁶⁵ When using logistic regression to account for the relationships between the variables, there were significant differences in OASys completion rates across all offender characteristics.⁶⁶ The odds ratios, set out in the final column of Table 9.6, are an indication of effect size, grouping the offenders by all of their characteristics and comparing the odds of OASys completion between the groups. In this instance, odds ratios of more than one indicate that OASys completion was more likely for offenders within the sub-group compared to offenders within the designated reference group. As shown, the odds of OASys completion were:

- seven times higher for offenders at OM Tiers 3 and 4 compared to offenders at Tier 1;
- 1.7 times higher for offenders subject to a Suspended Sentence compared to offenders subject to a Community Sentence;
- 1.5 times higher for offenders who had committed an offence classified as ‘violence against the person’ compared to offenders who had committed an offence classified as ‘other’; and
- 1.4 times higher for offenders with a high OGRS 2 likelihood of reconviction score compared to offenders with a low OGRS 2 likelihood of reconviction score.

These differences demonstrate selective targeting of OASys. Consequently, the representativeness of the OASys data corresponded to the lowest of the four levels (no evidence or unlikely to be representative) set out within the directory of clinical databases.⁶⁷

⁶⁵ While data completion was fairly good for these independent variables (incorporating OASys data where possible), a valid OGRS 2 score was unknown in 14% of cases.

⁶⁶ The logistic regression model had a chi-square value of 19006.975 with 24 degrees of freedom and a significance level of .000.

⁶⁷ The representativeness of the data at the regional level is considered in Appendix 12.

Table 9.6 also demonstrates that differences remained in OASys completion rates across all offender characteristics except gender when the analysis was restricted to those cases which met the post-sentence eligibility criteria and to those cases which did not meet the criteria.⁶⁸ In the latter cases, the odds of OASys completion were 3.8 times higher for offenders who had committed a 'sexual offence' compared to offenders who had committed an offence classified as 'other', and 3.2 times higher for offenders subject to a custodial sentence compared to offenders subject to a community sentence. The magnitude of these odds ratios demonstrate that OASys completion was being targeted at specific types of non-eligible case.

Focusing upon those probation commencements with completed assessments, Table 9.7 demonstrates that there were significant differences in the validity rates of these assessments across all offender characteristics.⁶⁹ As shown, the odds of valid OASys completion were:

- 4.8 times higher for offenders at OM Tier 4 compared to offenders at Tier 1; and
- 2.9 times higher for offenders with a high OGRS 2 likelihood of reconviction score compared to offenders with a low OGRS 2 likelihood of reconviction score.

These odds ratios suggest that greater efforts were being made to complete OASys fully in those high risk cases in which greater resources were being invested. Perhaps more surprisingly, the odds of the OASys assessments for Black offenders being valid were about one-third (34%) the odds of the assessments for White offenders being valid. This finding can be largely explained by the relatively low validity rate of OASys completions in London; 48% (n=13,595) of all Black offenders in the OMSAS form 20 commencements extract were being supervised within London.

Table 9.7 further demonstrates that many of the differences in OASys validity rates were particularly pronounced when restricting the analysis to those cases which did not meet the post-sentence eligibility criteria, suggesting that full OASys completion was being targeted at specific types of non-eligible case.⁷⁰ The odds of valid OASys completion were:

- 3.7 times higher for offenders who had committed an offence classified as 'robbery' compared to offenders who had committed an offence classified as 'other';
- 3.6 times higher for offenders with a high OGRS 2 likelihood of reconviction score compared to offenders with a low OGRS 2 likelihood of reconviction score; and
- 2.6 times higher for offenders subject to a Custodial Sentence compared to offenders subject to a Community Sentence.

68 The logistic regression model for the non-eligible cases had a chi-square value of 3259.073 with 22 degrees of freedom and a significance level of .000, while the model for the eligible cases had a chi-square value of 654.622 with 23 degrees of freedom and a significance level of .000.

69 The logistic regression model had a chi-square value of 5518.151 with 24 degrees of freedom and a significance level of .000.

70 The logistic regression model for the non-eligible cases had a chi-square value of 1187.364 with 22 degrees of freedom and a significance level of .000. The model for the eligible cases had a chi-square value of 1078.958 with 23 degrees of freedom and a significance level of .000.

Table 9.6: OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements)

	OASys post-sentence eligibility											
	Non-eligible (n=59,020)				Eligible (n=158,579)				All cases (n=231,143)			
	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio
Gender												
	Male	^		85.1%	^		76.1%	^		^		
	Female	n/s		85.7%	n/s		74.9%	***		***		1.073
Age groups												
	18-20	^		85.5%	^		76.3%	^		^		
	21-24	*	1.078	85.5%	***	.862	75.8%	*		*		.946
	25-40	***	1.135	85.3%	***	.888	76.6%	n/s		n/s		
	41+	***	1.185	84.0%	***	.843	74.1%	n/s		n/s		
Ethnicity												
	White	^		87.3%	^		78.4%	^		^		
	Black	***	1.465	75.2%	**	.878	67.6%	*		*		1.067
	Asian	***	1.406	81.6%	n/s		72.7%	***		***		1.162
	Mixed	***	1.297	81.7%	n/s		75.2%	n/s		n/s		
	Other	n/s		64.0%	*	.798	56.2%	n/s		n/s		
Offence category												
	Violence against the person	***	1.689	87.1%	***	1.203	80.0%	***		***		1.453
	Sexual offences	**	3.814	85.9%	n/s		85.3%	***		***		1.337
	Burglary	***	1.485	85.7%	***	1.155	81.4%	***		***		1.309
	Robbery	***	2.445	81.6%	*	1.208	81.1%	***		***		1.368
	Theft and handling	***	1.225	84.7%	n/s		77.0%	***		***		1.147
	Fraud and forgery	*	1.086	78.8%	n/s		63.0%	n/s		n/s		
	Criminal damage	***	1.484	86.7%	**	1.168	78.4%	***		***		1.371
	Drug offences	***	1.445	81.3%	n/s		75.1%	***		***		1.314
	Other offences	^		84.6%	^		70.5%	^		^		
OGRS band												
	Low	^		93.0%	^		81.6%	^		^		
	Medium	***	1.262	92.6%	**	.921	86.8%	***		***		1.120
	High	***	2.414	93.3%	n/s		91.9%	***		***		1.425

Table 9.6: OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements)
(continued)

		OASys post-sentence eligibility											
		Non-eligible (n=59,020)				Eligible (n=158,579)				All cases (n=231,143)			
		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio	
OM Tier	1	51.4%	^		-				51.4%	^			
	2	65.2%	***	2.067	84.6%	^		80.0%	***	4.666			
	3	-			85.2%	***	1.325	85.2%	***	7.015			
	4	-			86.2%	***	1.480	86.2%	***	6.873			
Sentence	Community sentence	50.4%	^		86.3%	^		73.1%	^				
	Custody/YOI	70.3%	***	3.204	78.6%	**	.613	76.1%	***	.919			
	Suspended sentence	69.9%	***	2.403	88.8%	**	1.216	84.4%	***	1.743			
	Other	70.3%	**	5.178	84.9%	n/s		82.2%	n/s				

^ Used as reference group within logistic regression. Asterisks indicate whether groups differ significantly (confidence levels * < .05, ** < .01, *** < .001).

Table 9.7: Validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements)

		OASys post-sentence eligibility											
		Non-eligible (n=32,079)				Eligible (n=135,050)				All cases (n=175,599)			
		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio	
Gender	Male	84.7%	^		96.9%	^			94.5%	^			
	Female	83.8%	n/s		96.6%	*	1.138		93.8%	***	1.193		
Age groups	18–20	86.6%	^		97.1%	^			94.9%	^			
	21–24	85.5%	n/s		97.3%	***	1.222		94.9%	***	1.146		
	25–40	83.7%	n/s		97.0%	***	1.196		94.6%	***	1.150		
	41+	82.9%	n/s		95.7%	n/s			92.9%	n/s			
	White	87.5%	^		97.4%	^			95.6%	^			
Ethnicity	Black	63.9%	***	.324	92.6%	***	.408		85.8%	***	.342		
	Asian	79.5%	***	.662	95.1%	***	.701		90.5%	***	.658		
	Mixed	78.8%	***	.564	95.8%	***	.604		92.8%	***	.567		
	Other	74.8%	***	.539	89.2%	***	.335		84.2%	***	.403		
	Violence against the person	86.9%	***	1.469	97.2%	***	1.480		95.3%	***	1.437		
	Sexual offences	94.3%	n/s		96.9%	n/s			96.7%	n/s			
Offence category	Burglary	89.5%	***	1.494	97.9%	***	1.448		96.9%	***	1.459		
	Robbery	95.5%	*	3.684	97.9%	***	1.847		97.8%	***	1.919		
	Theft and handling	84.7%	n/s		97.1%	*	1.131		95.1%	**	1.104		
	Fraud and forgery	79.7%	n/s		95.2%	n/s			88.6%	n/s			
	Criminal damage	87.7%	***	1.421	97.4%	***	1.446		95.5%	***	1.425		
	Drug offences	88.8%	***	1.648	96.8%	***	1.442		95.3%	***	1.626		
	Other offences	81.5%	^		95.7%	^			92.1%	^			
OGRS band	Low	85.4%	^		96.2%	^			92.9%	^			
	Medium	88.7%	***	1.398	97.2%	***	1.401		95.8%	***	1.448		
	High	95.5%	***	3.598	98.6%	***	2.630		98.3%	***	2.920		

Table 9.7: Validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements) (continued)

		OASys post-sentence eligibility									
		Non-eligible (n=32,079)			Eligible (n=135,050)			All cases (n=175,599)			
		% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	
OM Tier	1	85.1%	^		-			85.1%	^		
	2	83.2%	***	.849	96.1%	^		94.1%	***		2.279
	3	-			96.9%	**	1.124	96.9%	***		3.601
	4	-			98.3%	***	1.620	98.3%	***		4.740
Sentence	Community sentence	82.4%	^		96.5%	^		93.1%	^		
	Custody/YOI	93.2%	***	2.618	97.3%	**	1.153	96.5%	***		1.579
	Suspended sentence	90.1%	***	1.976	97.2%	***	1.310	96.0%	***		1.591
	Other	96.2%	n/s		98.4%	n/s		97.5%	**		2.958

^ Used as reference group within logistic regression. Asterisks indicate whether groups differ significantly (confidence levels * < .05, ** < .01, *** < .001).

Turning to the 2007 sentenced prisoner receptions, there were significant differences in OASys completion rates across all offender characteristics set out in Table 9.8 except gender.⁷¹ As shown, the odds of OASys completion were:

- 2.4 times higher for offenders categorised as ‘YOI open’ or ‘YOI short sentence’ compared to Category C offenders;
- 2.4 times higher for offenders sentenced to at least one year in custody but less than four years, compared to offenders sentenced to less than one year;
- 4.8 times higher for offenders whose ethnicity was recorded as ‘White’ compared to offenders whose ethnicity was recorded as ‘other’; and
- 1.7 higher for offenders who had committed a ‘sexual offence’ compared to offenders who had committed an offence classified as ‘other’, and 3.1 times higher for offenders who had committed an offence classified as ‘other’ compared to offenders who had committed an offence classified as ‘fraud and forgery’.

Once again, therefore, the representativeness of the OASys data corresponded to the lowest of the four levels (no evidence or unlikely to be representative) set out within the directory of clinical databases. With regard to ethnicity, the difference was largely explained by the relatively low OASys completion rate in London; 36% (n=1,078) of all offenders in the prisons receptions extract whose ethnicity was recorded as ‘other’ were in London establishments. Table 9.8 also demonstrates that differences remained in OASys completion rates across all offender characteristics except gender when the analysis was restricted to those cases which met the post-sentence eligibility criteria and to those cases which did not meet the criteria.⁷²

71 While data completion was fairly good for these independent variables (incorporating OASys data where possible), the security classification was unknown in 26% of cases. The logistic regression model had a chi-square value of 5612.780 with 25 degrees of freedom and a significance level of .000.

72 The logistic regression model for the non-eligible cases had a chi-square value of 1496.371 with 21 degrees of freedom and a significance level of .000, while the model for the eligible cases had a chi-square value of 2281.860 with 25 degrees of freedom and a significance level of .000.

Table 9.8: OASys completion rates by post-sentence eligibility and offender characteristics (prison receptions)

		OASys post-sentence eligibility														
		Non-eligible (n=46,695)					Eligible (n=38,013)					All cases (n=84,708)				
		% with OASys	Sig.	Odds ratio	% with OASys	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	
Gender	Male	57.6%	^		77.8%	^		66.9%	^							
	Female	56.3%	n/s		74.1%	n/s		62.3%	n/s							
Age groups	18–20	-			82.9%	^		82.9%	^							
	21–24	63.5%	^		78.7%	n/s		68.8%	**				.693			
	25–40	56.9%	***	.813	74.4%	n/s		62.5%	***				.569			
	41+	52.1%	***	.674	70.6%	**		59.3%	***				.467			
	White	61.2%	^		81.7%	^		70.2%	^							
Ethnicity	Black	39.7%	***	.524	60.0%	***		50.1%	***				.468			
	Asian	45.4%	***	.619	68.8%	***		56.2%	***				.617			
	Mixed	52.7%	***	.723	73.2%	***		63.9%	***				.675			
	Other	15.8%	***	.169	45.9%	***		29.5%	***				.207			
	Violence against the person	62.3%	***	1.278	81.7%	n/s		71.7%	***				1.248			
	Sexual offences	64.7%	***	1.574	80.4%	***		1.422	***				1.683			
	Burglary	62.5%	***	1.181	80.7%	n/s		74.1%	**				1.111			
Offence category	Robbery	64.6%	***	1.634	81.6%	**		1.214	***				1.395			
	Theft and handling	58.6%	n/s		80.1%	**		.830	n/s							
	Fraud and forgery	26.2%	***	.353	47.4%	***		.249	***				.327			
	Criminal damage	57.4%	n/s		87.2%	n/s		65.1%	n/s							
	Drug offences	61.6%	***	1.392	68.0%	***		.699	*				.913			
	Other offences	57.5%	^		81.6%	^		64.9%	^							

Table 9.8: OASys completion rates by post-sentence eligibility and offender characteristics (prison receptions)
(continued)

	OASys post-sentence eligibility											
	Non-eligible (n=46,695)				Eligible (n=38,013)				All cases (n=84,708)			
	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio
Sentence length	< 1 year	57.4%			82.7%	^		61.4%	^			
	1 year + , < 4 years	-			78.6%	n/s		78.6%	***		2.388	
	4 years +	-			66.9%	***	.562	66.9%	***		1.362	
	Life	-			75.8%	*	.767	75.8%	***		1.821	
Security category	Category A	57.7%	n/s		52.8%	***	.374	53.3%	***		.448	
	Category B	57.7%	*	.860	68.9%	***	.703	65.3%	***		.797	
	Category C	59.6%	^		75.7%	^		65.5%	^			
Security category	Category D	59.8%	n/s		78.9%	***	1.335	64.3%	n/s			
	YOI closed	79.1%	n/s		84.9%	n/s		84.8%	***		1.537	
	YOI open/short sentence	100.0%	n/s		87.3%	n/s		87.3%	***		2.394	
Uncategorised	59.9%	n/s		73.3%	*	.885	68.5%	n/s				

^ Used as reference group within logistic regression. Asterisks indicate whether groups differ significantly (confidence levels * < .05, ** < .01, *** < .001)

Focusing upon those prison receptions with completed assessments, Table 9.9 demonstrates that there were significant differences in the validity rates of these assessments across offender sub-groups.⁷³ The odds of valid OASys completion were:

- 1.6 times higher for offenders categorised as 'YOI closed' compared to Category C offenders;
- 1.5 times higher for offenders sentenced to at least one year in custody but less than four years, compared to offenders sentenced to less than one year; and
- 3.6 times higher for White offenders compared to offenders categorised as 'other' ethnic classification.

Once again, the differences relating to ethnic classification were largely explained by the relatively low validity rate of OASys completions in London; 26% (n=318) of the 'other' ethnic offenders with completed assessments were in London establishments. Table 9.9 also demonstrates that differences remained in OASys validity rates when the analysis was restricted to those cases which met the post-sentence eligibility criteria and to those cases which did not meet the criteria.⁷⁴

73 The logistic regression model had a chi-square value of 444.670 with 25 degrees of freedom and a significance level of .000.

74 The logistic regression model for the non-eligible cases had a chi-square value of 221.671 with 21 degrees of freedom and a significance level of .000, while the model for the eligible cases had a chi-square value of 221.465 with 24 degrees of freedom and a significance level of .000.

Table 9.9: Validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (prison receptions)

	OASys post-sentence eligibility											
	Non-eligible (n=26,825)				Eligible (n=29,469)				All cases (n=56,294)			
	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio
Gender												
	Male	^		92.8%	^		91.3%	^				
	Female	n/s		90.1%	n/s		89.8%	n/s				
	18-20	-		93.4%	^		93.4%	^				
Age groups												
	21-24	^		93.5%	n/s		92.0%	n/s				
	25-40	*	.862	92.2%	n/s		90.7%	n/s				
	41+	***	.722	90.2%	n/s		88.6%	n/s				
Ethnicity												
	White	^		93.5%	^		92.1%	^				
	Black	***	.403	87.2%	***	.606	84.5%	***				.489
	Asian	***	.601	91.1%	n/s		88.8%	***				.691
	Mixed	***	.574	92.3%	n/s		90.3%	**				.705
	Other	***	.338	76.4%	***	.280	75.5%	***				.279
	Violence against the person	n/s		93.5%	n/s		91.9%	n/s				
	Sexual offences	n/s		91.4%	*	.779	91.3%	n/s				
	Burglary	*	1.228	93.8%	n/s		93.2%	n/s				
	Robbery	*	1.686	92.8%	n/s		92.9%	n/s				
Offence category												
	Theft and handling	n/s		93.6%	n/s		90.5%	n/s				
	Fraud and forgery	n/s		82.4%	***	.389	84.5%	**				.689
	Criminal damage	n/s		94.5%	n/s		90.2%	n/s				
	Drug offences	n/s		89.5%	***	.626	89.7%	*				.851
	Other offences	^		94.0%	^		91.1%	^				

Table 9.9: Validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (prison receptions)(continued)

	OASys post-sentence eligibility									
	Non-eligible (n=26,825)			Eligible (n=29,469)			All cases (n=56,294)			
	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	Sig.
Sentence length	< 1 year	89.6%		93.5%	^		90.4%	^		
	1 year + , < 4 years	-		92.9%	n/s		92.9%	***		1.467
	4 years +	-		90.3%	n/s		90.3%	*		1.156
	Life	-		91.6%	n/s		91.6%	n/s		
Security category	Category A	100.0%	n/s	93.9%	n/s		94.5%	n/s		
	Category B	90.6%	n/s	91.3%	n/s		91.1%	n/s		
	Category C	89.9%	^	92.4%	^		90.9%	^		
YOI closed	Category D	89.2%	n/s	92.1%	n/s		90.0%	n/s		
	YOI closed	97.1%	n/s	93.5%	n/s		93.6%	*		1.641
	YOI open/short sentence	100.0%	n/s	92.4%	n/s		92.5%	n/s		
Uncategorised	93.6%	**	1.546	92.1%	n/s		92.5%	*		1.177

^ Used as reference group within logistic regression. Asterisks indicate whether groups differ significantly (confidence levels * < .05, ** < .01, *** < .001)

Score adjustments

To predict national risk levels and criminogenic needs for the probation commencements, the offender characteristics within Table 9.6 were entered into classification decision tree models (see Appendix 13 for the relative variable importance and initial splits for each model).⁷⁵ The goodness-of-fit of the models was checked by comparing the actual valid and predicted outcomes (where both were known) and the accuracy of the models checked by calculating the percentages correctly predicted (see Appendix 14). The likelihood of reconviction band was correctly predicted in 69% of cases, with an under-representation of low-likelihood offenders of 0.1% and high-likelihood offenders of 0.2% and an over-representation of medium-likelihood offenders of 0.3%. The risk of serious harm level was correctly predicted in 68% of cases, with an under-representation of medium risk offenders of 0.2% and of high/very high risk offenders of 1.0% and an over-representation of low risk offenders of 1.2%. Across the criminogenic needs, the percentage correctly predicted ranged from 64% for emotional wellbeing to 77% for drug misuse. The greatest residual (the difference between the actual and predicted rate) was for financial management and income with an under-prediction of 0.9%.

National risk and need levels were then calculated, using actual valid OASys risk and need scores where available and using the predicted risk and need scores in the remaining cases. Compared to the 165,830 offenders with valid OASys assessments, the calculated national figures for all 231,143 offenders differed as follows:

- a decrease in the prevalence of all criminogenic needs, ranging from a decrease of 2.9% for financial management and income to a decrease of 8.0% for thinking and behaviour;
- an increase in the percentage of low likelihood of reconviction offenders of 4.7%; a decrease in the percentage of medium likelihood offenders of 1.4%; and a decrease in the percentage of high likelihood offenders of 3.4%; and
- an increase in the percentage of low risk of serious harm offenders of 6.8%; a decrease in the percentage of medium risk offenders of 5.6%; and a decrease in the percentage of high risk offenders of 1.1%.

As shown by Table 9.10, the calculated national figures indicated that 18% of the probation commencements had a high likelihood of reconviction and 6.1% presented a high/very high risk of serious harm. Criminogenic need prevalence rates ranged from 26% for financial management and income to 53% for thinking and behaviour. Focusing upon those lower OM Tier offenders who did not meet the post-sentence eligibility for administering OASys, 2.0% had a high likelihood of reconviction and 0.5% presented a high/very high risk of serious harm (n=59,020), demonstrating that they were a relatively low risk group. Consequently, if fewer assessments had been completed either pre- or post-sentence for such cases, greater adjustments towards lower risk and need levels would have been required.

⁷⁵ The offenders' exact ages and OGRS 2 scores were entered. The other offender characteristics were entered as categorical variables as set out in Table 9.10.

Table 9.10: Actual, predicted and adjusted risk and need levels by OASys post-sentence eligibility (probation commencements)

Risk and need level	Actual levels (valid OASys assessments) by post-sentence eligibility			Predicted levels (no valid assessments) by post-sentence eligibility			Adjusted national levels (all cases) by post-sentence eligibility		
	Non-eligible (n = 27,113)	Eligible (n = 130,797)	All cases (n = 165,830)	Non-eligible (n = 31,887)	Eligible (n = 27,782)	All cases (n = 65,313)	Non-eligible (n = 59,020)	Eligible (n = 158,579)	All cases (n = 231,143)
Likelihood of reconviction:									
Low	60.8%	21.4%	28.2%	68.8%	21.7%	45.0%	65.1%	21.4%	32.9%
Medium	35.3%	53.4%	50.4%	30.7%	59.7%	45.6%	32.8%	54.5%	49.0%
High	3.8%	25.3%	21.5%	0.5%	18.6%	9.4%	2.0%	24.1%	18.1%
Risk of serious harm:									
Low	65.3%	34.3%	40.0%	87.0%	39.6%	63.5%	77.0%	35.4%	46.8%
Medium	33.8%	57.1%	52.7%	13.0%	52.9%	33.3%	22.5%	56.1%	47.1%
High/Very high	0.9%	8.7%	7.2%	0.0%	7.5%	3.2%	0.5%	8.5%	6.1%
Criminogenic need:									
Accommodation	19.1%	42.4%	38.4%	1.5%	41.9%	21.5%	10.0%	42.4%	33.8%
Education, training & employability	36.0%	63.3%	58.6%	19.3%	57.5%	37.9%	27.3%	62.4%	52.8%
Financial management & income	15.1%	31.1%	28.6%	2.8%	34.6%	18.2%	8.7%	31.7%	25.7%
Relationships	20.0%	47.8%	42.7%	4.1%	43.4%	23.0%	11.6%	47.0%	37.2%
Lifestyle & associates	19.7%	50.8%	45.6%	7.2%	50.4%	28.2%	13.1%	50.7%	40.7%
Drug misuse	10.9%	34.3%	30.4%	3.7%	33.7%	18.5%	7.1%	34.1%	27.0%
Alcohol misuse	30.9%	51.4%	47.4%	10.4%	44.9%	27.6%	20.1%	50.1%	41.8%
Emotional wellbeing	22.1%	47.3%	42.7%	10.1%	42.5%	25.4%	15.7%	46.3%	37.7%
Thinking & behaviour	31.9%	67.7%	61.3%	3.9%	65.8%	32.8%	17.6%	67.1%	53.3%
Attitudes	18.6%	41.0%	37.1%	2.6%	34.1%	18.0%	10.3%	39.8%	31.8%

To predict national risk levels and criminogenic needs for the prison reception cases, the offender characteristics within Table 9.8 were entered into classification decision tree models (see Appendix 15 for the relative variable importance and initial splits for each model).⁷⁶ The likelihood of reconviction band was correctly predicted in 60% of cases, with an over-representation of low-likelihood offenders of 0.1% and of high-likelihood offenders of 0.1% and an under-representation of medium-likelihood offenders of 0.2% and of high-likelihood offenders of 0.2% (see Appendix 16). The risk of serious harm band was correctly predicted in 61% of cases, with an under-representation of medium risk offenders of 0.4% and of high/very high risk offenders of 0.6% and an over-representation of low risk offenders of 0.9%. Across the criminogenic needs, the percentage correctly predicted ranged from 61% for accommodation to 72% for both (i) education, training and employability and (ii) thinking and behaviour. The greatest residual (the difference between the actual and predicted rate) was for drug misuse with an under-prediction of 2.2%. The models were thus limited in terms of their accuracy, with generally lower percentages correctly predicted than for the probation commencements models. In the latter models, the offenders' OGRS 2 scores and OM Tier levels were the most important variables for splitting the cases (see Appendix 13) but neither of these variables were available for the prison receptions.

National risk and need levels were then calculated, using actual valid OASys risk and need scores where available and using the predicted risk and need scores in the remaining cases. Compared to the 51,325 offenders with valid OASys assessments, the calculated national figures for all 84,708 offenders differed as follows:

- a decrease in the prevalence of six of the ten criminogenic needs. The greatest reduction was for thinking and behaviour which fell by 4.1%;
- an increase in the percentage of low likelihood of reconviction offenders of 1.9%; a decrease in the percentage of medium likelihood offenders of 1.2%; and a decrease in the percentage of high-likelihood offenders of 0.8%;
- an increase in the percentage of low risk of serious harm offenders of 5.7%; a decrease in the percentage of medium risk offenders of 3.9%; and a decrease in the percentage of high risk offenders of 1.8%.

⁷⁶ The offenders' exact ages and sentence lengths were entered into the models, with the other offender characteristics entered as categorical variables as set out in Table 9.10. The number of previous convictions and number of previous custodial sentences were also entered, despite some concerns regarding the reliability of the variables, as they were found to improve the accuracy of the models.

Table 9.11: Actual, predicted and adjusted risk and need levels by OASys post-sentence eligibility (prison receptions)

Risk and need level	Actual levels (valid OASys assessments) by post-sentence eligibility			Predicted levels (no valid assessments) by post-sentence eligibility			Adjusted national levels (all cases) by post-sentence eligibility		
	Non-eligible (n=24,040)	Eligible (n=27,285)	All cases (n=51,325)	Non-eligible (n=22,655)	Eligible (n=10,728)	All cases (n=33,383)	Non-eligible (n=46,695)	Eligible (n=38,013)	All cases (n=84,708)
Likelihood of reconviction:									
Low	10.5%	18.3%	14.7%	15.2%	28.8%	19.6%	12.9%	21.3%	16.6%
Medium	45.6%	47.7%	46.7%	43.5%	44.6%	43.8%	44.5%	46.8%	45.5%
High	43.9%	34.0%	38.6%	41.3%	26.6%	36.6%	42.7%	31.9%	37.8%
Risk of serious harm:									
Low	35.2%	26.9%	30.8%	48.0%	38.0%	44.8%	41.6%	30.2%	36.5%
Medium	55.2%	53.9%	54.5%	47.3%	40.1%	45.0%	51.1%	49.9%	50.6%
High/Very high	9.6%	19.2%	14.7%	4.7%	21.9%	10.2%	7.2%	19.9%	12.9%
Criminogenic need:									
Accommodation	54.8%	48.3%	51.3%	59.5%	38.0%	52.6%	56.9%	45.8%	51.9%
Education, training & employability	72.4%	67.9%	70.0%	74.7%	53.3%	67.8%	73.2%	64.2%	69.1%
Financial management & income	44.4%	37.0%	40.5%	42.6%	38.3%	41.2%	43.5%	37.2%	40.7%
Relationships	55.8%	46.0%	50.6%	66.9%	25.5%	53.6%	60.6%	40.7%	51.7%
Lifestyle & associates	63.7%	65.4%	64.6%	56.6%	54.4%	55.9%	60.3%	62.5%	61.3%
Drug misuse	48.4%	38.9%	43.3%	43.6%	37.7%	41.7%	45.8%	38.3%	42.4%
Alcohol misuse	50.3%	46.9%	48.5%	43.5%	32.8%	40.1%	46.9%	42.8%	45.1%
Emotional wellbeing	52.6%	40.9%	46.4%	60.5%	26.5%	49.6%	55.5%	37.1%	47.2%
Thinking & behaviour	75.7%	72.6%	74.0%	69.8%	52.4%	64.2%	72.3%	66.9%	69.9%
Attitudes	57.5%	51.7%	54.4%	59.6%	33.8%	51.3%	57.7%	47.0%	52.9%

As shown by Table 9.11, the calculated national figures indicated that 38% of the prison receptions had a high likelihood of reconviction and 13% presented a high/very high risk of serious harm. Criminogenic need prevalence rates ranged from 41% for financial management and income to 70% for thinking and behaviour. Focusing upon those offenders who did not meet the post-sentence eligibility for administering OASys (aged over 20 and sentence of less than one year), 43% had a high likelihood of reconviction and 7.2% presented a high/very high risk of serious harm (n=46,695). In comparison, 32% of those offenders meeting the post-sentence eligibility had a high likelihood of reconviction and 20% presented a high/very high risk of serious harm (n=38,013). Consequently, if fewer assessments had been completed either pre- or post-sentence for offenders serving the shorter custodial sentences, the adjustments towards lower risk of serious harm levels would have increased, whereas the likelihood of reconviction adjustments would have been reversed towards higher levels.

Figure 9.1 compares the calculated national risk levels of the probation commencements and prison receptions during 2007, demonstrating that offenders in the latter group were more likely to present a high/very high risk of serious harm and/or have a high likelihood of reconviction. As Figure 9.2 demonstrates, the prison receptions group also had higher levels of need across all ten OASys sections measuring offending-related needs.

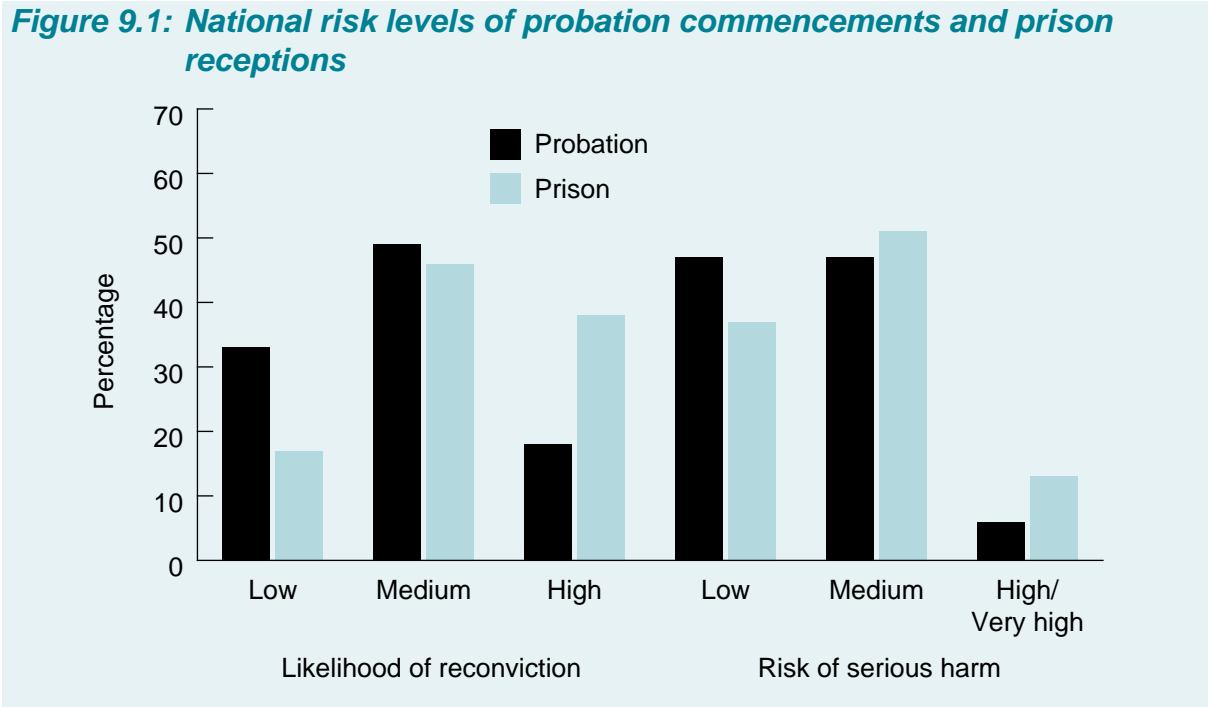
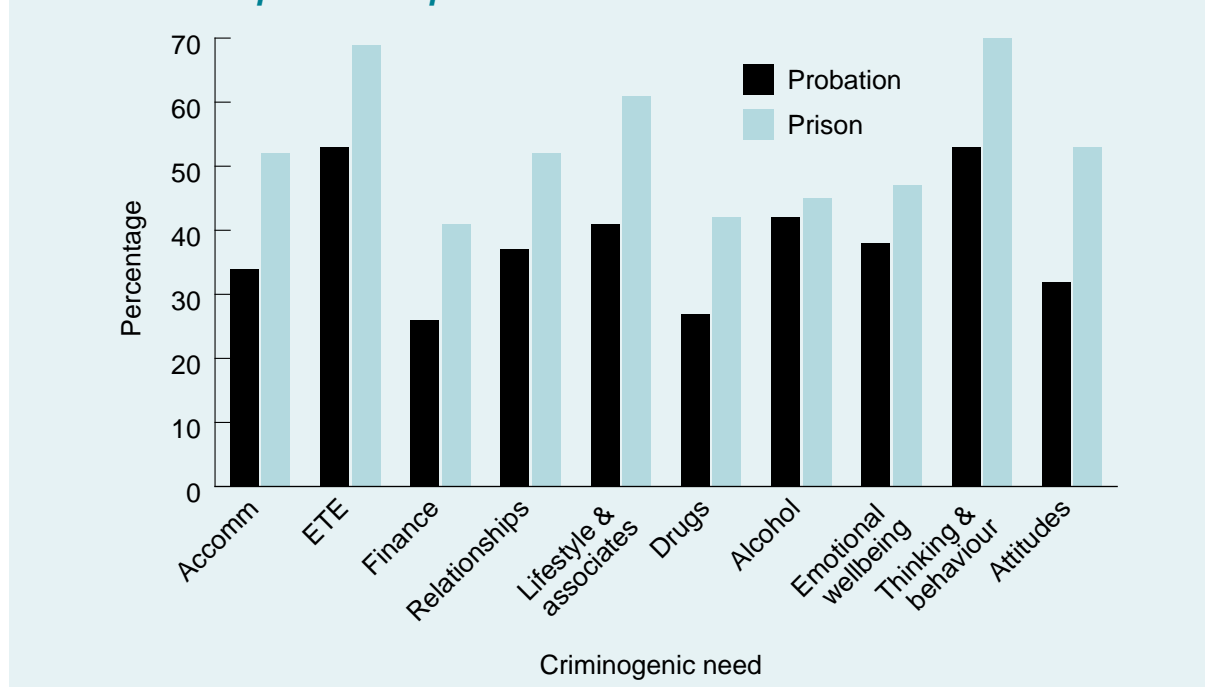


Figure 9.2: National criminogenic need levels of probation commencements and prison receptions



Implications

The research has the following implications for practitioners.

1. To improve the accuracy of offenders’ risk and need profiles, there is a need to ensure that OASys is completed when required. Completion of OASys will assist practitioners in making sound and evidence-based decisions in managing the risks posed by individual offenders.
2. It should also be ensured that OASys assessments are of sufficient quality for profiling offenders. As a minimum, each of the scored sections (1 to 12) within the core OASys assessment must have at least four-fifths of their scored items completed – ensuring that each criminogenic need is assessed properly. In the risk of serious harm component of OASys, the screening must be completed, the decision whether to complete a full risk analysis should be consistent with the information provided, and the four ratings of risk of harm in the community should be recorded in those cases in which a full analysis is required.

Implications for policy makers are as follows:

1. If commissioning and resource decisions are to be made upon the basis of OASys data, it needs to be recognised that OASys samples are not representative of the entire prison and probation caseloads and that adjustments are required towards lower overall risk and need levels.

2. When OASys data are disseminated, the adjustments in risk and need levels that might be needed for extrapolation to the wider caseloads should be indicated.
3. Any changes in the population targeted for assessment will affect the adjustments required to profile the full caseloads. For example, offenders serving short custodial sentences had higher overall likelihood of reconviction levels but lower overall risk of serious harm levels than those serving longer sentences. Any significant changes in the use of OASys with these two groups of offenders will result in different profiles of risk and need.
4. The ability to predict risk and need levels in non-assessed cases is restricted by the limited amounts of information collected in these cases. The recording of OGRS scores and OM Tier levels for all probation commencements and prison receptions would help to maximise the accuracy of the predictions.

Conclusion

The results suggest that while completion rates varied in line with the National Standards relating to the use of OASys, assessments were not always completed in eligible cases and were often completed in non-eligible cases. Not all assessments met the minimum standards of data completion required for profiling the risks and needs of offenders. Overall, however, offenders with an OASys were more likely to have committed a violent or sexual offence and to have a high likelihood of reconviction. The use of OASys was thus consistent with the expectation that resources should follow risk. In consequence, the risk and need levels of the complete probation and prison caseloads were lower than the risk and need levels of those for whom an assessment had been completed. The calculated national figures indicated that 18% of probation commencements had a high likelihood of reconviction and 6% presented a high/very high risk of serious harm, while 38% of prison receptions had a high likelihood of reconviction and 13% presented a high/very high risk of serious harm. Compared to the probation commencements, the prison receptions had higher levels of need across all ten OASys sections measuring offending-related need.

10. Exploratory research on the evidence boxes in the OASys 'relationships' and 'lifestyle/associates' sections

Introduction

The core OASys assessment identifies and classifies ten dynamic offending-related needs: accommodation; education; training and employability; financial management and income; relationships; lifestyle and associates; drug misuse; alcohol misuse; emotional wellbeing; thinking and behaviour; and attitudes.

Each of the ten sections contain fixed-response questions about an offending-related problem area which are scored, and free-text responses where assessors are asked to record information about their ratings to the fixed questions. There are also questions at the end of each section that allow the assessor to make a more clinical judgement regarding the links to: (i) offending behaviour; and (ii) the risk of serious harm, risks to the individual and other risks.

Previous research has concentrated on analysing the scored items and the quantitative content of OASys. An area that requires further exploration is the content of the free-text responses where assessors provide information and record evidence relevant to the section that has not been previously covered by the fixed responses. Exploring the textual information may yield some evidence of common themes for potentially revising and thus improving the reliability and validity of the core OASys assessment.

This chapter focuses on exploratory analysis of the information supplied in the free-text response boxes of OASys sections 6 and 7, addressing offenders' 'relationships' and 'lifestyle and associates'. This chapter presents the first piece of research on OASys using methods of qualitative data analysis. It addresses the two sections identified from other research on OASys reliability and validity as priorities for further development. Further similar exploratory research on other OASys sections covering different offending-related problems, are presented in the following chapter.

OASys section 6: Relationships

The OASys relationships section consists of the following seven fixed-response questions, the first six of which contribute to the criminogenic need score:

- 6.1 Current relationship with close family members (Relationships with parents, siblings, grandparents and any other family members/step family members with whom s/he has regular contact)
(0 = No problems; 1 = Some problems; 2 = Significant problems)

- 6.2 Close family member has criminal record
(0 = No; 2 = Yes)
- 6.3 Experience of childhood (Including any indication of physical, sexual or emotional abuse during childhood and adolescence)
(0 = No problems; 1 = Some problems; 2 = Significant problems)
- 6.4 Current relationship with partner (Level of support, mutual respect/affection, strength of relationship and difficulties. If currently single the level of satisfaction with state)
(0 = No problem; 1 = Some problems; 2 = Significant problems)
- 6.5 Current partner has criminal record
(0 = No; 2 = Yes)
- 6.6 Previous experience of close relationships (Quality, satisfaction of close relationships)
(0 = No problem; 1 = Some problems; 2 = Significant problems)
- 6.7 Evidence of domestic violence/partner abuse (Including threats and psychological abuse) (a. Victim/ b. Perpetrator)
(0 = No; 2 = Yes)

OASys section 7: Lifestyle and associates

The OASys lifestyle and associates section consists of the following five fixed-response questions, all of which contribute to the criminogenic need score and are scored **0 = No problems; 1 = Some problems; 2 = Significant problems**:

- 7.1 Community integration (Attachments to individual(s) or community groups. Participation in organised activities not linked to offending, including in prison, e.g. sports clubs, faith communities, etc.) **(Absence of any links = 2)**
- 7.2 Regular activities encourage offending (Do the leisure activities most commonly engaged in create opportunities to offend, or contribute to the need to offend e.g. gambling in prison?)
- 7.3 Easily influenced by criminal associates (Are most offences committed with others? When in the community does s/he spend a large amount of their time with other offenders?)
- 7.4 Manipulative/predatory lifestyle (Does s/he exploit others or abuse friendships, relationships, positions of trust? Does s/he use others, live off others without reciprocation, bully others?)

- 7.5 Recklessness and risk-taking behaviour (Lifestyle includes excessive thrill-seeking and risk-taking activities. Demonstrates intolerance for boring, unchallenging or unchanging situations. Needs excessive excitement or stimulation.)

Research reported earlier in this compendium, found that of the 11 scored sections within the core OASys assessment,⁷⁷ the relationships section did not have adequate internal reliability. In other words, the questions within the section were not measuring a discrete criminogenic need pertaining to relationships. Three of the questions (6.2, 6.4 and 6.5) were found to have low item-scale correlations (less than 0.3). Separate analysis of all start community/custodial sentence assessments during 2005/06 and 2006/07 revealed that, of all the scored questions, questions 6.2 and 6.5 were the least likely to be completed, suggesting that practitioners did not always have reliable information regarding the criminal records of partners and family members (see Chapter 3). Also, Howard, Clark, and Garnham (2006) found that whilst most of the criminogenic needs assessed by OASys were predictive of reconviction, the relationships section was not.⁷⁸

In terms of the lifestyle and associates section, analysis of OASys data found that it did not have adequate construct validity, that is, it did not appear to be measuring a single domain. More specifically, factor analysis revealed that two questions within the section (7.1: Community integration and 7.4: Manipulative/predatory lifestyle) did not fall into an underlying factor corresponding to lifestyle and associates but fell into employment and relationships factors respectively (see Chapter 4). Furthermore, only one question from the section (7.2: Regular activities encourage offending) is included within the new OASys predictors for general reoffending and violent reoffending (see Chapter 6). At present, therefore, the section has limited value in terms of predictive validity.

Previous research has concentrated on analysing the scored items and the quantitative content of OASys. An area that requires further exploration is the content of the free-text responses where assessors provide qualitative information for their ratings and record any evidence relevant to the section that has not been previously covered by the fixed responses. Exploring the textual information may yield some evidence of common themes for potentially revising and thus improving the reliability and validity of the 'relationships' and 'lifestyle/associates' sections and the construct validity of the core OASys assessment as a whole.

The key question this chapter seeks to address is:

What, if any, are the recurring themes within the textual information recorded in the OASys 'relationships' and 'lifestyle/associates' sections that are not covered in the fixed-response questions?

⁷⁷ Offending information is scored in addition to the ten dynamic criminogenic needs.

⁷⁸ When statistically controlling for the offenders' other needs.

Method

Sample

The OASys Data Evaluation and Analysis Team (O-DEAT) receive completed assessments from both the probation and prison services. The majority of assessments are completed by probation assessors and the predominance of such assessments is increasing with the roll-out of Offender Management and its requirement for assessments to be completed by community-based Offender Managers. For the purpose of this report, only probation assessments administered in June 2007 were reviewed. Furthermore, the assessments were restricted to those meeting the following standards of data completion.

- Each of the scored sections (1 to 12) within the core OASys assessment must have had at least four-fifths of their scored items completed – ensuring that each criminogenic need was assessed properly.
- In the risk of serious harm sections, the screening must have been completed, the decision whether to complete a full risk analysis should have been consistent with the information provided, and the four ratings of risk of serious harm in the community must have been recorded in those cases in which a full analysis was required.
- Some textual information must have been recorded in the evidence boxes of the (i) relationships and (ii) lifestyle and associates sections.

The samples were further restricted to one assessment (with earlier assessments taking precedence) in each individual ‘period of contact’. This de-duplication ensured that offenders could appear only once during a continuous period of supervision.

From these valid, de-duplicated assessments, random stratified samples of 300 assessments were extracted. As Table 10.1 shows, for each section, assessments were selected when:

- (i) there was no scored need but the assessor judged a link to offending; or
- (ii) there was a scored need but the assessor did not judge a link to offending.

It was thought that this approach would generate some additional themes pertaining to criminogenic or protective factors not currently recorded in the fixed responses.⁷⁹ The stratification also allowed for an adequate representation of different offender groups – the proportion of female and non-White⁸⁰ offenders being uplifted to guarantee sufficient numbers for analysis.

79 Protective factors have been defined as ‘those that moderate the effects of exposure to risk’ (Youth Justice Board, 2005).

80 Non-White offenders are those with a Black, Asian, Mixed or other ethnic classification. They were grouped together given the small numbers in the individual ethnic categories.

Table 10.1: The number and proportion of assessments for each stratum

Section score and assessor's judgement	Age		Ethnicity		Gender			
					Male (80%)		Female (20%)	
No scored need but assessor judged link to offending (50%)	18-24	(35%)	Non-White	(20%)	8	(3%)	11	(4%)
			White	(80%)	34	(11%)		
	25+	(65%)	Non-White	(20%)	16	(5%)	20	(6%)
			White	(80%)	62	(21%)		
Scored need but assessor judged no link to offending (50%)	18-24	(35%)	Non-White	(20%)	8	(3%)	11	(4%)
			White	(80%)	34	(11%)		
	25+	(65%)	Non-White	(20%)	16	(5%)	20	(6%)
			White	(80%)	60	(21%)		
Total							300 assessments	

Most of the 300 assessments showed extensive and comprehensive textual information. However, in the case of 16 assessments within the relationships section, the responses were truncated, as only 4,000 characters could pass between the probation and prison electronic systems. As a result, it was difficult to assess these responses given their lack of completeness.

Analysis

'Framework', an approach to qualitative data management developed by the National Centre for Social Research (NatCen), was used to analyse the textual information. This method involves the comprehensive and systematic analysis of qualitative data within a series of matrices or thematic charts, thus making it easier to identify recurrent key themes and allowing the accounts of different respondents or groups of respondents to be compared and contrasted (see Ritchie and Lewis (2003) for further explanation).

In this study, the rows in the matrix represented the individual assessments, whilst the columns represented the demographic, offence and risk-related information as well as the corresponding fixed response questions within the (i) relationships and (ii) lifestyle and associates sections (see appendices 17 and 18). The textual information from each sampled assessment was summarised into the appropriate cells. Responses that were not covered in the fixed responses were categorised into the 'other (additional themes)' cell.

Limitations

When noting information in the free-response text boxes, assessors may not have considered issues beyond those suggested by the fixed-response questions, thus wider aspects of the problem area may not have been recorded. It was evident that some assessors recorded information in one section that was more relevant to other sections in OASys, making it possible that details relevant to relationships and lifestyle and associates issues may have been recorded elsewhere. It is also possible that a larger sample may have led to more information and the emergence of additional themes.

Results

OASys section 6: Relationships

Within the textual information of the relationships section, the one recurrent theme which was not covered in the fixed response questions focused upon children. The following sub-themes were identified:

- parenthood;
- access to and contact with children;
- children as protective factors;
- relationships with children;
- child care and parenting; and
- single parent and coping with having children.

Parenthood: Offenders who were recorded as parents were more likely to have pre-school than teenage or grown-up children. The children were either from current or previous partners. Recurrent themes of parenthood included: planning to start a family with a current partner, partners having recently given birth or partners expecting their first or second child.

Access to and contact with children: The assessors' entries showed that access to and contact with children varied between offenders. There were those who were in new relationships but continued to have weekly contact and regular access to their children from previous relationships, whilst others wanted contact with their children but were refused by their ex-partners. The inability to maintain more frequent contact or the loss/refusal of contact caused some distress to offenders, including a suicide attempt following an ex-partner's refusal of contact. More generally, there were feelings of disappointment following the cancellation of supervised meetings, particularly in the case of offenders who had made the effort to initiate regular contact and whose children were under the care of local authorities. For the sake of their children, offenders would either 'please' or move nearer their ex-partners in order to see their children or they would resume their relationships with their ex-partners in order to 'work things through'.

In contrast, there were offenders who either had no further contact or only limited contact with their children who were living with their ex-partners. Some ex-partners had moved abroad and taken the children with them. Despite little contact, offenders continued to pay regular maintenance and had also taken on some parental responsibilities.

Children as protective factors: Children were considered one of the main motivating and stabilising factors in offenders' lives. They were the impetus to avoid offending, to abstain from drugs and to 'make a change' or for offenders to 'sort out' their lives. Spending time with their children and maintaining good working relationships with their current partners or ex-partners enabled offenders to change their lives for the better and to distance themselves from an offending lifestyle.

Relationships with children: Offenders generally described their relationships with their children as positive. Those with grown-up children, in particular, received a great deal of support. Similarly, those serving prison sentences also found regular visits from their partners and children beneficial. These offenders continued seeing their children following release even when their relationships with their partners had broken down. Reasons for valuing a positive relationship with family and children included offenders' own recollections of their unstable and traumatic childhood experiences.

Child care and parenting: Recurrent themes of child care and parenting included supporting partners with the care of their children, e.g. looking after them whilst their partners were at work or providing respite for their partners who would have been unable to cope. In terms of parenting, setting boundaries and standards for their children and improving parenting skills in order to be a positive influence on their children's behaviour were frequently mentioned.

Single parent and coping with having children: Offenders who were recorded as single parents were largely female and living in non-cohabitating relationships. These single mothers were either looking after children from different partners or they were caring for their children from a current relationship while receiving frequent visits, for example at weekends, from older children who were cared for by their ex-partners.

Gender differences

Generally, coping with children appeared to be more of an issue for female than male offenders. They were more likely to be raising two or more children on their own and to be experiencing great difficulties coping with their upbringing. Some of the children were in the care of local authorities/social services, having been placed on child protection registers due to inadequate supervision and neglect requiring medical attention. Others were either in foster care, had been put up for adoption or were being looked after by extended family members. Some children of women prisoners were not considered safe to return to their mothers following their release from prison. These offenders either had no formal contact with their children or were under supervision when visiting their children. However, some children had been taken off the child protection register and returned to the care of their mothers after they attended parenting classes and showed significant improvements in their parenting skills.

Women's difficulties coping with their children's upbringing appeared to be intrinsically linked with other issues such as alcohol dependency⁸¹ or poor mental health⁸² – some were experiencing suicidal thoughts⁸³ or had been hospitalised and prescribed antidepressants.⁸⁴

81 Quantitative analysis of the 300 assessments found that of the female offenders, 22% were assessed to have some/significant problems pertaining to current alcohol use (OASys 9.1).

82 56% were assessed to have some/significant psychological problems (OASys 10.2) and 22% were assessed to have some/significant psychiatric problems (OASys 10.6).

83 32% were reported to have self-harmed or attempted suicide (OASys 10.5).

84 13% were reported to have sought psychiatric treatment and 26% to have been prescribed medication to address their mental health problems (OASys 10.7 Psychiatric treatment; Medication).

Overall, the results showed a clear gender divide with female offenders more likely to experience difficulties coping with having children coupled with other issues relating to alcohol dependency and poor mental health. These findings appear to corroborate the research findings of others (Gelsthorpe, 1999; Sorbello *et al.*, 2002; Motiuk and Blanchette 2000). Male offenders, on the other hand, were more likely to report positive relationships with their children, supporting their partners with childcare and maintaining regular contact if separated.

OASys section 7: Lifestyle and associates

Within the textual information of the lifestyle and associates section, the only recurrent theme, not covered in the fixed response questions, related to gang involvement. Assessors recorded individual offenders as:

- having gang associations (*“He must be mixing with some serious criminal gangs if this was their response to failed business transactions but XXX is adamant that his diamond dealing is legitimate and not linked to criminals. He intends to return to his role of diamond valuator when he is released.”*);
- showing gang mentality (*“His offending behaviour is linked to a gang mentality that supports violent-related behaviour, revenge taking, and reckless behaviour”*); or
- being involved in some serious criminal gang activities (*“His previous offences have involved gang rivalry and XXX is known to this service as an offender with gang associations although the extent of this involvement is not yet known”*).

Other themes identified did not pertain to the lifestyle and associates section but to sections relating to employment, relationships, accommodation, drug misuse and attitudes. In terms of employment, meaningful full-time employment helped individual offenders to focus and engage in constructive activities. As a result, they had less leisure time to ‘hang around’ with their peers. Conversely, lack of employment resulted in boredom and socialising with the ‘wrong crowds’. In other words, employment was regarded as one of the motivators for offenders to avoid offending peers and situations.

In terms of relationships, offenders’ current partners and children were viewed as a positive motivator for offenders to change their criminal lifestyles. Spending all their free time with their partners and children helped offenders to distance themselves from their former offending peers. However, recorded evidence also showed that offenders with relationship problems at home were inclined to spend more time with their anti-social peers. Furthermore, single parents with driving offence convictions were seen as more likely to reoffend given their increased temptation to drive illegally.

Offenders' drug-taking lifestyles were compounded by difficulties relating to accommodation. Offenders were either homeless or reliant on their family members for accommodation. As a result, those who were homeless started to beg on the streets, as they were unable to claim benefits. Overall, moving into their own accommodation was considered a positive incentive 'to start afresh'.

Attending and participating in accredited offending behaviour programmes such as Think First (McGuire, 1995), Prison Partnership 12-step programme and Rehabilitation for Addicted Prisoners' Trust (RAPt) substance abuse treatment programme (see Martin, Player and Liriano (2003) for further information) helped individual offenders address their reckless/risk taking behaviour and acquire problem-solving skills in dealing with negative peer pressure. Assessors' records indicated that offenders generally enjoyed the programmes and had noticed a positive change in their behaviour.

The current results appear to corroborate Merrington and Skinns' findings (2000) that an offender's lifestyle and having pro-criminal friends are related to employment and accommodation. Similarly, factor analysis of the items within the core OASys assessment has found that two of the lifestyle and associates questions fall into underlying factors relating to employment and relationships (see Chapter 4). Although the information relating to accommodation, employment, and relationships is most relevant to those respective sections of OASys, their impact on an offender's lifestyle and associates explains the recording of these issues within the lifestyle and associates section.

Implications

OASys section 6: Relationships

Howard, Clark and Garnham (2006) reported in their evaluation of the OASys pilots that a question on offenders' relationships with their own children ('Relationship with child(ren) in parental role') was included in the first two pilots, but was subsequently removed given its relevance to fewer offenders than other questions. An item-by-item analysis, on the other hand, showed that the question relating to children was significantly associated with reconviction within 24 months. Furthermore, a question on 'looking after children' remains in the OASys self-assessment questionnaire – the only question that does not correspond to any question in the core OASys assessment (see Chapter 8).

The findings from the textual analysis demonstrate that there is some justification for recognising the issues surrounding children and introducing into the relationships section a new fixed-response question or questions relating to children, particularly for women offenders who are more often the primary caregivers. The textual analysis indicates that the key issues are: (i) whether offenders have parental responsibilities and (ii) whether there are any problems in their relationships with their children.

In light of the findings, a question on parental responsibility could be incorporated into OASys and the question 'relationship with child(ren) in parental role' which was previously removed from the pilot studies could be reinstated. The questions should focus on present parental responsibilities and current problems with their children. These questions could replace the current questions on close family members and current partners having a criminal record (6.2 and 6.5) which have been found to be problematic in studies of OASys completion rates and internal reliability (see Chapters 3 and 4).

The textual analysis also showed that the question on current relationship with partner (6.4) did not allow assessors to distinguish which offenders were in a relationship and those which were not. This question was also problematic in analysis of OASys construct validity (see Chapter 4). An alternative to removing the question, could be to split it according to status and quality of the current relationship.

The revised fixed-response questions in section 6 of OASys could be ordered as follows:

- 6.1 Experience of childhood
(0 = No problems; 1 = Some problems; 2 = Significant problems)

- 6.2 Previous experience of close relationships
(0 = No problems; 1 = Some problems; 2 = Significant problems)

- 6.3 Current relationship with close family members
(0 = No problems; 1 = Some problems; 2 = Significant problems)

- 6.4 Current relationship status
(0 = In a relationship; 2 = Single)

- 6.5 Quality of current relationship with partner
(0 = No problems; 1 = Some problems; 2 = Significant problems)

- 6.6 Parental responsibilities
(0 = No; 2=Yes)

- 6.7 Relationship with child(ren) in parental role
(0 = No problems; 1 = Some problems; 2 = Significant problems)

- 6.8 Evidence of domestic violence/partner abuse (a. Victim/ b. Perpetrator)
(0 = No; 2 = Yes)

OASys section 7: Lifestyle and associates

The findings from the textual analysis demonstrate that there is some justification for introducing into the lifestyle and associates section a new fixed-response question or questions relating to gangs, along with a clear definition for the term 'gang'. A focus on present gang membership rather than previous membership would be preferable.

Questions on gang membership and criminal associates could replace the current questions on community integration (7.1) and manipulative/predatory lifestyle (7.4), which have been found to be problematic in terms of the construct validity of the core OASys assessment (see Chapter 4), and neither of these questions are used in the new OASys predictors for general or violent reoffending (see Chapter 6). The revised five fixed-response questions in section 7 of OASys could be ordered as follows:⁸⁵

7.1 Has many criminal associates

(0 = No; 2 = Yes)

7.2 Easily influenced by criminal associates

(0 = No problems; 1 = Some problems; 2 = Significant problems)

7.3 Is a member of a gang

(0 = No; 2 = Yes)

7.4 Regular activities encourage offending

(0 = No problems; 1 = Some problems; 2 = Significant problems)

7.5 Recklessness and risk-taking behaviour.

(0 = No problems; 1 = Some problems; 2 = Significant problems)

Further research will be required to test the reliability and validity of the new questions alongside the existing OASys questions.

Conclusion

Exploratory analysis of the information recorded in the text box sections of the OASys sections on 'relationships' and 'lifestyles and associates', identified several themes that suggested ideas for amending the section content with additional questions. Inclusion of amended questions in new versions of OASys would need to be subjected to later analysis of reliability and validity, as data accumulate, in order to verify their value in the assessment of offending-related risks and needs.

⁸⁵ Offenders who are in gangs are likely to score highly on the revised 7.1, 7.2, and 7.4. However, many offenders will have a number of criminal associates without being members of gangs.

11. Exploratory research on the evidence boxes in eight OASys sections assessing offending-related problems

Introduction

This chapter follows on from the previous chapter which focused on exploratory analysis of the information supplied in the free-text response boxes of two OASys sections. The findings presented in this chapter are based on qualitative analysis of the remaining OASys dynamic risk factors: accommodation; education training and employment (ETE); financial management and income; drug misuse; alcohol misuse; emotional wellbeing; thinking and behaviour; and attitudes.

The key question this research sought to address was:

What, if any, are the recurring themes within the textual information recorded in each OASys section that are not covered in the current fixed-response questions?

Method

The sampling and method used in this research is identical to that used in the analysis of OASys sections 6 and 7, presented in the previous chapter. The reader is thus referred to the paragraphs under 'method' in Chapter 10 for details on method, sampling and limitations which similarly apply here.

Results

Section 3: Accommodation

Themes from the textual analysis

Within the evidence recorded in the sampled assessments for accommodation, a recurring theme was offenders with a history of domestic violence perpetration. There were issues with offenders returning to live in the same house as their victims or for potential new victims living in the same house. For instance, one offender was reported to live “*with his wife and daughter at the family home. Wife has allowed him to stay despite the assault and papers suggest she feels that he can change his behaviour if he gets specialist help. Nevertheless accommodation must be linked to a risk to his wife and possibly his daughter.*”

Information noted about another offender included: *“He tells me that it is his intention to return to live at the address with his partner after the bail conditions have been removed. He informs me that he has sought a reconciliation with his wife. In my assessment this will heighten the risk of further offending and I have notified the Domestic Violence Unit accordingly of his intentions and these risks will need to be monitored closely.”*

Some offenders did not have accommodation scored as a criminogenic need; however, their offence had been an acquisitive crime to cover household expenditure (for instance, mortgage payments or rent) and therefore assessors considered that the offender’s accommodation situation was linked to offending behaviour. For instance one assessment stated that *“The accommodation issues are related to offending as Mr X was in arrears with his mortgage and stole to alleviate this”* and another *“The home and upkeep of the family was subsidised by benefit fraud.”* This information is currently scored under the financial management and income section of OASys.

There were some assessments where the offender’s OASys score indicated accommodation to be a criminogenic need, although the assessor disagreed with this. For instance, one offender was pregnant and living temporarily with friends. The assessor remarked that it was important for permanent accommodation to be found, but that the unsatisfactory accommodation situation was not linked to the offender’s offending behaviour: *“This accommodation is uninhabitable and [she] has informed me that given the late stage of her pregnancy both her health visitor and social services have become involved in trying to find Ms X and her partner more appropriate accommodation.”* One assessment reported that *“[He] could live with his mother in the short term but she suffers from mental health issues and this is not suitable for him at this time. Accommodation is a priority, he needs his own self-contained flat which would provide him with security and a base to make further progress.”*

Implications for OASys

OASys question 3.4 (suitability of accommodation) includes consideration of whether the offender’s victim lives in the same house. However, the guidance does not currently include any consideration of whether the offender has future plans to move back in with their victim. The research evidence demonstrates a high rate of repeat offending for domestic violence (Hester and Westmarland, 2006). This supports considering the offender’s plans for future accommodation as there may be a risk of repeat offending in domestic violence cases. It would not be necessary to add an extra question to measure this theme, but instead the OASys guidance could be adapted.

Section 4: Education, Training and Employability (ETE)

Themes from the textual analysis

Within the evidence recorded in the sampled assessments for ETE, there was one main theme that is not currently covered in the entire OASys assessment. This related to the

appropriateness of employment or skills that the offender had. In some cases either education or employment provided the opportunity for offending behaviour. For instance, one offender's education in computers was felt to be a risk for future offending: *"He clearly has extensive technical knowledge about computers which in my view adds to the risk of further offending via the internet."* Another assessment reported concerns about the offender remaining working in the same industry that he had exploited to his advantage to commit offences: *"I raised concerns with Mr X that he has continued to work in the same trade for which the index offences were committed but he assures me that he has taken specific measures to manage his business better and avoid taking risks with potentially stolen cars."* A final example of this theme concerned offenders who had stolen from their place of work; for instance, one assessment stated: *"This offence is obviously clearly linked to her employment and represents a breach of trust which is exacerbated by the fact that she planned the offence and committed it over a long period of time taking large amounts of money."*

A second issue with the appropriateness of employment or offender skills was the ability of the offender to cope with their job. There were reports of stressful working environments, for instance: *"He describes this type of work as stressful and believes that the travelling and long hours have played a role in his current mental ill-health, increased alcohol intake and consequently his offending."*

The other main reason for a discrepancy between no scored criminogenic need and the assessor judging that ETE was linked to offending behaviour was that, for some offenders, being currently unemployed (but scores indicating no problems with education or employment skills) was enough to increase the likelihood of further offences. This was because of the way unemployment affected the offender, for instance causing boredom or increased alcohol/drug use. One assessment stated that: *"X acknowledges that boredom is a factor which influences his offending behaviour and believes if he were to obtain employment and re-establish a structure in his life his risk of further offending would reduce further."* Other assessments reported the impact of unemployment on alcohol consumption: *"When out of work tended to get bored and drink all day which puts him at higher risk of reoffending"* and drug use: *"Relates his loss of employment to the commencement of his offending as without a regular income he was no longer able to fund his entrenched drug dependency."*

Implications for OASys

None of the current OASys questions measure the **appropriateness** of the employment sector that the offender works in. A revised version of OASys might pilot a question on the appropriateness of an offender's employment or employment skills for future study or its value in assessing offending-related risks and needs.

Section 5: Financial management and income

Themes from the textual analysis

Within the evidence recorded in the sampled assessments for financial management and income, addiction was a common theme. Addiction was given as a reason for problems managing a budget and for committing crime to fund addictions. For instance, an offender was assessed with alcohol problems leading to financial difficulties: *“X has previously prioritised purchasing alcohol rather than budgeting. One of the index offences is theft – illegal earnings were some source of income at time as he stole to fund alcohol use.”*

Gambling was another addiction that impacted upon offenders’ ability to manage their finances without resorting to acquisitive crime, for instance: *“Measures that could decrease his risk of reoffending in the future would be for him to seek professional help in addressing his gambling problem to understand the thoughts, feelings and beliefs he has that contributed to his gambling addiction.”*

Many offenders reported problems with drug addiction that impacted on finances. For instance: *“the only time he has a problem with money is when he is attempting to fund his drug habit.”* Another reported that: *“X continues to take risk on his own part in offending due to lack of funds to assist him with his [drug] lifestyle”*. Some offenders were dealing drugs in order to fund their own drug use: *“Illegal earnings are an issue as displayed in the committal of the second drug offence where he dealt heroin to pay for his own habit.”*

Another theme mentioned were instances where the offender owed debts to criminal associates and had committed crimes to raise money to repay the debt. In some cases the debts had not been fully repaid and therefore there was a risk of further offending. One example was: *“Of concern is X’s financial position regarding debts to criminal associates (£6000+). He told me during interview that he committed the offence in an attempt to repay some of his debt and this clearly shows an inappropriate response to the situation. He tells me that having returned to live with his mother he will now seek to repay this debt by legitimate means as quickly as possible to remove any potential threats to his own safety.”* Another assessor wrote: *“He is culpable for the sale of the drugs on this sentence and it must be assumed that there may have been some financial reward for this. Although X states that he did it to pay off a debt.”*

Implications for OASys

OASys question 5.6 (severe impediment to budgeting) includes consideration of financial difficulties which includes debts to loan sharks and long-term gambling problems. The guidance advises assessors to not consider alcohol and drug abuse as these are covered in sections 8 and 9 of OASys. Specific questions on addictive demands that impact upon financial management or any other new questions under the financial management and income section do not appear necessary at present.

Section 8: Drug misuse

Themes from the textual analysis

Within the evidence recorded in the sampled assessments for drug misuse, cannabis use was regularly mentioned. Some offenders were using large quantities of cannabis (examples included £40 per day and £150 per week) which may be more problematic than the current weighting given by OASys for cannabis use. Aside from crime being committed to fund large cannabis use, there were also examples of cannabis causing psychological problems that had led to offending behaviour. For instance: *“Mr X self-reports previously smoking cannabis. He stated in interview that it was at this time that he began doing crazy stuff and robbing people and suffering from paranoia.”* and: *“Mr Y felt that his heavy use of cannabis in general and on the day of his offence in particular, did cloud [his] judgement.”* There were also other examples where offenders reported cannabis use that they did not feel was problematic. However, it is apparent that there are offenders for whom cannabis use contributes to offending behaviour.

Protective factors against continued drug use were also reported in this section. For instance, some offenders were scored to have a drug misuse criminogenic need; however, the assessors commented that pregnancy and children were motivational factors that would reduce drug use and offending behaviour. For instance: *“When she discovered she was pregnant she significantly reduced her drug intake and is not currently using any illicit drugs”*, whilst another reported that *“Prior to the birth of X, he and his partner were heavily using cocaine... once they found out they were going to have a baby this motivated them to stop immediately.”*

Another protective factor was a recent reduction or abstinence from drugs. For instance, some offenders were using drugs at the time of their offence and their drug behaviour was substantial enough to be scored with a drug criminogenic need; however, the offender had demonstrated a recent reduction. For instance: *“Previous offending is linked to problems with drug use in the past. She tells me that she has greatly reduced her drug use”* and: *“Using drugs was a contributory factor in his offending... since his arrest he has abstained from using drugs.”*

Implications for OASys

Children and pregnancy as a protective factor is a theme that also emerged from analysis of the textual data of the relationships section of OASys. Smith-Yau (Chapter 10) recommended that questions to capture if the offender has childcare or caring responsibilities should be included, alongside whether the offender had the capabilities to match their responsibilities.

In summary, the analysis of the textual data does not point to a need to include additional questions in OASys. Cannabis use is not problematic for all offenders and for those where it is a serious problem, this should be captured by the other questions in the drug misuse section (e.g. frequency of use, violent behaviour related to drug use, drug use and obtaining drugs a major occupation). The two protective factors identified would be covered by the proposal for the changes to the relationships section and by the frequency of use question already in OASys.

Section 9: Alcohol misuse

Themes from the textual analysis

Within the evidence recorded in the sampled assessments for alcohol misuse, one theme that was apparent was some offenders' use of alcohol to cope with emotional problems. For instance, one assessment said: *"When faced with an emotional situation X engages in a binge-drinking session in an attempt to deal with the situation."* Another assessment reported: *"His alcohol intake has increased recently which he states is as a result of his using alcohol as a means to deal with his feelings regarding the breakdown of his relationship and associated difficulties."* This may have an effect on the ability and motivation to reduce alcohol use.

There were other examples of offenders who were not scored as having an alcohol offending-related need, but who had a history of offences resulting from alcohol consumption, namely driving whilst under the influence. For instance, one assessment reported: *"[He] tells me he is a social drinker... this having been his third appearance for drink driving it is clear that alcohol is linked to his offending behaviour."* Another assessment stated: *"He drinks when he goes out at the weekends with his friends and when he plays pool in the week. Clearly alcohol was related to both of X's two previous offences as they were both driving whilst under the influence. X stated no dissatisfaction with his current alcohol intake and suggested that his convictions result from poor decision making rather than alcoholism."*

Implications for OASys

None of the current OASys questions on alcohol use covers the reasons for alcohol use, such as the emotional response that emerged in the analysis of the text. Similarly, none of the scored questions covers whether alcohol use has led directly to offending behaviour. It may be useful to consider piloting the following questions within the alcohol misuse section:

1. Has the offender ever been convicted for a drink-driving offence?
2. Is alcohol used to escape from life, e.g. in response to emotional stress?

Section 10: Emotional wellbeing

Themes from the textual analysis

Within the evidence recorded in the emotional wellbeing section of the sampled assessments, offenders' mental health problems were associated with issues relating to:

- accommodation (*"It would appear that X had experienced difficulties relating to stable and secure accommodation and this impacted upon his wellbeing"*);
- employment (*"He told me that his feelings of depression are based around his current situation i.e. lack of housing and employment"*);
- relationships primarily relating to children (*"X described how she suffered from depression and anxiety for a number of years and this has been exacerbated by the adoption of her three sons"*);

- traumatic events such as the death of family members (*“The defendant’s grandmother had recently died and that X had enjoyed a close relationship with her. He tells me that his feelings of grief have not really been expressed until now and that they were also a contributing factor in these offences”*);
- changes in personal circumstances (*“[He was] overwhelmed by the change in his circumstances of being single to having a partner who was expecting their first child”*);
- issues specific to women such as termination of pregnancy (*“It appears that her depression originated from the termination of a pregnancy in November 2004”*).

Another recurrent theme (mirrored in section 9 – see above) was self-medication i.e. offenders using drugs (particularly cannabis) and/or alcohol to cope with emotional stress such as depression.

- *“X reports that he does not cope very well with emotional stress and so to some extent has been self-medicating over the years by using drugs.”*
- *“[He] has recently disclosed ongoing issues with depression that he states he was ashamed to admit and so attempted to self-medicate with alcohol.”*
- *“It would seem that he is now self-medicating with both alcohol and the ‘manics’... he states that he takes one of these tablets each day and that they help him deal with his depression and calm him down.”*

Alongside the self-medication theme, depression, psychosis and/or paranoia as a result of drug and/or alcohol consumption was also frequently mentioned.

- *“[He] admits to feeling depressed in the past when he was taking drugs.”*
- *“His previous crack cocaine and amphetamine use could potentially be responsible for his feelings of paranoia particularly in relation to authority figures.”*
- *“[He] states that since he stopped using drugs he is much more stable and feels that his depression was largely brought on by misusing drugs.”*

Implications for OASys

Similar to the alcohol misuse section, none of the current OASys questions on emotional wellbeing covers the reasons for the offender’s difficulties coping with emotional stress. One of the proposed new questions in the alcohol misuse section asks whether the offender uses alcohol in response to emotional stress. Consequently, it would not be necessary to add an extra question on the reasons for an offender’s emotional wellbeing problems as the themes identified above can be captured by questions already covered in other OASys sections.

Section 11: Thinking and behaviour

Themes from the textual analysis

Within the evidence recorded in the thinking and behaviour section of the sampled assessments, offenders' expression of remorse and regret was a recurring theme. This theme was also evident in the attitudes section.

"X demonstrated remorse for his actions and for his foul language. He seemed both embarrassed and ashamed that his life has reached this point."

"X expressed regret regarding his actions, stating on one occasion that if he were able to meet his victim, he would like to say how sorry he was and that it should not happen again."

Another identified theme was the minimisation by offenders of the severity of their offences. For example, *"X does minimise the offence stating he is not a dealer and the cannabis was for him and his brother."*

Some offenders also denied the effects of their offences on the victims.

"X seemed to minimise his behaviour by stating that the victim got their i-pod back so did not lose out but failed to acknowledge the potential psychological effects."

"X seemed dismissive suggesting the worst that could have happened would have been an injury to his victim's nose."

This theme was also identified in the attitudes section and is detailed further below.

Implications for OASys

A question on an offender's lack of remorse is already captured in question 2.6 (Does the offender recognise the impact and consequences of offending on victim, community/wider society?). No new questions are identified for inclusion at this time.

Section 12: Attitudes

Themes from the textual analysis

As mentioned in relation to section 11 above, minimisation of offence(s) was a recurring theme within the evidence recorded in the attitudes section. Assessors' entries largely related to the minimisation of offence(s) by domestic violence offenders. One example stated as follows:

"He does however continue to minimise his behaviour by asserting that her injuries were an accident and unintentional. His strong denial that he intentionally struck X leads me to believe he is using a number of 'techniques of neutralisation' such as victim blaming and a belief that he acted in self-defence in order to relieve himself of the stigma and guilt he would otherwise have felt as a result of assaulting a female."

Implications for OASys

It would not be necessary to include a question on an offender's denial and minimisation of offence(s) and/or victim blaming as this theme is currently covered in the current question 2.11 (Does the offender accept responsibility for the current offence?). No new questions are identified for inclusion at this time.

Implications

The analysis of OASys textual data resulted in recommendations to pilot new questions or to revise the OASys guidance where necessary. Table 11.1 summarises proposed changes for each section of the core OASys assessment. Further research will be required to test the reliability and validity of any new questions alongside the existing OASys questions.

Table 11.1: Summary of recommendations from textual analysis

Section	Proposed changes
Accommodation	Adapt the guidance for question 3.4: suitability of accommodation to consider if the offender has future plans to change their accommodation situation, which would risk repeat offending in domestic violence cases.
ETE	Pilot a new question on the appropriateness of an offender's employment or employment skills to reduce the likelihood of reoffending.
Financial management	No changes
Drug misuse	No changes
Alcohol misuse	Pilot the following new questions: Has the offender ever been convicted for a drink-driving offence? Is alcohol used to escape from life, e.g. in response to emotional stress?
Emotional wellbeing	No changes
Thinking and behaviour	No changes
Attitudes	No changes

Conclusion

Exploratory textual analysis of the information recorded in the text box sections of the OASys sections on accommodation, education, training and employment, and alcohol misuse identified several themes that suggested ideas for amending those sections' content with additional questions. Analysis of textual information recorded in the sections assessing financial management and income, drug misuse, emotional wellbeing, thinking and behaviour and attitudes did not point to the need for any additional questions.

12. OASys statistics: 2008 probation and prison assessments

Introduction

The potential benefits from using OASys data as a source of management information are set out in the OASys user manual:

OASys has the capacity to provide valuable management information, some of which will be used by practitioners to develop profiles of the offenders they are working with and to evaluate overall outcomes. Information will also be of use to local managers, to enable them to identify which risk factors are most common within their local offender population and to help ensure that adequate provision has been made for them. When applied on a national basis, OASys will provide a profile of offenders and their needs, and will permit resources to be allocated effectively

(Home Office, 2002:3-4)

This chapter presents data on offenders assessed in 2008, from the database of completed OASys assessments held by O-DEAT. The presentation consists of profiles for the sample as a whole and for a number of offender sub-groups.

The 2008 sample

The 2008 probation and prison assessments held within the O-DEAT database were cleansed and de-duplicated by selecting valid assessments and prioritising the earliest such assessments in each individual contact period. For an OASys assessment to be held valid, the following standards of data completion had to be satisfied.

- Each of the scored sections (1 to 12) within the core OASys assessment must have had at least four-fifths of their scored items completed – ensuring that each criminogenic need was assessed properly.
- In the risk of serious harm sections, the screening must have been completed, the decision whether to complete a full risk analysis should have been consistent with the information provided, and the four ratings of risk of serious harm in the community must have been recorded in those cases in which a full analysis was required.

This sampling left 325,863 assessments, 305,483 (94%) of which were completed by the probation service and 20,380 by the prison service. 150,444 (46%) of these assessments included a fully completed SAQ.⁸⁶ Nearly nine-tenths (88%) of the offenders were male, their average age was 32, and 84% were of White ethnic classification. Over half (53%) had been in contact with the police prior to the age of 18, while 45% had over five previous convictions. The offence category of violence against the person was recorded in 31% of the cases. Over two-fifths (44%) had received a community sentence, with a further 33% having received a custodial

⁸⁶ There is no national standard for completion of the SAQ and the paper-based data may not always be transferred to the electronic system.

sentence and 18% having received a suspended sentence. At a regional level, the sample sizes ranged from 18,798 (5.8% of the national sample) for Wales to 52,885 (16%) for the North West.

A second sample of 2008 probation and prison assessments was used to assess the extent to which identified needs were being addressed through planned interventions. This sample was restricted to assessments which were recorded as having been administered at the start of a community sentence, suspended sentence or a custodial sentence. In addition to the standards of data completion set out above, information must have been recorded within the objectives and plans section of the OASys sentence plan. This sampling left 110,943 assessments, 73,491 (66%) of which were start community sentence assessments, 27,982 (25%) of which were start suspended sentence assessments and 9,470 (8.5%) of which were start custodial sentence assessments.

Analysis

OASys profile data are presented for the sample as a whole and for a number of offender subgroups. Previous research has found that there are differences between the criminogenic risk factors of males and females (e.g. Motiuk and Blanchette, 2000) and of different ethnic groups (e.g. Calverley *et al.*, 2004). Of particular interest to policy makers, in terms of reducing both the frequency and seriousness of offending, have been persistent offenders, serious offenders and early-onset offenders, with recent research identifying differences between the profiles of these groups (e.g. Motiuk, 2000; Hanson and Morton-Bourgon, 2004; Ge, Donnellan and Wenk, 2001). In this paper, the following variables have been used to group the offenders:

- Gender (Male/Female);
- Ethnicity (White/Black/Asian/Mixed/Other);
- Age at assessment (18–20; 21–24; 25–30; 31–40; 41+);
- Age of first police contact (Under 14; 14-17; 18+);⁸⁷
- Number of previous convictions (0; 1–5; 6–10; 11+);⁸⁸
- Offence category (Violence against the person; Sexual offences; Burglary; Robbery; Theft and handling; Fraud and forgery; Criminal damage; Drug offences; Other indictable offences; Summary motoring offences; Other summary offences).⁸⁹

There are no commonly accepted definitions of early-onset offending, persistent offending and serious offending, but the latter three variables set out above are used to provide an indication of each. A further breakdown is provided by sentence category and region. When presenting data from the Self Assessment Questionnaire (SAQ), a breakdown is provided by

87 OASys records both age at first conviction (Q1.7) and age first in contact with the police (Q1.8). The latter was preferred as an indicator of early-onset offending due to its inclusion of youth justice reprimands and final warnings.

88 OASys records the number of previous convictions aged under 18 years (Q1.5) and 18 years and over (Q1.6). These two fields were combined to provide the overall number of previous convictions.

89 Information regarding previous offence types is also recorded within section 1 of the core OASys assessment and the risk of serious harm screening. These items could have been used to group the offenders, ensuring that previous serious offences were not hidden by relatively minor current offences. However, as the timescales between the offences was unknown, it was thought preferable to focus upon current offences so that the offenders' current risk/need profiles were not linked to non-recent types of offending behaviour.

the likelihood of reconviction score bands produced by the core OASys assessment and the practitioners' risk of serious harm ratings, enabling a comparison to be made between the offenders' views and those of the practitioners.

While the tables in this chapter do not set out data for the variables in combination, there were some important overlaps between the subgroups. For example, 56% of the 50,279 early-onset offenders (first police contact prior to the age of 14) were highly persistent (more than ten previous convictions).⁹⁰ The most significant findings for the variables in combination are set out in the text.

To identify relatively high and low levels of risk and need for specific offender groups, odds ratios were used, comparing the risk/need rates for each group against the rates for the sample as a whole. Odds ratios of 0.67 or less were used to indicate low levels of risk/need and odds ratios of 1.5 and above used to indicate high levels of risk/need.⁹¹

Data limitations

OASys is not completed with all offenders, and previous analysis has found that offenders with an OASys were more likely to have committed a violent offence and to have a high likelihood of reconviction (see Chapter 9). OASys data should not be read as representative of the entire offending population and care should be taken in generalising the results.

The value of OASys risk and need profile information is also dependent upon the assessment tool being both reliable and valid. The research presented in this compendium presents findings on several types of reliability and validity studies of OASys with recommendations for improvements made where necessary. The data presented in this chapter is included with confidence about the value of OASys as a reliable and valid assessment.

Findings

As shown by Table 12.1, the criminogenic need prevalence rates of the complete 2008 sample ranged from 24% for financial management and income to 57% for thinking and behaviour. Female offenders had relatively high levels of need for relationships and emotional wellbeing, and relatively low levels of need for both (i) thinking and behaviour and (ii) attitudes. The youngest offenders (aged 18–20) had relatively high levels of need for (i) education, training and employability and (ii) lifestyle and associates, and a relatively low level of need for emotional wellbeing. The oldest offenders (aged over 40) had relatively low levels of need for (i) education, training and employability and (ii) drug misuse. Asian offenders had relatively low levels of need across five of the ten OASys sections and Black offenders had relatively low levels of need across three of these sections – relationships, alcohol misuse and emotional wellbeing.

90 Previous research has identified an early age of first arrest as one of the strongest predictors of persistent offending (see, for example, Blumstein *et al.*, 1986; Farrington, 1992; Ge, Donnellan and Wenk, 2001).

91 For example, if a specific group had a criminogenic need prevalence rate of 40% and the rate for the whole sample was 60%, their corresponding odds of having the need would be two-thirds ($0.4/(1-0.4)$) and 1.5 ($0.6/(1-0.6)$) respectively. Consequently, the odds ratio would be less than half ($0.67/1.5 = 0.44$).

Those 'early-onset' offenders whose first contact with the police was prior to the age of 14 had relatively high levels of need across six of the ten OASys sections, while the most persistent offenders (those with more than ten previous convictions) had relatively high levels of need across eight sections. In contrast first-time offenders with no previous convictions had relatively low levels of need across nine of the sections. Clear differences were evident when grouping the offenders according to their current offence. Those who had committed an offence of burglary or robbery had relatively high levels of need across seven sections, while those who had committed a summary motoring offence or an offence of fraud/forgery had relatively low levels of need across eight sections – although those offenders committing fraud/forgery offences had a relatively high level of need for financial management and income. Those who had committed an offence of violence against the person had a relatively high level of need for alcohol misuse, while those who had committed a sexual offence had relatively high levels of need for (i) lifestyle and associates and (ii) attitudes.

The offenders could be distinguished even more clearly through combinations of the grouping variables. For example, the criminogenic need prevalence rates for the early-onset, most persistent offenders committing an offence of burglary or robbery were as high as 90% for education, training and employability (n=7,100). In contrast, the prevalence rates for first time, late-onset offenders committing a summary motoring offence were as low as 1.5% for drug misuse (n=4,973).

Table 12.2 presents data from the OASys sentence plan, setting out the planned intervention prevalence rates for those assessed at the start of sentence. For each OASys section, the analysis is restricted to those offenders for whom some form of intervention appeared necessary, as indicated by a scored criminogenic need. As shown by Table 12.2, the planned intervention prevalence rates for the complete 2008 sample ranged from 13% for financial management and income to 69% for drug misuse.⁹² There were relatively high planned intervention rates across seven of the ten needs for those offenders assessed at the start of a custodial sentence and across six of the ten needs for those who had committed an offence of robbery. The youngest offenders (aged 18–20) had relatively high planned intervention rates for (i) education, training and employability, (ii) lifestyle and associates, (iii) thinking and behaviour and (iv) attitudes, with relatively low planned intervention rates for (i) financial management and income, (ii) relationships, (iii) drug misuse and (iv) alcohol misuse. Female offenders had relatively high planned intervention rates for (i) financial management and income and (ii) drug misuse, and relatively low planned intervention rates for (i) relationships and (ii) lifestyle and associates.

92 The number of coded interventions varies greatly between the criminogenic needs: thinking and behaviour has 28 corresponding intervention codes while accommodation has just two corresponding intervention codes. Previous analysis has revealed that practitioners most commonly enter three intervention codes, and the current analysis included up to 12 intervention codes, taking into account the practical limitations upon how much can be delivered within an individual period of contact. Practitioners can indicate whether the need is to be addressed currently or in the future, but no distinction was employed in the analysis.

Table 12.1: Criminogenic need prevalence rates by offender subgroups

	n	Criminogenic needs by section									
		Accommodation	ETE	Finance	Relationships	Lifestyle & associates	Drugs	Alcohol	Emotional wellbeing	Thinking & behaviour	Attitudes
All offenders	325,863	36.5%	53.9%	24.4%	41.9%	42.3%	27.0%	43.5%	41.2%	56.8%	34.9%
Gender											
Male	285,648	36.6%	53.4%	23.8%	39.5%	43.2%	27.2%	44.4%	38.1%	58.4%	36.1%
Female	40,184	35.9%	57.9%	28.6%	59.0%	35.7%	25.5%	37.3%	63.2%	45.7%	26.2%
Age group											
18–20	35,651	37.1%	67.9%	25.9%	35.7%	54.6%	25.2%	48.3%	30.7%	65.6%	41.4%
21–24	59,756	35.8%	60.3%	24.9%	39.2%	46.3%	26.5%	47.7%	34.3%	60.7%	36.8%
25–30	71,642	36.8%	56.1%	27.3%	42.5%	43.6%	33.3%	42.2%	39.7%	56.9%	34.7%
31–40	85,246	38.4%	52.3%	26.1%	45.7%	40.7%	32.3%	42.0%	46.7%	55.8%	33.8%
41+	73,400	34.3%	41.8%	18.2%	42.3%	33.5%	15.9%	40.8%	47.2%	50.3%	31.6%
Ethnicity											
White	243,249	37.9%	56.2%	25.0%	45.3%	44.0%	28.5%	48.4%	44.1%	58.8%	36.4%
Black	22,859	39.2%	52.8%	27.2%	32.0%	43.9%	29.2%	17.9%	26.6%	53.9%	37.1%
Asian	14,030	25.5%	41.8%	21.8%	23.3%	36.5%	21.3%	22.3%	25.9%	50.2%	30.3%
Mixed	8,188	40.2%	62.4%	28.6%	44.6%	49.7%	34.5%	32.6%	35.1%	61.8%	41.8%
Other	2,282	39.3%	51.2%	27.7%	24.0%	36.9%	15.0%	20.3%	31.2%	46.6%	28.0%
Age of first police contact											
Under 14	50,279	45.9%	78.1%	33.3%	55.5%	61.3%	40.0%	51.0%	41.9%	72.2%	52.2%
14-17	121,648	40.4%	65.2%	27.8%	46.3%	50.9%	34.6%	48.7%	40.8%	64.9%	41.8%
18+	153,881	30.4%	37.1%	18.8%	34.1%	29.3%	16.7%	37.0%	41.3%	45.4%	23.8%
Number of previous convictions											
0	61,557	23.8%	28.4%	14.0%	25.6%	20.6%	7.5%	25.4%	35.0%	33.8%	14.3%
1–5	116,871	30.8%	45.2%	18.1%	35.0%	32.9%	17.6%	43.7%	38.0%	51.6%	26.3%
6–10	53,965	40.7%	62.1%	25.6%	46.9%	49.4%	31.1%	51.5%	42.2%	65.6%	41.7%
11+	92,468	49.9%	77.2%	38.5%	58.8%	64.5%	49.5%	50.7%	48.8%	73.6%	55.7%

Table 12.1: Criminogenic need prevalence rates by offender subgroups (continued)

Offence category	n	Criminogenic needs by section									
		Accommodation	ETE	Finance	Relationships	Lifestyle & associates	Drugs	Alcohol	Emotional wellbeing	Thinking & behaviour	Attitudes
Violence against the person	100,480	36.4%	49.6%	14.0%	46.5%	32.8%	17.9%	59.5%	42.1%	63.7%	32.3%
Sexual offences	16,055	44.6%	42.2%	10.8%	50.4%	55.5%	9.2%	26.5%	48.9%	63.7%	48.9%
Burglary	23,501	50.9%	78.1%	46.8%	50.1%	68.0%	52.9%	36.7%	42.2%	67.8%	50.2%
Robbery	14,183	48.8%	70.9%	41.3%	45.2%	69.8%	48.0%	43.1%	37.1%	69.2%	48.0%
Theft and handling	42,702	43.4%	69.6%	46.0%	46.2%	57.1%	47.9%	33.0%	47.3%	58.4%	41.9%
Fraud and forgery	11,609	22.6%	37.8%	34.2%	28.5%	25.6%	10.4%	11.8%	37.6%	30.1%	18.7%
Criminal damage	10,655	44.0%	62.9%	20.6%	51.1%	44.1%	23.1%	67.1%	53.1%	70.6%	38.5%
Drug offences	30,699	32.7%	53.1%	29.8%	31.8%	45.2%	45.8%	17.7%	29.4%	37.3%	26.1%
Other indictable offences	25,028	29.5%	46.7%	16.1%	34.9%	39.7%	17.5%	39.2%	37.1%	53.3%	32.9%
Summary motoring offences	31,868	20.2%	36.0%	12.2%	25.5%	22.6%	10.0%	48.0%	34.6%	38.8%	22.5%
Other summary offences	18,708	35.0%	55.4%	20.0%	44.8%	36.5%	18.4%	51.8%	49.0%	63.3%	38.1%

Table 12.1: Criminogenic need prevalence rates by offender subgroups (continued)

Sentence category	n	Criminogenic needs by section									
		Accommodation	ETE	Finance	Relationships	Lifestyle & associates	Drugs	Alcohol	Emotional wellbeing	Thinking & behaviour	Attitudes
C-JA 03 Community Order	123,472	33.6%	52.6%	22.1%	41.3%	34.5%	24.8%	48.0%	44.2%	54.7%	29.3%
Custody/YOI	93,117	45.7%	60.3%	30.1%	45.2%	58.4%	34.6%	37.5%	36.7%	61.8%	45.5%
Suspended sentence	51,470	31.4%	50.3%	21.1%	41.9%	36.2%	22.5%	48.5%	42.6%	56.7%	31.8%
Other	15,401	33.7%	54.4%	20.9%	43.3%	40.5%	22.6%	41.1%	44.6%	56.0%	37.4%
Region											
North West	52,885	34.6%	53.8%	23.4%	40.8%	42.6%	27.7%	45.8%	40.7%	56.7%	35.1%
North East	22,985	36.7%	54.8%	24.7%	42.8%	42.8%	25.7%	48.4%	41.4%	56.0%	37.4%
Yorkshire and Humberside	33,811	38.9%	57.7%	25.6%	45.2%	47.0%	28.8%	45.5%	41.6%	59.8%	39.0%
East Midlands	29,746	34.5%	52.5%	23.5%	40.4%	39.7%	25.1%	41.8%	38.9%	55.2%	33.9%
East of England	28,583	34.3%	49.1%	21.1%	40.9%	36.8%	23.3%	42.7%	41.4%	52.7%	31.1%
West Midlands	37,170	34.7%	55.5%	25.5%	40.1%	43.6%	26.2%	40.1%	38.8%	59.0%	35.3%
South East	40,381	37.3%	51.0%	22.6%	44.6%	39.9%	24.8%	46.5%	43.3%	54.3%	33.1%
South West	23,360	39.1%	52.7%	23.1%	46.2%	41.2%	27.3%	48.7%	45.3%	58.0%	36.3%
London	38,050	38.7%	56.0%	28.2%	37.0%	44.2%	31.2%	31.1%	38.0%	56.6%	33.8%
Wales	18,798	38.2%	56.8%	26.5%	44.7%	44.4%	28.9%	50.4%	46.7%	61.2%	35.0%

Key: Odds ratio (group % vs. all %) <= 0.67 Odds ratio (group % vs. all %) >=1.5

Table 12.2: Planned intervention prevalence rates for offenders with scored criminogenic needs by offender subgroups

	n	Planned interventions by section									
		Accommodation	ETE	Finance	Relationships	Lifestyle & associates	Drugs	Alcohol	Emotional wellbeing	Thinking & behaviour	Attitudes
All offenders	110,943	29.7%	46.8%	13.4%	16.0%	32.5%	68.5%	65.7%	25.6%	43.5%	37.3%
Gender											
Male	94,245	29.0%	47.7%	12.0%	17.4%	33.8%	66.9%	65.2%	25.4%	44.6%	38.5%
Female	16,693	33.5%	42.1%	19.5%	10.7%	24.2%	77.2%	69.5%	26.4%	36.2%	28.9%
Age group											
18–20	14,753	28.3%	57.5%	8.3%	8.8%	43.0%	48.3%	55.8%	25.9%	54.0%	48.7%
21–24	20,937	29.0%	52.1%	12.2%	14.0%	38.6%	58.4%	59.9%	26.2%	49.3%	42.5%
25–30	24,726	29.8%	48.1%	13.1%	16.6%	30.6%	73.5%	66.6%	25.7%	42.4%	34.8%
31–40	29,036	31.4%	41.6%	13.7%	17.8%	26.6%	76.3%	71.1%	24.9%	38.3%	31.3%
41+	21,469	28.9%	34.4%	19.0%	18.5%	25.1%	70.7%	70.9%	26.1%	37.2%	33.2%
Ethnicity											
White	84,423	30.4%	46.6%	13.1%	15.2%	32.0%	69.5%	66.4%	25.5%	43.0%	36.4%
Black	6,262	31.2%	53.7%	13.4%	18.0%	41.9%	63.7%	67.2%	27.2%	49.6%	46.4%
Asian	4,313	19.3%	45.6%	13.5%	17.5%	36.4%	71.1%	63.5%	24.0%	47.8%	40.0%
Mixed	2,624	31.6%	52.3%	12.9%	16.2%	38.2%	64.1%	64.5%	24.4%	49.7%	46.7%
Other	633	20.7%	47.6%	13.0%	22.1%	24.1%	73.6%	60.6%	29.7%	37.5%	33.9%
Age of first police contact											
Under 14	16,663	30.7%	49.7%	8.6%	11.0%	37.6%	66.1%	64.3%	24.9%	47.7%	40.8%
14-17	42,199	30.8%	48.0%	11.0%	14.9%	34.0%	68.6%	65.3%	25.3%	45.2%	37.9%
18+	52,064	28.2%	43.1%	18.4%	19.5%	27.0%	70.0%	66.7%	26.2%	39.6%	34.0%
Number of previous convictions											
0	19,072	23.9%	47.8%	23.8%	20.5%	28.3%	57.1%	58.6%	27.2%	40.3%	36.2%
1–5	41,207	26.6%	47.7%	16.7%	19.3%	33.4%	59.6%	62.8%	26.5%	44.5%	39.5%
6–10	18,494	29.3%	47.4%	11.2%	17.1%	35.0%	65.2%	65.4%	26.3%	45.7%	40.7%
11+	31,844	34.3%	45.5%	9.5%	11.5%	31.5%	75.2%	72.1%	23.7%	42.4%	34.7%

Table 12.2: Planned intervention prevalence rates for offenders with scored criminogenic needs by offender subgroups (continued)

Offence category	n	Criminogenic needs by section									
		Accommodation	ETE	Finance	Relationships	Lifestyle & associates	Drugs	Alcohol	Emotional wellbeing	Thinking & behaviour	Attitudes
Violence against the person	34,837	23.7%	43.9%	10.5%	29.2%	29.9%	48.6%	60.6%	29.1%	42.0%	35.8%
Sexual offences	1,978	24.7%	44.4%	13.7%	6.4%	31.5%	55.1%	64.0%	26.3%	75.9%	67.0%
Burglary	6,480	36.4%	55.7%	9.9%	6.1%	39.8%	78.2%	74.5%	24.0%	48.1%	41.5%
Robbery	1,603	35.6%	59.8%	14.3%	7.9%	62.6%	75.8%	78.0%	34.6%	72.3%	67.5%
Theft and handling	20,893	37.0%	47.3%	13.1%	6.5%	27.8%	79.7%	72.2%	20.9%	37.4%	29.9%
Fraud and forgery	3,693	30.0%	48.5%	30.3%	8.9%	34.9%	70.2%	59.0%	21.1%	43.5%	39.2%
Criminal damage	4,443	29.6%	44.4%	9.1%	16.4%	31.5%	46.0%	62.8%	30.5%	43.7%	37.4%
Drug offences	8,713	33.0%	54.0%	14.3%	5.9%	35.8%	77.7%	75.5%	22.7%	45.3%	39.2%
Other indictable offences	7,178	29.6%	45.5%	11.6%	11.3%	37.0%	58.3%	57.7%	29.5%	52.6%	46.7%
Summary motoring offences	13,295	23.0%	42.6%	17.7%	7.2%	35.6%	48.9%	75.7%	21.8%	42.5%	38.6%
Other summary offences	7,702	28.2%	41.0%	13.0%	15.2%	27.9%	50.2%	62.0%	26.9%	39.7%	35.1%

Table 12.2: Planned intervention prevalence rates for offenders with scored criminogenic needs by offender subgroups (continued)

Sentence category	n	Criminogenic needs by section									
		Accommodation	ETE	Finance	Relationships	Lifestyle & associates	Drugs	Alcohol	Emotional wellbeing	Thinking & behaviour	Attitudes
C-JA 03 Community Order	73,491	30.0%	44.5%	12.9%	16.2%	26.6%	67.5%	64.6%	24.9%	38.5%	31.1%
Start custody	9,470	35.0%	65.2%	15.2%	8.7%	67.8%	84.2%	83.0%	34.4%	80.3%	73.6%
Start suspended sentence	27,982	26.6%	46.0%	13.7%	17.6%	29.8%	63.8%	64.5%	25.3%	43.8%	37.2%
Region											
North West	17,580	27.7%	36.9%	10.5%	17.4%	31.1%	69.8%	66.7%	26.1%	47.1%	36.2%
North East	10,097	35.5%	52.0%	17.4%	12.4%	63.3%	64.7%	58.9%	30.9%	37.2%	28.5%
Yorkshire and Humberside	11,806	30.4%	46.3%	12.5%	16.3%	26.5%	67.9%	64.0%	26.1%	47.6%	35.1%
East Midlands	10,749	34.1%	55.1%	18.8%	17.2%	30.2%	70.1%	65.3%	30.5%	43.7%	38.3%
East of England	10,543	30.6%	48.3%	15.5%	15.3%	30.8%	66.1%	63.7%	22.7%	45.5%	39.9%
West Midlands	13,126	23.9%	44.9%	15.6%	12.2%	30.7%	68.3%	62.9%	26.6%	45.2%	41.3%
South East	12,577	27.4%	48.1%	14.1%	17.6%	24.3%	68.5%	68.1%	23.1%	40.0%	33.6%
South West	6,898	32.2%	55.2%	12.0%	17.0%	30.1%	71.7%	65.6%	19.2%	43.3%	45.8%
London	10,474	27.2%	45.7%	7.2%	18.6%	31.6%	67.2%	73.9%	24.9%	41.9%	41.7%
Wales	6,985	32.2%	44.8%	11.4%	14.7%	27.3%	70.9%	70.3%	25.4%	38.2%	36.7%

Key: Odds ratio (group % vs. all %) <= 0.67 Odds ratio (group % vs. all %) >=1.5

As shown by Table 12.3, the offenders in the 2008 sample had an average of four criminogenic needs. Approximately one-fifth (19%) of the 2008 sample were scored as having a high likelihood of reconviction and one in ten was rated as presenting a high/very high risk of serious harm to the community. Early-onset offenders, persistent offenders, those who had committed an offence of burglary, robbery or theft, and those sentenced to custody were most commonly scored as having a high likelihood of reconviction, while those who had committed a sexual offence or robbery and those sentenced to custody were most commonly rated as presenting a high/very high risk of serious harm to the community. At the extremes, one-half of the most persistent offenders (those with over ten previous convictions) had a high likelihood of reconviction compared to just 0.1% of the first-time offenders, and approximately one-half (49%) of those who had committed a sexual offence presented a high/very high risk of serious harm compared to just 0.7% of those committing an offence of fraud or forgery.

Greater distinctions were evident when combining the grouping variables. Notably, 76% of the early-onset, most persistent offenders committing an offence of burglary or robbery had a high likelihood of reconviction (n=7,100).

The risk of serious harm ratings presented in Table 12.3 band together the risks to four specific groups (children/public/known adult/staff), presenting the highest rating across the four groups. A breakdown of the risks to each of these groups is provided in Table 12.3.⁹³ As shown, 6.0% of the 2008 sample were rated as presenting a high/very high risk of serious harm to the public, 4.4% were rated as presenting such a risk to a known adult and 3.0% were rated as presenting such a risk to children. The most prominent differences were evident when grouping the offenders according to their current offence. Over one-third (36%) of those who had committed a sexual offence were rated as presenting a high/very high risk of serious harm to children, while over one-quarter (27%) of those who had committed an offence of robbery were rated as presenting a high/very high risk of serious harm to the public.

93 Risk to the public covers harm of a general nature or to a specific group, while risk to known adults focuses upon harm to specific individuals (e.g. previous victims, partners).

Table 12.3: Likelihood of reconviction and risk of serious harm levels by offender subgroups

	n	Mean no. needs	Likelihood of reconviction			Risk of serious harm (highest risk in community – all categories)		
			Low (0–40)	Medium (41–99)	High (100–168)	Low	Medium	High/Very high
			All offenders	325,863	4.0	32.1%	49.1%	18.8%
Gender								
Male	285,648	4.0	31.0%	49.3%	19.7%	33.3%	55.3%	11.4%
Female	40,184	4.1	40.2%	47.4%	12.5%	56.1%	40.7%	3.2%
Age group								
18–20	35,651	4.3	28.5%	53.5%	18.1%	35.0%	57.5%	7.5%
21–24	59,756	4.1	31.5%	49.8%	18.7%	34.2%	56.7%	9.1%
25–30	71,642	4.1	29.6%	48.8%	21.6%	36.1%	54.3%	9.6%
31–40	85,246	4.1	29.9%	48.7%	21.4%	36.5%	53.2%	10.4%
41+	73,400	3.6	39.5%	46.9%	13.6%	37.7%	48.5%	13.7%
Ethnicity								
White	243,249	4.2	28.5%	50.1%	21.4%	35.4%	54.2%	10.4%
Black	22,859	3.6	34.5%	49.7%	15.8%	33.1%	52.5%	14.4%
Asian	14,030	3.0	48.1%	44.3%	7.6%	40.6%	49.4%	10.0%
Mixed	8,188	4.3	25.4%	51.4%	23.2%	29.4%	57.7%	12.9%
Other	2,282	3.2	48.3%	45.1%	6.6%	49.2%	40.6%	10.2%
Age of first police contact								
Under 14	50,279	5.3	5.7%	49.6%	44.7%	26.7%	59.0%	8.5%
14–17	121,648	4.6	14.9%	58.3%	26.8%	30.5%	58.3%	11.2%
18+	153,881	3.1	54.4%	41.6%	4.0%	43.6%	47.9%	14.4%
Number of previous convictions								
0	61,557	2.3	77.0%	22.9%	0.1%	48.9%	43.5%	7.6%
1–5	116,871	3.4	42.9%	53.4%	3.8%	38.9%	52.7%	8.4%
6–10	53,965	4.6	10.0%	70.6%	19.4%	30.1%	58.0%	11.9%
11+	92,468	5.7	1.5%	48.5%	50.0%	27.4%	58.6%	14.0%
Offence category								
Violence against the person	100,480	3.9	35.9%	51.2%	12.9%	16.6%	70.0%	13.4%
Sexual offences	16,055	4.0	39.2%	50.3%	10.6%	4.9%	46.3%	48.8%
Burglary	23,501	5.4	7.7%	44.9%	47.4%	37.9%	55.8%	6.2%
Robbery	14,183	5.2	14.2%	48.4%	37.4%	6.3%	65.5%	28.1%
Theft and handling	42,702	4.9	17.0%	50.2%	32.8%	58.8%	39.0%	2.2%
Fraud and forgery	11,609	2.6	57.9%	36.2%	5.9%	82.9%	16.4%	0.7%
Criminal damage	10,655	4.8	21.8%	58.5%	19.7%	25.5%	62.9%	11.7%
Drug offences	30,699	3.5	32.5%	52.1%	15.4%	62.5%	35.7%	1.8%
Other indictable offences	25,028	3.5	41.5%	44.6%	13.9%	30.3%	58.3%	11.4%
Summary motoring offences	31,868	2.7	49.4%	43.0%	7.6%	60.0%	38.6%	1.4%
Other summary offences	18,708	4.1	31.8%	53.7%	14.5%	36.2%	58.0%	5.8%

Table 12.3: Likelihood of reconviction and risk of serious harm levels by offender subgroups (continued)

	n	Mean no. needs	Likelihood of reconviction			Risk of serious harm (highest risk in community – all categories)		
			Low (0–40)	Medium (41–99)	High (100–168)	Low	Medium	High/Very high
Sentence category								
CJA 03 Community Sentence	123,472	3.9	34.9%	51.1%	14.0%	43.1%	53.8%	3.2%
Custody/YOI	93,117	4.6	22.8%	47.4%	29.8%	22.7%	52.9%	24.4%
Suspended sentence	51,470	3.8	33.9%	51.4%	14.7%	37.2%	58.1%	4.7%
Other	15,401	3.9	33.7%	49.5%	16.7%	39.1%	51.2%	9.7%
Region								
North West	52,885	4.0	30.9%	49.7%	19.4%	33.0%	55.1%	12.0%
North East	22,985	4.1	31.4%	47.2%	21.4%	37.4%	54.6%	8.0%
Yorkshire and Humberside	33,811	4.3	26.5%	50.9%	22.5%	36.1%	51.0%	12.9%
East Midlands	29,746	3.9	34.7%	47.6%	17.7%	36.0%	55.5%	8.5%
East of England	28,583	3.7	36.9%	47.9%	15.1%	40.9%	50.9%	8.2%
West Midlands	37,170	4.0	32.7%	49.1%	18.2%	37.6%	52.0%	10.5%
South East	40,381	4.0	33.3%	49.5%	17.1%	41.0%	49.8%	9.2%
South West	23,360	4.2	31.3%	48.8%	19.9%	34.0%	56.5%	9.5%
London	38,050	3.9	33.9%	49.2%	16.9%	30.7%	56.5%	12.8%
Wales	18,798	4.3	29.2%	49.1%	21.7%	36.4%	53.8%	9.8%

Key: Odds ratio (group % vs. all%) ≤ 0.67 Odds ratio (group % vs. all%) ≥ 1.5

Table 12.4: Risk of serious harm rates (full community breakdown) by offender subgroups

	n	Risk of serious harm (in the community) to...											
		Children			Public			Known adult			Staff		
		Low	Medium	High/ Very high	Low	Medium	High/ Very high	Low	Medium	High	Low	Medium	High/ Very high
All offenders	325,863	83.7%	13.3%	3.0%	51.5%	42.5%	6.0%	70.5%	25.1%	4.4%	92.5%	6.9%	0.5%
Gender													
Male	285,648	83.5%	13.2%	3.3%	49.0%	44.4%	6.6%	68.6%	26.6%	4.8%	92.5%	7.0%	0.6%
Female	40,184	85.5%	13.7%	0.9%	69.3%	28.9%	1.8%	84.0%	14.6%	1.4%	93.1%	6.6%	0.3%
Age group													
18–20	35,651	90.9%	7.7%	1.4%	42.2%	52.1%	5.7%	76.9%	20.6%	2.5%	92.3%	7.2%	0.5%
21–24	59,756	88.4%	10.0%	1.5%	43.8%	49.5%	6.7%	72.6%	23.7%	3.7%	92.7%	6.8%	0.5%
25–30	71,642	85.1%	13.2%	1.7%	49.3%	44.3%	6.4%	70.4%	25.3%	4.3%	92.9%	6.6%	0.5%
31–40	85,246	81.5%	15.9%	2.7%	53.8%	40.3%	5.8%	67.4%	27.6%	5.0%	91.9%	7.5%	0.6%
41+	73,400	77.7%	15.7%	6.6%	61.9%	32.8%	5.4%	69.3%	25.5%	5.3%	92.9%	6.5%	0.6%
Ethnicity													
White	243,249	82.8%	13.9%	3.2%	51.1%	43.2%	5.7%	69.9%	25.7%	4.4%	92.3%	7.2%	0.5%
Black	22,859	88.1%	9.8%	2.1%	42.6%	46.4%	11.1%	73.8%	20.9%	5.4%	90.4%	8.7%	0.9%
Asian	14,030	87.8%	10.2%	2.0%	54.8%	38.9%	6.2%	72.3%	23.1%	4.7%	95.2%	4.4%	0.4%
Mixed	8,188	86.6%	11.7%	1.8%	38.8%	51.6%	9.6%	70.0%	24.6%	5.4%	89.1%	10.3%	0.7%
Other	2,282	88.7%	9.4%	1.9%	60.6%	32.9%	6.6%	76.7%	18.3%	5.0%	94.3%	5.3%	0.4%
Age of first police contact													
Under 14	50,279	84.5%	12.8%	3.6%	35.2%	54.3%	3.4%	68.8%	25.5%	3.5%	88.3%	10.7%	0.3%
14–17	121,648	85.3%	12.4%	2.3%	41.6%	51.0%	7.4%	68.8%	26.2%	5.0%	90.8%	8.5%	0.6%
18+	153,881	82.2%	14.1%	2.7%	64.7%	31.9%	10.5%	72.4%	24.1%	5.6%	95.3%	4.4%	1.0%

Table 12.4: Risk of serious harm rates (full community breakdown) by offender subgroups (continued)

	Risk of serious harm (in the community) to...												
	n	Children			Public			Known adult			Staff		
		Low	Medium	High/Very high	Low	Medium	High/Very high	Low	Medium	High	Low	Medium	High/Very high
Number of previous convictions													
0	61,557	81.8%	14.1%	4.1%	73.9%	23.7%	2.4%	77.4%	20.2%	2.4%	97.9%	1.9%	0.1%
1-5	116,871	84.8%	12.3%	2.9%	56.3%	39.5%	4.2%	70.7%	25.8%	3.6%	94.9%	4.8%	0.3%
6-10	53,965	83.6%	13.7%	2.7%	42.6%	50.0%	7.4%	66.5%	28.3%	5.2%	91.3%	8.1%	0.6%
11+	92,468	83.7%	13.8%	2.5%	35.7%	54.4%	9.9%	67.9%	25.8%	6.3%	86.6%	12.3%	1.1%
Offence category													
Violence against the person	100,480	79.5%	18.7%	1.8%	42.2%	50.2%	7.6%	47.2%	44.5%	8.3%	89.9%	9.4%	0.7%
Sexual offences	16,055	27.6%	36.8%	35.7%	56.4%	26.8%	16.8%	73.9%	16.6%	9.5%	93.3%	5.6%	1.1%
Burglary	23,501	91.5%	7.9%	.6%	42.8%	52.2%	5.0%	81.5%	16.3%	2.3%	92.5%	7.0%	0.5%
Robbery	14,183	88.0%	10.2%	1.8%	7.6%	65.4%	27.1%	77.6%	18.2%	4.2%	88.9%	9.9%	1.2%
Theft and handling	42,702	91.6%	7.9%	0.5%	65.4%	33.2%	1.4%	87.0%	12.1%	0.9%	94.6%	5.2%	0.2%
Fraud and forgery	11,609	95.9%	3.9%	0.2%	87.6%	12.1%	0.3%	93.9%	5.7%	0.4%	98.3%	1.7%	0.0%
Criminal damage	10,655	83.4%	15.1%	1.5%	41.0%	50.9%	8.1%	54.7%	39.0%	6.3%	88.5%	10.5%	1.0%
Drug offences	30,699	94.5%	5.2%	0.2%	67.2%	31.5%	1.3%	90.0%	9.2%	0.8%	96.7%	3.1%	0.2%
Other indictable offences	25,028	80.2%	15.8%	4.1%	46.4%	47.1%	6.4%	74.5%	21.3%	4.2%	92.3%	7.0%	0.7%
Summary motoring offences	31,868	93.8%	5.9%	0.4%	64.0%	35.2%	0.8%	88.3%	10.9%	0.7%	97.1%	2.8%	0.1%
Other summary offences	18,708	86.1%	12.7%	1.2%	53.2%	44.1%	2.7%	62.2%	33.9%	3.9%	88.8%	10.7%	0.5%

The risk of serious harm component of OASys also considers the risks to the offenders themselves, recognising that some offenders are vulnerable and have the potential for self-harm. The levels of current concerns are set out in Table 12.5. As shown, for the complete 2008 sample, concerns about the offender's ability to cope in custody were recorded in 9.6% of the cases, concerns about vulnerability were recorded in 9.1% of the cases, concerns about self-harming were recorded in 7.7% of the cases, and concerns about suicide were recorded in 7.6% of the cases. Across all four measures, the levels of concern were relatively high for female offenders, with nearly one in five (18%) female offenders assessed as vulnerable and over one in ten (13%) assessed as presenting a risk of suicide. In contrast, there were relatively low levels of concern across all four measures for Black offenders and in relation to suicide and self-harm for Asian offenders.

There were relatively high levels of concern across all four measures for those whose current offence was criminal damage, and in relation to vulnerability and the ability to cope in custody for those who had committed a sexual offence. For these latter offenders, concerns about vulnerability were recorded in nearly one-quarter (23%) of the cases.

Table 12.5: Current concerns regarding risks to self by offender subgroups

	n	Risks to the individual – current concerns			
		Suicide	Self harm	Coping in custody	Vulnerability
All offenders	325,863	7.6%	7.7%	9.6%	9.1%
Gender					
Male	285,648	6.8%	6.7%	8.5%	7.8%
Female	40,184	13.3%	14.8%	17.5%	18.2%
Age group					
18–20	35,651	6.3%	7.5%	9.9%	10.5%
21–24	59,756	6.7%	8.0%	9.6%	9.1%
25–30	71,642	6.9%	7.7%	8.6%	8.0%
31–40	85,246	8.4%	8.0%	9.4%	8.4%
41+	73,400	8.6%	7.0%	10.6%	10.1%
Ethnicity					
White	243,249	8.2%	8.4%	9.8%	9.4%
Black	22,859	3.0%	3.0%	5.2%	5.9%
Asian	14,030	3.7%	4.0%	6.4%	6.4%
Mixed	8,188	5.6%	6.2%	7.8%	8.0%
Other	2,282	5.8%	5.7%	7.9%	8.9%
Age of first police contact					
Under 14	50,279	6.8%	7.7%	8.1%	8.3%
14–17	121,648	7.1%	7.8%	8.3%	7.8%
18+	153,881	8.2%	7.6%	11.1%	10.3%

Table 12.5: Current concerns regarding risks to self by offender subgroups (continued)

	n	Risks to the individual – current concerns			
		Suicide	Self harm	Coping in custody	Vulnerability
Number of previous convictions					
0	61,557	7.5%	6.4%	12.4%	11.6%
1–5	116,871	7.7%	7.6%	10.4%	9.5%
6–10	53,965	7.7%	8.1%	8.8%	8.4%
11+	92,468	7.5%	8.3%	7.1%	7.2%
Offence category					
Violence against the person	100,480	8.8%	8.8%	11.1%	9.2%
Sexual offences	16,055	8.5%	7.8%	14.6%	23.1%
Burglary	23,501	6.0%	7.2%	7.3%	7.4%
Robbery	14,183	5.5%	6.7%	7.5%	8.1%
Theft and handling	42,702	7.2%	8.1%	7.3%	7.8%
Fraud and forgery	11,609	6.5%	5.2%	8.6%	7.1%
Criminal damage	10,655	13.0%	13.4%	14.0%	13.2%
Drug offences	30,699	3.9%	3.8%	5.5%	5.5%
Other indictable offences	25,028	7.8%	7.5%	11.7%	10.5%
Summary motoring offences	31,868	6.3%	5.5%	7.6%	5.6%
Other summary offences	18,708	10.5%	10.2%	12.1%	10.6%
Sentence category					
CJA 03 Community Order	123,472	8.5%	8.6%	10.0%	9.2%
Custody/YOI	93,117	5.0%	5.6%	6.3%	8.1%
Suspended sentence	51,470	9.0%	8.7%	12.3%	9.4%
Other	15,401	8.6%	8.8%	11.4%	11.9%
Region					
North West	52,885	6.9%	7.0%	9.4%	9.2%
North East	22,985	8.0%	8.2%	10.0%	9.4%
Yorkshire and Humberside	33,811	7.8%	7.8%	10.2%	9.2%
East Midlands	29,746	7.9%	8.3%	9.8%	8.9%
East of England	28,583	8.4%	7.6%	9.2%	8.3%
West Midlands	37,170	7.7%	7.5%	9.4%	8.2%
South East	40,381	8.1%	8.1%	10.2%	9.6%
South West	23,360	8.0%	7.9%	9.3%	9.4%
London	38,050	5.7%	6.6%	8.3%	8.9%
Wales	18,798	8.8%	8.9%	10.8%	10.1%

Key: Odds ratio (group % vs. all %) <= 0.67

Odds ratio (group % vs. all %) >=1.5

Table 12.6 presents findings from the OASys self-assessment questionnaire (SAQ). Questions 1 to 27 of the SAQ address a range of ‘external’ social problems encompassing accommodation, employment and finances, relationships and lifestyle, as well as ‘internal’ individual characteristics, covering values, perceptions, reasoning, beliefs, attitudes and goals. All 27 questions are prefixed by the phrase ‘Are any of these a problem for you?’. As

shown by Table 12.6, approximately half (52%) of the 2008 sample answered positively to no more than five of these questions. Relatively high positive response rates were given by the most persistent offenders (those with over ten previous convictions), and those whose current offence was burglary, theft or criminal damage, with approximately one-third of each of these subgroups responding positively to over ten of the 27 questions. Differences were also evident between the ethnic groups, with relatively low proportions of Black and Asian offenders responding positively to over ten of the 27 questions – 15% and 12% respectively.⁹⁴

The final question within the SAQ (Q28) asks offenders whether they think that they are likely to offend in the future, with a four-scale response ranging from **definitely not** to **very likely**. Approximately half (53%) of the 2008 sample responded **definitely not**. Relatively high response rates for **quite likely** or **very likely** were provided by early-onset offenders, the most persistent offenders and those whose current offence was burglary or theft. At the extremes, 16% of those with over ten previous convictions responded **quite likely** or **very likely**, compared to just 1.9% of those with no previous convictions. Some further distinctions were evident when combining the grouping variables. For example, 21% of the early-onset, most persistent offenders committing an offence of theft or burglary thought that further offending was **quite likely** or **very likely**, with 40% of these offenders responding positively to over ten of the preceding 27 questions (n=5,156).

Table 12.6 also provides a breakdown for the SAQ responses by the likelihood of reconviction score bands produced by the core OASys assessment and the practitioners' risk of serious harm ratings. As shown, there were clear links between the offenders' views and those of the practitioners, with relatively high positive response rates to the SAQ questions for those offenders with a high likelihood of reconviction or a very high risk of serious harm. Nearly one-half (46%) of those with a high likelihood of reconviction and approximately one-third (34%) of those with a very high risk of serious harm responded positively to over ten of the 27 problems questions, and approximately one-fifth of these two groups (21% and 17% respectively) thought that further offending was **quite likely** or **very likely**. However, many offenders appeared to be more optimistic regarding their future desistance than indicated by their OASys scores, with over one-quarter (28%) of those with a high likelihood of reconviction score responding that they would **definitely not** offend again.

94 A previous study of offenders on probation in England and Wales found less evidence of crime-prone attitudes and beliefs and lower levels of self-reported problems for the three minority ethnic groups (Black, Asian and Mixed) compared to the White offenders (Calverley *et al.*, 2004).

Table 12.6: Number of self-assessed problems and perceived likelihood of reoffending by offender subgroups

	n	Number of problems				Likely to offend in the future			
		0	1–5	6–10	11+	Definitely not	Unlikely	Quite likely	Very likely
All offenders	150,444	16.5%	35.3%	25.2%	23.0%	53.5%	38.2%	7.3%	1.0%
Gender									
Male	132,514	17.1%	35.7%	24.9%	22.4%	52.1%	39.3%	7.6%	1.0%
Female	17,919	12.1%	32.3%	27.9%	27.6%	63.6%	30.2%	5.1%	1.1%
Age group									
18–20	16,436	11.9%	35.2%	27.6%	25.4%	47.8%	41.2%	9.9%	1.0%
21–24	29,331	14.7%	35.3%	25.7%	24.3%	52.1%	39.9%	7.0%	1.0%
25–30	33,900	15.9%	34.2%	24.9%	25.0%	51.2%	40.1%	7.7%	1.0%
31–40	38,943	16.5%	34.2%	25.3%	24.0%	52.4%	38.8%	7.7%	1.1%
41+	31,803	21.3%	37.8%	23.7%	17.2%	61.6%	32.5%	5.0%	0.8%
Ethnicity									
White	115,131	15.1%	33.9%	26.0%	25.1%	51.5%	39.4%	8.0%	1.1%
Black	11,320	21.6%	41.1%	22.1%	15.2%	50.6%	42.2%	6.5%	0.7%
Asian	6,579	29.4%	40.4%	18.3%	11.9%	64.6%	31.3%	3.6%	0.5%
Mixed	3,989	16.3%	35.8%	25.6%	22.3%	48.8%	41.8%	8.5%	0.9%
Other	919	18.9%	40.5%	24.8%	15.8%	60.2%	36.1%	3.3%	0.4%
Age of first police contact									
Under 14	24,525	11.2%	30.9%	27.9%	30.0%	39.5%	47.1%	11.8%	1.6%
14–17	57,675	13.7%	33.4%	26.3%	26.6%	45.2%	44.2%	9.4%	1.2%
18+	68,228	20.7%	38.5%	23.4%	17.4%	65.6%	30.1%	3.8%	0.6%
Number of previous convictions									
0	26,899	26.3%	43.2%	20.2%	10.4%	75.7%	22.3%	1.7%	0.3%
1–5	52,695	18.5%	38.1%	24.6%	18.8%	61.5%	33.7%	4.2%	0.6%
6–10	25,777	13.9%	33.4%	27.3%	25.5%	47.7%	43.5%	7.8%	1.0%
11+	44,791	9.8%	28.3%	27.8%	34.1%	34.1%	50.1%	14.0%	1.9%
Offence category									
Violence against the person	45,564	17.3%	36.0%	24.8%	21.8%	57.6%	36.2%	5.4%	0.7%
Sexual offences	7,063	23.9%	43.7%	20.4%	12.0%	67.2%	28.9%	3.7%	0.3%
Burglary	11,688	8.8%	29.3%	29.0%	32.9%	38.2%	46.5%	13.6%	1.7%
Robbery	7,415	16.7%	36.1%	24.1%	23.1%	45.4%	42.8%	11.2%	0.5%
Theft and handling	19,915	9.0%	27.3%	28.9%	34.9%	40.5%	45.9%	11.5%	2.1%
Fraud and forgery	5,161	23.6%	40.7%	21.7%	14.0%	74.1%	23.3%	2.2%	0.4%
Criminal damage	4,755	10.7%	28.1%	27.5%	33.6%	47.6%	42.7%	8.3%	1.4%
Drug offences	15,986	23.7%	38.4%	21.4%	16.4%	47.3%	43.6%	8.3%	0.9%
Other indictable offences	11,259	20.4%	39.2%	23.3%	17.1%	60.3%	33.5%	5.6%	0.6%
Summary motoring offences	13,345	16.6%	40.1%	26.5%	16.8%	65.0%	30.9%	3.6%	0.6%

Table 12.6: Number of self-assessed problems and perceived likelihood of reoffending by offender subgroups (continued)

	n	Number of problems				Likely to offend in the future			
		0	1–5	6–10	11+	Definitely not	Unlikely	Quite likely	Very likely
Sentence category									
CJA 03									
Community Sentence	54,548	12.3%	33.3%	27.7%	26.6%	55.2%	37.7%	5.9%	1.2%
Custody/YOI	50,984	22.9%	36.9%	21.4%	18.8%	43.8%	44.2%	11.2%	0.8%
Suspended sentence	24,263	12.9%	35.4%	27.7%	24.1%	60.5%	34.4%	4.3%	0.8%
Other	5,427	14.9%	34.1%	25.9%	25.2%	59.8%	33.3%	5.9%	1.1%
Region									
North West	18,698	17.6%	35.2%	25.5%	21.7%	50.2%	40.1%	8.6%	1.0%
North East	10,717	15.2%	34.9%	26.5%	23.5%	49.1%	40.2%	9.5%	1.2%
Yorkshire and Humberside	17,116	14.6%	35.3%	26.7%	23.4%	52.0%	39.5%	7.5%	1.0%
East Midlands	17,251	17.5%	35.0%	25.3%	22.1%	53.8%	38.4%	6.9%	0.9%
East of England	10,361	17.7%	36.0%	23.8%	22.5%	49.8%	41.3%	8.0%	0.9%
West Midlands	16,131	16.7%	35.4%	25.3%	22.5%	57.5%	34.4%	7.3%	0.8%
South East	21,532	16.3%	34.4%	24.4%	24.9%	53.1%	39.1%	6.8%	1.0%
South West	10,371	14.5%	31.9%	26.1%	27.5%	54.6%	37.8%	6.2%	1.3%
London	18,762	18.4%	38.9%	23.2%	19.5%	59.8%	34.3%	4.9%	0.9%
Wales	9,428	14.6%	34.0%	26.4%	25.0%	52.0%	38.9%	8.1%	1.0%
Likelihood of reconviction									
Low (0–40)	45,493	29.9%	45.7%	18.0%	6.3%	75.1%	23.4%	1.3%	0.2%
Medium (41–99)	74,989	12.8%	35.1%	28.4%	23.8%	50.7%	42.1%	6.4%	0.8%
High (100+)	29,962	5.3%	20.0%	28.2%	46.4%	27.8%	51.1%	18.4%	2.7%
Risk of serious harm									
Low	50,999	19.5%	36.9%	24.3%	19.2%	59.3%	34.3%	5.5%	0.9%
Medium	84,141	15.0%	34.7%	25.8%	24.5%	50.8%	40.2%	8.0%	1.0%
High	14,815	14.5%	33.5%	25.3%	26.7%	49.2%	40.9%	8.9%	1.1%
Very high	489	13.7%	25.8%	26.2%	34.4%	42.7%	40.5%	13.5%	3.3%

Key: Odds ratio (group % vs. all %) <= 0.67 Odds ratio (group % vs. all %) >=1.5

Implications

Implications for practitioners and policy makers from the 2008 profiles are as follows.

- Combinations of interventions are required to address the co-occurring internal and external risk factors exhibited by early-onset persistent offenders.
- Interventions and risk management plans are required to prevent further serious offending by violent and sexual offenders, with alcohol misuse problems prominent amongst violent offenders and lifestyle and attitudinal problems prominent amongst

sexual offenders. Those committing offences of robbery were most commonly rated as presenting a high/very high risk of serious harm to the public generally, with combinations of interventions required to address their co-occurring internal and external risk factors.

- Addressing relationships and emotional wellbeing issues appears particularly important for female offenders, for whom concerns about vulnerability and self-harm were more prominent. For the youngest adult offenders, meeting education, training and employment needs and addressing problems relating to lifestyle and associates appears particularly important.
- Non-serious, non-persistent offenders were distinguishable through their absence of assessed problems, supporting the use of limited interventions. Non-White offenders were also distinguishable through their absence of assessed problems. While this could be seen as supporting the use of limited interventions, it may be that OASys is insufficiently capturing the specific offending-related problems of Minority Ethnic groups. Attention should thus be given to whether alternative questions should be incorporated within OASys to identify the needs of Minority Ethnic offenders.
- Offenders' self-assessment of their own likelihood of reoffending suggests that attention should be paid to their perceptions of the links between various problems and offending behaviour.

Conclusion

The completion of OASys assessments across the prison and probation services during 2008 and the collation of these assessments within a central database has enabled offender profiles to be generated using a sample of over 300,000 cases. These profiles aid understanding of offenders' differing risk levels and the underlying causes behind their offending, providing information pertinent to the targeting of interventions and the allocation of resources. Current levels of provision are indicated through the information recorded in the OASys sentence plan.

During 2009, changes are to be made to the content and scoring of OASys, implementing a number of the recommendations arising from the research presented in this volume on the tool's reliability and validity. These changes will impact upon the risk/need profiles produced for 2009. The OASys likelihood of reconviction score is to be replaced by two new improved predictors, one for general reoffending (OGP) and one for violent reoffending (OVP), while the questions used for scoring criminogenic needs are to be amended and the cut-off points recalibrated in relation to reoffending rates.

13. Use of OASys data elsewhere and compendium conclusions and recommendations

Compendium conclusions

The chapters of this compendium presented the research on OASys reliability and validity completed between 2006 and 2009, along with exploratory qualitative research on OASys content and a statistical summary of OASys data in the form of offender profiles. The research on reliability and validity provides evidence about the strengths and shortcomings of OASys as an assessment of offending-related risks and needs. This research enables OASys to clearly demonstrate the extent to which it meets the criterion for good systems of offender assessment described by Bonta *et al.* (2001), and summarised in the first chapter of this compendium. This gives reassurance to its continued use in assessment and offender management and to the use of OASys data in management information, research on offenders and evaluation of interventions designed to reduce reoffending. The following section in this chapter sets out the use of OASys data outside of O-DEAT research.

The use of these findings to inform the development of a layered OASys, including a reduced full-length assessment and shortened versions of OASys, ensures that the actions arising from the Strategic Review of OASys led to defensible changes that maintain the rigorous aspects of OASys. The table included at the end of this chapter describes the recommendations and responses arising from the O-DEAT research programme on OASys from 2006–2009.

While OASys continues to be used with the population targeted for assessment since its inception, and if the population remains essentially unchanged, it can be assumed that the reliability and validity of OASys remains fairly robust. Advice to international colleagues working in offender management, who express interest in adopting OASys, always includes a statement about the need to test the reliability and validity of its use with the intended population in order to make any necessary amendments to the content. In time, findings from such undertakings will provide useful comparisons for OASys use in England in Wales.

The established evidence base should assist with continued use of OASys as a robust system, when faced with possible pressure for change for cost-savings. Good practice in offender management begins with good assessment that fully informs decisions about individual offenders, in order to reduce risk and protect the public. OASys is a good assessment of offenders, demonstrated by several different types of research evidence from the literature and from direct empirical research of its use in practice.

There is always possible further research that can be undertaken on OASys. Current research underway at the time of publication includes monitoring use of layered OASys, trend analysis, and development of predictive validity for specific types of offending, such

as domestic violence. Some of the resources previously focused on empirical research on OASys have now been redirected to providing information to various users of OASys data, from many different sources, as the listing that follows illustrates.

The use of OASys data in strategy, policy and research reports

Completed OASys assessments provide large amounts of standardised information about offenders while they are being supervised by the prison and probation services.

Now that OASys is both automated and in general use, the collated data are being widely used by a range of bodies. The data have been used: (i) at the parliamentary level by the National Audit Office; (ii) at the centre of government by the Cabinet Office; (iii) within NOMS at both the national and regional levels; (iv) by the Prison and Probation Inspectorates in various thematic reviews; (v) within research reports conducted by or on behalf of the Home Office; (vi) by independent charities; and (vii) within academic reports.

National Audit Office

At the parliamentary level, OASys data were used by the National Audit Office in their 2008 review of the supervision of community penalties across England and Wales. The report focused upon changes over the course of offenders' orders, concluding that many presented a lower risk of serious harm and demonstrated positive changes in the factors contributing to their offending behaviour following a community order.

Cabinet Office

OASys data are being used by the Cabinet Office to measure levels of performance against two of the eight indicators underpinning the Public Service Agreement (PSA 16) for socially excluded adults (HM Government, 2007). Baseline OASys data for 2006/07 (n=51,488) indicated that 77% of offenders under probation supervision were living in settled and suitable accommodation at the end of their order or licence (National Indicator 143), while 36% were in employment at the end of their order or licence (National Indicator 144). The data are further broken down by regions and local authority areas.

National Offender Management Service

OASys data have been used by NOMS to inform national strategies and guides. For example, the 2006 delivery strategy for 'Working with Alcohol Misusing Offenders' used data from over 120,000 probation OASys assessments completed during 2004/05. The strategy noted that over one-third (37%) of these offenders had a current problem with alcohol use and a similar proportion (37%) had a problem with binge drinking. Nearly half (47%) had misused alcohol in the past and approximately one-third (32%) had violent behaviour related to their alcohol use. Nearly two-fifths (38%) were found to have a criminogenic need relating to alcohol misuse. Finally, over a quarter (27%) of the offenders had problems with their levels of motivation for tackling their alcohol misuse.

The 2008 National Service Framework for 'Improving Services to Women Offenders' used OASys data for the 12 months ending September 2007. The data indicated that women were six times more likely to be carers than men and twice as likely to live in a house with children. Women offenders were more likely to have emotional wellbeing and relationships needs than men; 46% of women offenders had been the victim of domestic violence compared to 6% of men; 27% were considered a suicide risk; and 27% were considered at risk of self-harm.

The 2008 'Offender Management Guide to Working with Women Offenders' included further OASys findings. It was reported that 62% of women offenders supervised by the probation service had a relationships need compared to 40% of men, and that 28% of women offenders identified finance as an issue contributing to their offending compared to 21% of men. It was further noted that poor scores on the attitudes section were associated with a 33% reduction in completion rates on programmes for women.

OASys data have also been used within NOMS to inform the regional and national commissioning plans which set out the priorities for investing and disinvesting in commissioned prisons and probation services. For example, the 2008/09 North West regional commissioning plan set out the criminogenic need prevalence rates for community-sentenced offenders alongside the percentages who received interventions designed to address these needs. A more detailed needs assessment compared OASys scores and criminogenic needs by region and by various offender groups within the North West.

At the national level, the 2008/09 commissioning plan for the high security estate reported findings based upon OASys assessments completed within the high security prisons during 2006/07 (n=643). Comparisons to the Category B estate (n=664) and all prisons (n=10,137) indicated that the prisoners within the high security estate were more likely to have a high risk of serious harm and a high likelihood of reconviction. The data were also used to indicate that there were varying degrees of misalignment between the levels of criminogenic needs and the levels of planned interventions.

Prison and Probation Inspectorates

OASys data have been used by the independent inspectorates of probation and prisons in a number of thematic reviews. For example, the 2006 review by HM Inspectorate of Probation on substance misuse work examined 687 cases that had started supervision in the community during 2004/2005. Within this sample, 38% of the offenders were found to have an OASys score of four or above for alcohol, and 21% had a similar score for drugs (with 9% having such a score for both alcohol and drugs). The most commonly misused substances after alcohol in the previous six months had been cannabis and heroin, with previous patterns of misuse also including frequent use of amphetamines, crack and cocaine. In 28% of the cases where OASys indicated a significant substance misuse problem, there was no corresponding sentence plan objective.

The 2007 review by HM Inspectorate of Prisons on the care and support of prisoners with mental health needs included analysis from OASys assessments completed during the financial year 2005/06. The OASys data were used to indicate that more women (55%) than men (30%) or young adults (25%) had problems in the area of emotional wellbeing. Those with emotional wellbeing needs had greater needs in all other areas associated with reoffending, with relationships emerging as the dominant need for those women (but not men) with emotional problems.

More recently, a 2008 joint inspection by the probation and prisons inspectorates on the indeterminate sentence for public protection (IPP) compared 54,785 valid OASys assessments for all prisoners (up to September 2007) against a subset of 2,204 assessments for IPP prisoners. The analysis indicated that the IPP prisoners had an average of 6.3 criminogenic needs compared to 4.4 for other prisoners. About two-thirds (68%) of the IPP prisoners presented a high risk of serious harm, but only 6% presented a very high risk of serious harm, while a quarter were assessed as low or medium risk. The IPP prisoners were found to have more mental health problems, and both IPP prisoners and lifers had a raised risk of self-harm and suicide (37%) compared to other prisoners (23%).

Home Office Research Reports

OASys data have been used within a number of research reports conducted by or on behalf of the Home Office. A 2005 review of 'What Works' in reducing reoffending reported early OASys data (10,000 assessments from 19 probation areas) which showed that offenders had an average of four criminogenic needs, with offenders in custody tending to have a greater number of needs (Harper *et al.*, 2005). Over half of the offenders had criminogenic needs relating to (i) education, training and employment and (ii) thinking and behaviour. Additionally, just over half of the custodial offenders had a need relating to lifestyle and associates, and they were more likely to have drug misuse problems than offenders on community sentences. The OASys data also indicated that female offenders had markedly higher levels of criminogenic need in the areas of relationships and emotional wellbeing, while male offenders had higher levels of need with regard to alcohol misuse, thinking and behaviour, and attitudes.

Another 2005 report used OASys data to inform an evaluation of the Intensive Control and Change Programme (ICCP), an intensive community sentence designed as an alternative to custody for 18- to 20-year-old offenders. OASys scores were used to target young adult offenders to the programme – eligible offenders were initially defined as those with an OASys score of 40 or more (medium to high risk of reoffending), which was then increased to 79+ with four or more previous convictions. The analysis revealed that offenders in the 11 pilot areas (April 2003 to March 2004) had an average OASys score of 83, and that they displayed higher levels of need in terms of (i) accommodation and (ii) education, training and employment than 18- to 20-year-olds serving other community sentences (Partridge *et al.*,

2005). The ICCP offenders also had higher levels of risk and needs in comparison to their custodial counterparts. When entering the offenders' OASys scores into a logistic regression model to predict breach and revocation, the analysis found that those with a low to medium OASys likelihood of reconviction score had odds of revocation 25% lower (using the upper confidence interval) than those with a high score.

A 2006 Home Office report used OASys data from 2004/05 to identify problem drug users, opiate users, crack cocaine users and those injecting drugs (Hay *et al.*, 2006). The data were then used alongside other local and national data sources to estimate the prevalence of problem drug misuse across England, employing capture/recapture and multiple indicator methods. The analysis led to an estimate of nearly 330,000 problem drug users (defined as 'opiate and/or crack cocaine users') across England.

A 2007 report on the national Prolific and other Priority Offender (PPO) programme explored the OASys profile of PPOs, using the most recent OASys assessments for the 4,067 PPOs (about 40% of the entire PPO population) who had an OASys assessment recorded between January and September 2005 (Dawson, 2007). To provide a comparator group, the most recent assessments (during the same period) of a random selection of 3,412 other offenders were used. The analysis found that the PPOs had greater education, training and employability needs and their accommodation needs were judged by OASys assessors to be more strongly linked to their offending behaviour than for the other offenders. PPOs were more likely to have misused drugs, but their alcohol misuse was typical of the wider offending population.

Independent charity reports

Independent charities have used OASys data to augment a number of policy reports. For example, in a 2007 paper, the National Association for the Care and Resettlement of Offenders (NACRO) used OASys data when setting out their position on offenders with mental health needs. Using different OASys samples, it was reported: (i) that 45% of offenders were identified as having an emotional wellbeing need, with women more likely to report problems such as feeling stressed, depressed, anxious or lonely; (ii) that 7% of offenders were at risk of suicide; and (iii) that 7% were at risk of self-harm. It was also noted that a third of offenders completing the OASys self-assessment questionnaire had said that they felt depressed and that one in ten said that this had contributed to their offending behaviour.

The Centre for Crime and Justice Studies at King's College London, an independent charity that informs and educates about crime and criminal justice, used OASys data in their 2008 'Community Sentences Digest'. In attempting to provide a picture of the multiple social needs of offenders on community sentences, the criminogenic need levels of those assessed during 2007–2008 were reported. In summary, it was stated that

“over half have basic education and training deficits, more than half are unemployed, nearly a third have an accommodation problem, nearly half have a mental health problem, close to a quarter have some kind of drug problem and almost half have an alcohol problem”

(Solomon and Silvestri, 2008:8).

Academic work

Finally, a number of academics have used OASys data in their work. For example, in their consideration of difference and diversity in probation, Gelsthorpe and McIvor (2007) reported OASys findings when considering the issue of mental health. They noted that

“within a sample of about 203,000 offenders (NPS 2005-2006 data), 13 per cent were recorded as having significant psychological problems/depression and 22 per cent were recorded as having some problems. A further 6 per cent were recorded as having significant psychiatric problems, with a further 9 per cent some problems.”

Recommendations

The findings that arose from the research presented in this compendium were discussed with policy leads within the OASys business team and many were used to inform the development of layered OASys, following the OASys Strategic Review published in 2008.

The O-DEAT research on OASys examined the extent to which it possesses the characteristics highlighted by Bonta in 2001 as desirable in any risk and needs assessment tool. Importantly, the research has contributed to modifications to OASys to improve its prediction of general and violent reoffending and to reduce its length and create layers for layered OASys, without compromising the other types of reliability and validity in which it has shown strength. A table of recommendations and responses is set out below.

Mia Debidin 2009

Table 13.1: Recommendations from O-DEAT research findings and responses from OASys business team

	Recommendations in the OASys research	Response
1. Data completion	<p>The completion rates in 2006/07 were generally good with the majority of items complete in over 85 per cent of assessments, and completion rates have improved since April 2005. The relationships section had the poorest completion rate in both prison and probation assessments. Family and partner criminal records as well as experience of childhood were the most problematic.</p> <ol style="list-style-type: none"> 1. Completion rates should be further raised so all items are complete in 90% of assessments. 2. Managers should explore the reasons for components with poor completion quality and establish improvement plans. 3. Individuals using OASys data to inform policy decisions, research or resource allocation should be aware of which OASys components are potentially less accurate because of poor completion rates. 4. Monitoring of OASys completion should be continued to maintain data knowledge and identify where further improvements may be necessary. 	<ol style="list-style-type: none"> 1. O-DEAT continues to monitor completion rates and to provide data at the individual level when requested, to assist management exploring the reasons for poor completion. 2. Currently completion reports are produced and disseminated to probation areas quarterly and prison data bi- annually. 3. The data included in management information produced by O-DEAT is filtered to include only data of sufficiently high quality.

	Recommendations in the OASys research	Response
2. Internal reliability and construct validity	<p>Analysing approximately 230,000 valid OASys assessments completed during 2006/07, six of the 11 scored OASys scales were found to have high internal reliability, clearly measuring a discrete characteristic, and four had adequate internal reliability. 'Relationships' was the only section with non-adequate internal reliability, indicating that the questions within the section were failing to measure a single factor.</p> <ol style="list-style-type: none"> 1. The construction of OASys could be improved through a reduction from 73 scored questions across 11 scales to 47 scored questions across the ten individual-level and social problem scales (sections 3 to 12). 2. The cut-off points for identifying criminogenic needs should be set in relation to reoffending rates. Adjustments in the allocation of resources would be required to ensure that interventions were available to address the revised criminogenic need levels. Offenders with 'high' levels of need should be distinguished from offenders with 'medium' levels of need to assist with the targeting of interventions, maximising the use of resources. 3. The optimum criminogenic need cut-off points for different age and gender subgroups should be recalculated once larger samples are available. Any widening in the targeting of OASys would increase the validity of the calculations to the complete prison and probation caseloads. 4. Across sections 3 to 12 of OASys, nine of the currently scored questions are not needed within the revised individual-level or social problem scales or the new violent and general reoffending predictors. These questions could be removed from OASys unless: (i) they are found to be helpful in assessing risk of serious harm; (ii) they serve another specific purpose for practitioners; and/or (iii) further research reveals that they could be improved through amendments to their wording or accompanying guidance. 5. The potential value of additional questions should be considered following an evaluation of the textual information recorded by assessors within each of the OASys sections. The initial focus should be placed upon the relationships, the lifestyle and associates and the emotional wellbeing sections, identifying alternative questions which are amenable to change and have stronger independent associations with reoffending. 	<ol style="list-style-type: none"> 1. These findings were considered in determining the content of layered OASys, and they helped to inform which questions to keep and to exclude, in order to retain an assessment with high internal reliability and validity. 2. Further research explores possible new questions for inclusion, covering protective factors. 3. Recommendation on criminogenic needs cut-off points has been considered as part of the development of layered assessments

	Recommendations in the OASys research	Response
3. Inter-rater reliability	<p>This study measured the ability of OASys to deliver consistent OASys assessments by asking multiple assessors to rate the same offender. Results showed that the reliability of OASys was moderate. The most reliable sections were: accommodation; lifestyle and associates; and drug misuse. Moderately reliable sections included: education, training and employability; relationships; emotional wellbeing; and attitudes. The least reliable sections were: financial management; alcohol misuse; thinking and behaviour; and risk of serious harm. The implication for sections with poor agreement is that similar offenders may be assessed differently and as a result experience different supervision and different interventions. This may also result in poor targeting of resources.</p> <p>1. The five questions with the poorest consensus should be removed in a revised version of OASys. These five questions are: 6.1 (current relationship with close family), 9.3 (level of alcohol use in the past), 11.3 (aggressive/controlling behaviour), 11.10 (concrete/abstract thinking), and 12.5 (attitude to community/society).</p> <p>2. The OASys guidance manual should be revised to clarify definitions for the sections on alcohol misuse and thinking and behaviour.</p> <p>3. The variation in the risk of serious harm component should be further explored, and problems addressed by revising the section or improving assessor training as necessary</p>	<p>1. These findings were considered in determining the content of layered OASys, and they helped to inform which questions to keep and to exclude, in order to retain an assessment with high inter-rater reliability.</p> <p>2. Specific advice on clarifying definitions are being considered for updating the help text/online version of the manual.</p>

	Recommendations in the OASys research	Response
4. Coverage and representativeness	<p>During 2007, OASys assessments had been completed in approximately three-quarters (76%) of all probation commencements and approximately two-thirds (66%) of all sentenced prisoner receptions. Assessments had not always been completed in those cases which were known to meet the post-sentence eligibility criteria for OASys and had often been completed in those cases which were known not to meet the criteria. For the probation commencements, the respective completion rates were 85% and 54%. For the sentenced prisoner receptions, the completion rates were 78% and 57%. In many of the 'non-eligible' cases the assessments had been completed pre-sentence alongside a standard delivery report for the court.</p> <p>Overall, offenders who had committed a violent or sexual offence or who had a high likelihood of reconviction were more likely to have had an assessment. The use of OASys was thus consistent with the expectation that resources should follow risk. There were differences in completion rates between subgroups when the analysis was restricted to those cases which did not meet the post-sentence eligibility criteria, demonstrating that OASys completion had been targeted at specific types of non-eligible case.</p> <p>The risk and need levels of all those offenders commencing supervision were lower than the risk and need levels of those for whom an assessment had been completed. The proportions of offenders who presented a low risk of serious harm increased by 7% for the probation commencements and 6% for the sentenced prisoner receptions. Consequently, the OASys samples were not fully representative of the entire offender caseloads.</p> <p>One in 50 of those probation commencements which did not meet the post-sentence eligibility for administering OASys had a high likelihood of reconviction and just 0.5% presented a high/very high risk of serious harm, demonstrating that they were a relatively low risk group. In contrast, over two-fifths (43%) of the sentenced prisoner receptions who did not meet the post-sentence eligibility for administering OASys had a high likelihood of reconviction.</p>	<ol style="list-style-type: none"> 1. These findings did not produce specific recommendations. 2. The effective targeting of OASys will be addressed to some extent by layered OASys. 3. There is a need to remind users of OASys that data are not representative of the offender population. 4. Continue to monitor non-completion through a prison backlog report; national standards in probation also driving this forward.

	Recommendations in the OASys research	Response
5. Predictive validity	<p>The new OASys Violence Predictor (OVP) greatly improves prediction of violence against the person, weapons, robbery, criminal damage and public order (“violent-type”) offences. The new OASys General Reoffending Predictor (OGP) improves prediction of other non-sexual (“general”) offences. Dynamic risk factors for violent-type offending are (strongest first): alcohol misuse; employability; attitudes; temper control; failure to recognise impact of offending; accommodation; current psychiatric treatment. Static factors are: previous OVP-type offending; young age; any criminal history; previous other offending; being male. Dynamic risk factors for general reoffending are (strongest first): drug misuse; accommodation; employability; regular activities encourage offending; attitudes; thinking and behaviour. OGRS 3 provides information on static risk.</p> <p>For both OVP and OGP, thinking and behaviour underlies most other risk factors, so scores should help targeting to offending behaviour programmes (which focus on thinking skills). Both predictors should help assessors to prioritise offenders under offender management and produce sentence plans which effectively target key risk factors. OVP addresses most serious further offending and can guide risk of serious harm assessment. NOMS should adopt them in place of the current OASys summary score, including necessary IT development, as a high priority. Consultation with OASys assessors – which has already begun – will ensure that OGP and OVP are presented in a user-friendly manner and support good practice.</p>	<p>1. These findings were used to support the early inclusion of OGP/OVP in new releases of OASys, specifically in Release 4.3.1.</p> <p>2. The tools were piloted in order to develop the guidance.</p>

	Recommendations in the OASys research	Response
<p>OASys textual analysis: Relationships</p>	<p>Analysis of the textual information recorded within the relationships section revealed a recurrent theme relating to children which is not covered in the fixed-response questions. The following sub-themes identified were: (i) parenthood; (ii) access to and contact with children; (iii) children as protective factors; (iv) relationships with children; (v) child care and parenting; and (vi) single parent and coping with having children. There was a clear gender divide with female offenders more likely to experience difficulties coping with having children coupled with other issues relating to poor mental health and alcohol dependency.</p> <p>In light of the findings, incorporating questions relating to offenders' parental responsibilities and problems in their relationships with their children should be considered. To ensure that the questions are fully dynamic, focusing on current parental responsibilities and current problems with their children would be preferable.</p> <p>These questions could replace the current questions on close family members and current partners having a criminal record (6.2 and 6.5) which have been found to be problematic in terms of completion by practitioners and, when completed, the internal reliability of the relationships section.</p> <p>The textual analysis also showed that the question on current relationship with partner (6.4) did not allow assessors to distinguish which offenders are in a relationship and those which were not. This question could be split according to status and quality of the current relationship.</p>	<p>These findings were considered in determining the content of layered OASys, and they helped to inform which new questions to include in OASys, in order to capture commonly occurring themes within the text boxes.</p> <p>These issues to be added to the OASys-R project issues log for future consideration post-release of the new IT system.</p>

	Recommendations in the OASys research	Response
OASys textual analysis: Lifestyle and associates	<p>Analysis of the free-text responses recorded within the lifestyle and associates section revealed a recurrent theme of ‘gangs’ which is not covered in the fixed-response questions.</p> <p>A question on gang membership could therefore be incorporated into OASys, along with a clear definition for the term ‘gang’. To ensure that the question is fully dynamic, focusing on current rather than previous gang membership and involvement would be preferable.</p> <p>Given the on-going concerns expressed in some research findings about accurately identifying gang members and assessing gang involvement, a more general question on the presence of ‘many’ criminal associates could also be incorporated into OASys.</p> <p>These questions could replace the current questions on community integration and manipulative/predatory lifestyle (7.1 and 7.4) which have been found to have limited predictive validity and to be problematic in terms of the construct validity of the core OASys assessment.</p> <p>Through these changes, there will be an increased focus on ‘associates’ rather than ‘lifestyle’, adhering to the research evidence which has concluded that ‘anti-social peers and associates’ is one of the four major risk factors for offending.</p> <p>Prior to making any changes to the ‘relationships’ and ‘lifestyle/associates’ sections of OASys, the OASys business team and potentially the OASys user group will need to be consulted. Further research will be required to test the reliability and validity of the new questions alongside the existing OASys questions, assessing how well the items measure a single distinct domain (internal reliability/construct validity) and how well the items predict reoffending (predictive validity).</p>	To be added to issues log for consideration of OASys-R post release.

References

- Andrews, D.A. (1983) 'The Assessment of Outcome in Correctional Samples'. In M.L. Lambert, E.R. Christenson, and S.S. De Julio (eds.) *The measurement of psychology outcome*. New York: Wiley.
- Andrews, D.A. (1995) 'The psychology of Criminal Conduct and Effective Treatment'. In J. McGuire (ed.) *What works: reducing offending*. Chichester: Wiley.
- Andrews, D.A. and Bonta, J.L. (1994) *The psychology of criminal conduct*. Cincinnati: Anderson.
- Andrews, D.A. and Bonta, J.L. (1995) *LSI-R: The Level of Service Inventory – Revised*. Toronto: Multi-Health Systems.
- Andrews, D. and Bonta, J. (2003) *The Psychology of Criminal Conduct* (3rd edition). Cincinnati: Anderson.
- Andrews, D. and Bonta, J. (2005) *The Level of Service Inventory – Revised Manual*. Toronto: Multi-Health Systems Inc.
- Andrews, D.A., Bonta, J. and Wormith, J.S. (2006) 'The recent past and near future of risk and/or need assessment', *Crime & Delinquency*, 52(1): 7-27.
- Baker, K., Jones, S., Roberts, C. and Merrington, S. (2002) *Validity and reliability of ASSET: findings from the first two years of the use of ASSET*. Oxford: Probation Studies Unit, Centre for Criminological Research, University of Oxford.
- Black, N. and Payne, M. on behalf of the DoCDat Development Group (2003) 'Directory of clinical databases: improving and promoting their use', *Quality and Safety in Health Care*, 12: 348-352.
- Blanchette, K. (2002) 'Classifying female offenders for effective intervention: Application of the case-based principles of risk and need'. *Forum on Correctional Research*, 14, 31-35.
- Blumstein, A., Cohen, J., Roth, J. and Visher, C. (1986) *Criminal careers and career "criminals"*, Washington DC: National Academy Press.
- Boles, S.M., and Miotto, K. (2003). 'Substance abuse and violence: a review of the literature'. *Aggression and Violent Behavior*, 8, 155-174.

Bonta, J., Law, M. and Hanson, K. (1998) 'The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis'. *Psychological Bulletin*, 123, 123-142.

Bonta, J., Pang, B. and Wallace-Capretta, S. (1995) 'Predictors of recidivism among incarcerated female offenders'. *The Prison Journal*, 75, 277-294.

Bonta, J., Bogue, B., Crowley, M. and Motiuk, L. (2001) Implementing offender classification systems: lessons learned. In Bernfeld, G.A., Farrington, D. P., and Leschied, A.W. *Offender Rehabilitation in Practice*. John Wiley and Sons.

Bowles, R.A. and Florackis, C. (2007) 'Duration of the time to reconviction: Evidence from UK prisoner discharge data'. *Journal of Criminal Justice*, 35, 365-378.

Broadhurst, R. (2000) 'Criminal careers, sex offending and dangerousness'. In Brown, M. and Pratt, J. (eds.) *Dangerous Offenders: Punishment and Social Order*, London: Routledge.

Brown, S. and Motiuk, L.L. (2005) *The Dynamic Factor Identification and Analysis (DFIA) Component of the Offender Intake Assessment (OIA) Process: A Meta-Analytic, Psychometric and Consultative Review*. Ottawa: Correctional Service Canada.

Bureau of Justice Statistics (1999) *Women offenders*. Washington, DC: US Department of Justice, Office of Justice Programs.

Butler, G., Hodgkinson, J., Holmes, E. and Marshall, S. (2004) *Evidence Based Approaches to Reducing Gang Violence: A Rapid Evidence Assessment for Aston and Handsworth Operational Group*. AHOG Paper.

Calverley, A., Cole, B., Kaur, G., Lewis, S., Raynor, P., Sadeghi, S., Smith, D., Vanstone, M. and Wardak, A. (2004) *Black and Asian Offenders on Probation*, Home Office Research Study 277, London: Home Office.

Campbell, M. A., French, S. and Gendreau, P. (2007) Assessing the utility of risk assessment tools and personality measures in the prediction of violent recidivism for adult offenders. *Public Safety Canada Research Reports 2007-04*.

Carter, P. (2003) *Managing Offenders, Reducing Crime*. London: Strategy Unit.

Chitty, C. (2004) 'The impact of corrections on reoffending: conclusions and the way forward', in Harper, G. and Chitty, C. (eds.) *The Impact of Corrections on Reoffending: A Review of 'What Works'*, Home Office Research Study 291, London: Home Office.

- Cooke, D. (1989) 'Containing violent prisoners: an analysis of the Barlinnie special unit'. *British Journal of Criminology*, 29, 129-143.
- Cooke, D. and Michie, C. (2001). 'Refining the construct of psychopathy: Towards a hierarchical model'. *Psychological Assessment*, 13, 171-188.
- Copas, J.B., Ditchfield, J. and Marshall, P. (1994) *Development of a new risk predictor score*. Research Bulletin 36. London: Home Office.
- Copas, J.B. and Marshall, P. (1998) 'The offender group reconviction scale: a statistical reconviction score for use by probation officers'. *Journal of the Royal Statistical Society: Applied Statistics*, 47(1), 159-171.
- Copas, J. (1992, unpublished) *Statistical analysis for a risk of reconviction predictor*, Report to the Home Office, Coventry: University of Warwick.
- Cornish, D. and Clark, R.V. (1975) *Residential treatment and its effects on delinquency*. Home Office Research Unit Study 32. London: HMSO.
- Cottle, C., Lee, R. and Heilbrun, K. (2001) 'The prediction of criminal recidivism in juveniles'. *Criminal justice and behaviour*, 28(3) 367-394.
- Coulbeck, R. (2004) *Risk of harm issues in the National Probation Service: Challenges, perspectives and ideas for change*. London: National Probation Directorate.
- Coulson, G., Ilacqua, G., Nutbrown, V., Giulekas, D. and Cudjoe, F. (1996) 'Predictive utility of the LSI for incarcerated female offenders'. *Criminal Justice and Behaviour*, 23, 427-439.
- Covington, S.S. (2003) 'A woman's journey home: Challenges for female offenders'. In J. Travis and M.Waul (eds.) *Prisoners once removed: The impact of incarceration and reentry on children, families and communities* (pp. 67-103). Washington, DC: The Urban Institute.
- Crawford, A. (2007) 'What impacts on quality assessment in using OASys?' *Probation Journal* 54(2): 157-170.
- Cunliffe, J. and Shepherd, A. (2007). *Reoffending of adults: results from the 2004 cohorts*. Home Office Statistical Bulletin 06/07. London: Home Office.
- Daly, K. (1994) *Gender, crime and punishment*. New Haven, CT: Yale University Press.

Davies, P. (2003) *The Magenta Book: Guidance notes for policy evaluation and analysis*. London: Cabinet Office.

Dawson, P. (2007) *The National PPO evaluation – research to inform and guide practice*, Home Office Online Report 09/07, London: Home Office.
<http://www.homeoffice.gov.uk/rds/pdfs07/rdsolr0907.pdf>

Douglas, K. S., Guy, L. S., and Weir, J. (2006). HCR-20 violence risk assessment scheme: Overview and annotated bibliography. Burnaby, Canada: Department of Psychology, Simon Fraser University.

Douglas, K.S., Yeomans, M. and Boer, D.P. (2005) 'Comparative Validity Analysis of Multiple Measures of Violence Risk in a Sample of Criminal Offenders'. *Criminal Justice and Behavior*, 32, 479-510.

Ebenson, F. and Huizinga, D. (1993) 'Gangs, drugs, and delinquency in a survey of urban youth'. *Criminology*, 35, 561-575.

Ellickson, P.L., Hays, R.D. and Bell, R.M. (1992) 'Stepping through the drug use sequence: Longitudinal scalogram analysis of initiation and regular use'. *Journal of Abnormal Psychology*, 101(3), 441-451.

Endicott, J., Tracy, K., Burt, D., Olson, E., and Coccaro, E.F. (2002). 'A novel approach to assess inter-rater reliability in the use of the Overt Aggression Scale-Modified'. *Psychiatry Research* 112 (2) 153– 59.

Farrington, D.P. (1989) *Cambridge Study in delinquent development: long-term follow-up*. Cambridge: University of Cambridge Institute of Criminology.

Farrington, D.P. (1990) Implications of criminal career research for the prevention of offending. *Journal of adolescence*, 13, 93-113.

Farrington, D. (1992) 'Criminal career research in the United Kingdom', *British Journal of Criminology*, 32: 521-536.

Farrington, D.P., Gallagher, B., Morley, L., St. Ledger, R.J. and West, D.J. (1986) 'Unemployment, school leaving and crime'. *British Journal of Criminology*, 26(4) 335-356.

Farrington, D.P. (2002) 'Developmental criminology and risk-focussed prevention'. In M. Maguire, R. Morgan and R. Reiner (eds.) *The Oxford Handbook of Criminology*. Oxford: Clarendon Press.

Fitzgibbon, D.W. (2007) 'Risk analysis and the new practitioner. Myth or reality?' *Punishment & Society* 9(1): 87-97

Fitzgibbon, D.W. (2008) 'Fit for purpose? OASys assessments and parole decisions'. *Probation Journal* 55(1):55-69.

Fleiss, J. L. (1986) *The Design and Analysis of Clinical Experiments*. New York: Wiley.

Francis, B., Soothill, K., and Humphreys, L. (2007) *Development of a reoffending measure using the Police National Computer database*. Lancaster University.

Ge, X., Donnellan, M.B., Wenk, E. (2001) 'The development of persistent criminal offending in males', *Criminal Justice and Behavior*, 28(6): 731-755.

Gelsthorpe, L.R. (1999) 'Reconstructing a Woman's Prison. The Holloway Redevelopment Project 1968-88'. *The British Journal of Sociology*, 50 (2), 359-360.

Gelsthorpe, L. and McIvor, G. (2007) 'Difference and diversity in probation', in L. Gelsthorpe and R. Morgan (eds.) *Handbook of Probation*, Cullompton, Devon: Willan Publishing.

Gendreau, P., Little, T. and Goggin, C. (1996) 'A meta-analysis of the predictors of adult recidivism: What works?' *Criminology*, 34(4), 575-608.

Girard, L. and Wormith, J.S. (2004) 'The predictive validity of the Level of Service Inventory – Ontario Revision on general and violent recidivism among various offender groups'. *Criminal Justice and Behaviour*, 31, 150-181.

Harcourt, B.E. (2007) *Against Prediction. Profiling, Policing and Punishing in an Actuarial Age*. Chicago and London: University of Chicago Press.

Hare, R.D., McPherson, L.M. and Forth, A.E. (1993) 'Male psychopaths and their criminal careers'. *Journal of consulting and clinical psychology*, 56, 710-714.

Hare, R.D. (2004) *Hare Psychopathy Checklist – Revised (PCL-R): 2nd Edition*. Toronto: Multi-Health Systems Inc.

Harer, M. D. (1995) *Prison Education Program Participation and Recidivism: A Test of the Normalization Hypothesis*. ORE report.

Harper, G., Man, L., Taylor, S. and Niven, S. (2005) 'Factors associated with offending'. In G. Harper and C. Chitty (eds.) *The impact of corrections on reoffending: a review of 'what works'*. Home Office Research Study 291. London: Home Office.

Hart, S.D., Kropp, P. R. and Hare, R.D. (1988) Performance of male psychopaths following conditional release from prison. *Journal of Consulting and Clinical Psychology*, 56, 227-232.

Hart, S.D., Cox, D.N., and Hare, R.D. (1995). *The Hare Psychopathy Checklist: Screening Version (PCL-SV)*. Toronto: Multi-Health Systems Inc.

Haslewood-Pócsik, I. (2001) *Practitioner ACE Assessments and Offender Self-Assessments: A Comparison*, ACE Practitioner Bulletin 5, Oxford: Probation Studies Unit.

Hay, G., Gannon, M., MacDougall, J., Millar, T., Eastwood, C. and McKeganey, N. (2006) 'Local and national estimates of the prevalence of opiate use and/or crack cocaine use (2004/05)'. In N. Singleton, R. Murray and L. Tinsley (eds.) *Measuring different aspects of problem drug use: methodological developments*, Home Office Online Report 16/06, London: Home Office.

Heidensohn, F. (2002) 'Gender and Crime'. In M. Maguire, R. Morgan and R. Reiner (eds.) *The Oxford Handbook of Criminology*. Oxford: Clarendon Press.

Hester, M. and Westmarland, N. (2006) 'Domestic Violence Perpetrators'. *Criminal Justice Matters*, 66(1), 34-35.

HM Chief Inspector of Prisons and HM Chief Inspector of Probation (2008) *The indeterminate sentence for public protection*, A thematic review, London: HM Inspectorate of Prisons and HM Inspectorate of Probation.

HM Government (2006) *A Five Year Strategy for Protecting the Public and Reducing Reoffending*. London: The Stationery Office.

HM Government (2007) *PSA Delivery Agreement 16: Increase the proportion of socially excluded adults in settled accommodation and employment, education or training*, London: Home Office.

HM Inspectorate of Prisons (2007) *The mental health of prisoners: A thematic review of the care and support of prisoners with mental health needs*, London: HM Inspectorate of Prisons.

HM Inspectorate of Probation (2006) *'Half Full and Half Empty': An Inspection of the National Probation Service's substance misuse work with offenders*, An Effective Supervision Inspection Programme Thematic Report, London: HM Inspectorate of Probation.

HM Prison Service (2005) *Offender Assessment and Sentence Management – OASys*, Prison Service Order 2205, London: HM Prison Service.

- Hollin, C.R. (2002) 'Risk-Needs assessment and allocation to offender programmes'. In J. McGuire (ed.) *Offender rehabilitation and treatment*. Chichester: Wiley.
- Hollin, C.R. and Palmer, E.J. (1995) *Education and Work Programmes in Prisons: Effect on Recidivism. Report for the Planning Group HM Prison Service*. HM Prison Service: London.
- Hollin, C.R and Palmer, E.J. (2006a) 'The Level of Service Inventory – Revised profile of English prisoners: risk and reconviction analysis'. *Criminal Justice and Behaviour*, 33, 347-366.
- Hollin, C.R. and Palmer, E.J. (2006b) 'Criminogenic need and women offenders: A critique of the literature', *Legal and Criminological Psychology*, 11: 179-195.
- Home Office (1999) *A Joint Risk/Needs Assessment System for the Prison and Probation Services*, Probation Circular 16/1999, London: Home Office.
- Home Office (2001) *Making Punishments Work: Report of a Review of the Sentencing Framework for England and Wales*, London: Home Office.
- Home Office (2002) *Offender Assessment System: OASys: User manual*. London: Home Office.
- Home Office (2006) Criminal Statistics 2005 England and Wales. *Home Office Statistical Bulletin 19/05*. London: Home Office.
- Horsefield, A. (2003) 'Risk assessment. Who needs it?' *Probation Journal* 50(4) 374-379.
- Hosmer, D.W. and Lemeshow, S. (2000) *Applied Logistic Regression*, New York: John Wiley and Son.
- Howard, P., Clark, D. and Garnham, N. (2006) *An evaluation and validation of the Offender Assessment System (OASys)*. OASys Central Research Unit. Report to HM Prison Service and National Probation Service.
- Howard, P. (2006). *The Offender Assessment System: an evaluation of the second pilot*. Home Office Research Findings 278. London: Home Office.
- Howard, P., Francis, B., Soothill, K., and Humphreys, L. (2009). *OGRS 3: the revised Offender Group Reconviction Scale*. London: Ministry of Justice, Research Summary 7/09.

- Howard, P. (2009) *Improving the prediction of reoffending using the offender assessment system*. London: Ministry of Justice, Research Summary 2/09.
- Howard, P. and Moore, R. (2009) *Measuring changes in risk and need over time using OASys*. London: Ministry of Justice Research Summary 10/09.
- Huebner, B.M., Varano, S.P. and Bynum, T.S. (2007) 'Gangs, guns, and drugs: Recidivism among serious young offenders'. *Criminology and Public Policy*, 6, 187-222.
- Kaestner, R. (1993) 'The effect of illicit drug use on the labor supply of young adults'. *Journal of Human Resources*, 29, 126-155.
- Kandel, D.B. and Yamaguchi, K. (1993) 'From beer to crack: Developmental patterns of drug involvement'. *American Journal of Public Health*, 83(6), 851-855.
- King, R.D., Massoglia, M. and Macmillan, R. (2007) 'The context of marriage and crime: gender, the propensity to marry and offending in early adulthood'. *Criminology*, 45, 33-65.
- Kubrin, C.E. and Stewart, E.A. (2006) 'Predicting who reoffends: the neglected role of neighbourhood context in recidivism studies'. *Criminology*, 44(1), 165-197.
- Kurlycheck, M.C., Brame, R. and Bushway, S.D. (2006) 'Scarlet letters and recidivism: does an old criminal record predict future offending?' *Criminology and Public Policy*, 5(3), 483-504.
- Lancaster, E. and Lumb, J. (2006) 'The assessment of risk in the National Probation Service in England and Wales'. *Journal of Social Work* 6(3) 275-291.
- Landis, J. R., Koch, G. G. (1977) 'The measurement of observer agreement for categorical data'. *Biometrics* 33:159-174.
- Langan, P.A. and Levin, D.J. (2002) *Recidivism of prisoners released in 1994*. Bureau of Justice Statistics Special Report. Washington, DC: US Department of Justice, Office of Justice Programs.
- Lanza-Kaduce, L., Parker, K. and Thomas, C. (1999) 'A Comparative Recidivism Analysis of releases form Private and Public Prisons'. *Crime & Delinquency*, 45, 28-47.
- Laub, J.H., Nagin, D.S. and Sampson, R.J. (1998) 'Trajectories of change in criminal offending: Good marriages and the desistance process'. *American Sociological Review*, 63, 225-238.
- Laub, J.H. and Sampson, R.J. (2003) *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*. Cambridge, MA: Harvard University Press.

Lipsey, M. (1995) 'What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents?' In J. McGuire (ed.) *What works: reducing offending*. Chichester: Wiley.

Lloyd, C., Mair, G. and Hough, M. (1995) *Explaining reconviction rates: a critical analysis*. Home Office Research Study 136. London: HMSO.

Lowenkamp, C.T., Holsinger, A.M. and Latessa, E.J. (2001) 'Risk/need assessment, offender classification, and the role of childhood abuse'. *Criminal Justice and Behaviour*, 28, 543-563.

Loza, W., Cumbleton, A., Shahinfar, A., Hong Neo, L., Evans, M., Conley, M. and Summers, R. (2004) 'Cross-Validation of the Self-Appraisal Questionnaire (SAQ): An Offender Risk and Need Assessment Measure on Australian, British, Canadian, Singaporean, and American Offenders', *Journal of Interpersonal Violence*, 19(10):1172-1190.

Maden, A., Rogers, P., Watt, A., Lewis, G., Amos, T., Gournay, K. and Skapinakis, P. (2006) *Assessing the utility of the Offenders Group Reconviction Scale-2 in predicting the risk of reconviction within 2 and 4 years of discharge in English and Welsh Medium Secure Units*. National Forensic Mental Health R&D Programme report. Available at: <http://www.nfmhp.org.uk/Maden/%20Lay%20summary.doc>

Mair, G., Burke, L. and Taylor, S. (2006) 'The worse tax form you've ever seen'? Probation officers' views about OASys. *Probation Journal* 53(1): 7-23.

Martin, C., Player, E. and Liriano, S. (2003) *Results of evaluations of the RAPt drug treatment programme*. Home Office Research Study 267. London: Home Office.

May, C. (1999) *Explaining reconviction following a community sentence: the role of social factors*. Home Office Research Study 192. London: Home Office.

Merrington, S. (2004) 'Assessment tools in probation: their development and potential'. In Burnett, R. and Roberts, C. (eds.) *What works in Probation and Youth Justice*, Cullompton: Willan Publishing.

McCarthy, P., Laing, K. and Walker, J. (2004) *Offenders of the Future? Assessing the Risk of Children and Young People Becoming Involved in Criminal or Antisocial Behaviour*, Nottingham, DfES Research Report RR545.

McGraw, K. O. and Wong, S. P. (1996) 'Forming Inferences About Some Intra-class Correlation Coefficients'. *Psychological Methods* 1: 30-46.

- McGuire, J. (1995) *What Works: Reducing Reoffending Guidelines from Research and Practice*. Chichester: Wiley.
- Merrington, S. and Skinnis, J. (2000) *Using ACE to profile criminogenic needs*. Probation Studies Unit: ACE Practitioner Bulletin No. 1.
- Mills, J.F., Kroner, D.G. and Hemmati, R. (2004) 'The Measures of Criminal Attitudes and Associates (MCAA): The prediction of general and violent recidivism'. *Criminal Justice and Behaviour*, 31(6), 717-733.
- Moore, T.M. and Stuart, G.L. (2005) 'A review of the literature on marijuana and interpersonal violence'. *Aggression and Violent Behaviour*, 10, 171-192.
- Moore, R. (2007). *Adult offenders' perceptions of their underlying problems: findings from the OASys self-assessment questionnaire*. Home Office Research Findings 284. London: Home Office.
- Moore, R. (2009) *The coverage and representativeness of OASys risk and need offender profiles: 2007 probation commencements and sentenced prisoner receptions*. London: Ministry of Justice, Research Summary 3/09.
- Moore, R. (2009) *Predicting reoffending with the OASys self-assessment questionnaire*. London: Ministry of Justice, Research Summary 5/09.
- Moore, R. (2009) *The internal reliability and construct validity of the Offender Assessment System (OASys)*. London: Ministry of Justice, Research Summary 6/09.
- Morton, S. (2009) *Can OASys deliver consistent assessments of offenders? Results from the inter-rater reliability study*. London: Ministry of Justice, Research Summary 1/09.
- Morton, S. (2009) *Findings from the measurement of OASys completion rates*. London: Ministry of Justice, Research Summary 4/09.
- Motiuk, M. S., Motiuk, L. L. and Bonta, J. (1992) 'A comparison between self-report and interview-based inventories in offender classification', *Criminal Justice and Behavior*, 19(2), 143-159.

Motiuk, L.L. and Blanchette, K. (2000) Assessing women offenders: What works. In M McMahon (ed.) *Assessment to Assistance: Programs for women in community corrections*. Virginia: American Correctional Association.

NACRO (2007) *Effective mental healthcare for offenders: the need for a fresh approach*, Mental health and crime policy briefing. London: NACRO.

Nagin, D.S. and Paternoster, R. (2000) 'Population heterogeneity and state dependence: State of the evidence and directions for future research'. *Journal of Quantitative Criminology*, 16, 117-144.

National Audit Office (2008) *The National Probation Service: the supervision of community orders in England and Wales*, London: The Stationery Office.

National Offender Management Service (2004) *National Reducing Reoffending Delivery Plan*. London: Ministry of Justice.

National Offender Management Service (2006) *Working with Alcohol Misusing Offenders – a strategy for delivery*, London: National Offender Management Service.

National Offender Management Service (2007a). *National Standards for the Management of Offenders. Standards and Implementation Guidance*. London: Home Office.

National Offender Management Service (2007b) *Pre-Sentence Report Decision Tool*. Probation Circular 12/2007, London: National Offender Management Service.

National Offender Management Service (2008a) *National Rules for Tiering Cases and Associated Guidance*, Probation Circular 08/2008, London: National Offender Management Service.

National Offender Management Service (2008b) *National Commissioning Plan 2008-2009: High Security Estate*, London: NOMS.

National Offender Management Service (2008c) *National Service Framework: Improving Services to Women Offenders*, London: NOMS.

National Offender Management Service (2008d) *Regional Commissioning Plan for the North West 08-09*, London: NOMS.

National Offender Management Service (2008e) *The Offender Management Guide to Working with Women Offenders*, London: NOMS.

National Probation Directorate (2005) *National Standards 2005*, London: National Probation Directorate.

Nilsson, A. (2003) 'Living conditions, social exclusion and recidivism among prison inmates'. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 4, 57-83.

Niven, S. and Olagundoye, J. (2002) *Jobs and homes – a survey of prisoners nearing release*. Home Office Findings 173. London: Home Office.

Niven, S. and Stewart, D. (2005) *Resettlement outcomes on release from prison in 2003*. Home Office Findings 248. London: Home Office.

Novaco, R.W. (1997) 'Remediating anger and aggression with violent offenders'. *Legal and criminological psychology*, 2, 77-88.

Nuttall, C.P. (1996) *A note on first and second prison sentences with special reference to parole*. Home Office Research Unit. London: Home Office.

Oddone-Paolucci, E., Violato, C. and Schofield, M.A. (2000) *A review of marital and family variables as they relate to adult criminal recidivism*. National Foundation for Family Research and Education.

Oldfield, M. (1996) *The Kent reconviction survey: a 5-year survey of reconvictions amongst offenders made subject to probation orders in 1991*. Maidstone: Kent Probation Service.

Oliver, M.E., Wong, S.C., Nicholaichuk, T. and Gordon, A. (2007). 'The validity and reliability of the Violence Risk Scale – Sexual Offender version: assessing sex offender risk and evaluating therapeutic change'. *Psychological Assessment* 19, 318-329.

Oppenheim, A.N. (1996) *Questionnaire Design, Interviewing and Attitude Measurement*, London: PINTER.

Owen, B. (1998) *In the mix: Struggle and survival in a woman's prison*. Albany, NY: State University of New York Press.

Palmer, E.J. and Hollin, C.R. (2004) 'Predicting reconviction using the Psychological Inventory of Criminal Thinking Styles with English Prisoners'. *Legal and Criminological Psychology*, 9, 57-68.

Palmer, E.J. and Hollin, C.R. (2007) 'The Level of Service Inventory – Revised with English Women Prisoners: a needs and reconviction analysis'. *Criminal Justice and Behaviour*, 34(8), 971-984.

Partridge, S., Harris, J., Abram, M. and Scholes, A. (2005) *The Intensive Control and Change Programme pilots: a study of implementation in the first year*, Home Office Online Report 48/05, London: Home Office.

Peersen, M., Sigurdsson, J.F., Gudjonsson, G.H. and Gretarsson, J. (2004) 'Predicting reoffending: a 5-year prospective study of Icelandic prison inmates'. *Psychology, Crime & Law*, 10(2), 197-204.

Phillips, H.K., Gray, N.S., MacCulloch, S.I., Taylor, J., Moore, S.C., Huckle, P. and MacCulloch, M.J. (2005) 'Risk assessment in offenders with mental disorders. Relative efficacy of personal demographic, criminal history, and clinical variables'. *Journal of interpersonal violence*, 20(7), 833-847.

Quinsey, V.L., Harris, G.T., Rice, M.E. and Cormier, C. (1998) *Violent Offenders: Appraising and Managing Risk*. Washington, DC: American Psychological Association.

Raynor, P., Kynch, J., Roberts, C. and Merrington, S. (2000) *Risk and need assessment in probation services: an evaluation*. Home Office Research Study 211. London: Home Office.

Raynor, P. (2007) 'Risk and need assessment in British probation: the contribution of LSI-R', *Psychology, Crime & Law*, 13(2): 125-138.

Reisig, M.D., Holtfreter, K. and Morash, M. (2006) 'Assessing recidivism risk across female pathways to crime'. *Justice Quarterly*, 23, 384-405.

Ritchie, J. and Lewis, J. (2003) *Qualitative research practice: A guide for social science students and researchers*. London, UK: Sage Publications.

Roberts, C., Burnett, R., Kirby, A. and Hamill, H. (1996) *A System for Evaluating Probation Practice*, Probation Studies Unit Report 1, Oxford: Centre for Criminological Research.

Robinson, G., (2003) 'Implementing OASys. Lessons from Research into LSI-R and ACE'. *Probation Journal* 50(1) 30-40.

Robinson, D. (1995) *The impact of cognitive skills training on post-release recidivism among Canadian Federal Offenders*. Ottawa: Correctional Services of Canada.

Ross, R.R. and Fabiano, E.P. (1985) *Time to think: a cognitive model of delinquency prevention and offender rehabilitation*. Johnson City: Institute of Social Sciences and Arts.

Rutter, M., Giller, H. and Hagell, A. (1998) *Antisocial behaviour by young people*. Cambridge: Cambridge University Press.

Salmelainen, P. (1995) *The Correlates of Offending Frequency: A Study of Juvenile Theft Offenders in Detention*. NSW Bureau of Crime Statistics and Research, Sydney.

Sampson, R.J. and Laub, J.H. (1993) *Crime in the making: Pathways and turning points through life*. Cambridge, MA: Harvard University Press.

Schroeder, R.D., Giordano, P.C. and Cernkovich, S.A. (2007) 'Drug use and desistance processes'. *Criminology*, 45(1), 191-222.

Shepherd, A. and Whiting, E. (2006) *Reoffending of adults: results from the 2003 cohort*, Home Office Statistical Bulletin 20/06, London: Home Office.

Shrout, P. E., and Fleiss, J. L. (1979). 'Intra-class correlations: Uses in assessing rater reliability'. *Psychological Bulletin*, 86, 420-428.

Sigmour, D.J. (1997) 'The Criminal Sentiments Scale-Modified and Pride in Delinquency'. *Criminal Justice and Behaviour*, 24(1), 52-70.

Sigmour, D.J. (2004) Use of dynamic risk/need assessment instruments among long-term incarcerated offenders. *Criminal Justice and Behaviour*, 31(3), 306-323.

Social Exclusion Unit (2002) *Reducing reoffending by ex-prisoners*. London: Office of the Deputy Prime Minister.

Solomon, E. and Silvestri, A. (2008) *Community Sentences Digest*, London: Centre for Crime and Justice Studies.

Soothill, K., Francis, B., and Fligelstone, R. (2002). 'Patterns of offending behaviour: a new approach'. *Home Office Research Findings*, 171. London: Home Office.

Sorbello, L., Eccleston, L., Ward, T. and Jones, R. (2002) 'Treatment needs of female offenders: a review'. *Australian Psychologist*, 37(3), 198-205.

Spicer, K. and Glicksman, A. (2004) *Adult reconviction: results from the 2001 cohort*. Home Office Online Report 59/04. London: Home Office.

- Spivak, A.L. and Damphousse, J.R. (2006) 'Who returns to prison? A survival analysis of recidivism among adult offenders released in Oklahoma, 1985-2004'. *Justice Research and Policy*, 8(2), 57-88.
- Spohn, C. and Holleran, D. (2002) 'The effect of imprisonment on recidivism rates of felony offenders: a focus on drug offenders'. *Criminology*, 40(2), 329-358.
- Sussman, S., Stacy, A.W., Dent, C.W., Simon, T.R. and Johnson, C.A. (1996) 'Marijuana use: Current issues and new research directions'. *Journal of Drug Issues*, 26(4), 695-733.
- Sutherland, E.H. and Cressey, D.R. (1970) *Criminology*. Philadelphia: Lippincott.
- Taylor, S.P. and Hulsizer, M.R. (1998) 'Psychoactive drugs and human aggression'. In R.G. Geen and E. Donnerstei (eds.) *Human aggression: Theories, research and implications for social policy* (pp. 139-165). San Diego, CA: Academic Press.
- Taylor, R. (1999) *Predicting reconvictions for sexual and violent offences using the revised Offender Group Reconviction Scale*. Research Findings 104. London: Home Office.
- Walters, G.D. (1992) *Foundations of Criminal Science 2: The use of knowledge*. New York: Praeger.
- Walters, G.D. (1996) 'The Psychological Inventory of Criminal Thinking Styles. Part III. Predictive validity'. *International Journal of Offender Therapy and Comparative Criminology*, 40, 105-112.
- Webster, C.D., Douglas, K.S., Eaves, D. and Hart, S.D. (1997) *HCR-20: Assessing risk for violence*. Burnaby, Canada: Simon Fraser University.
- Webster, C., MacDonald, R. and Simpson, M. (2006) 'Predicting criminality? Risk factors, neighbourhood influence and desistence'. *Youth Justice*, 6(1) 7-22.
- Wright, R.T. and Decker, S.H. (1997) *Armed robbers in action: Stickups and street culture*. Boston: Northeastern University Press.
- Wright, K.N. and Wright, K.E. (1992) 'Does getting married reduce the likelihood of criminality? A review of the literature'. *Federal Probation*, 3, 50-56.
- Youth Justice Board (2005) *Risk and Protective Factors*, London: Youth Justice Board for England and Wales.

Zamble, E. and Quinsey, V.L. (1999) *The criminal recidivism process*. Cambridge: Cambridge University Press.

Zarkin, G.A., Mroz, T., Bray, J.W. and French, M.T. (1998) 'The relationship between substance use and labour supply for young men'. *Journal of Labour Economics*, 5, 385-409.

Appendix 1: Example of a Quality Assessment Tool and guidance notes on completion

Quality Assessment Tool

Quality indicator	Level of quality	Grade
1. Sample		
a) size	Whole population or 100+ participants in both treatment and control groups	1
	70% of population or 50–100 participants in both treatment and control groups	2
	Less than 50 participants in both treatment and control groups	3
	Not reported	5
b) method	Whole population or random samples	1
	Purposive samples with potential impact adequately controlled for statistically	2
	Purposive samples with potential impact not adequately controlled for statistically, or not controlled for at all	3
	Not reported	5
2. Bias		
a) response/refusal bias	No bias	1
	Some bias but adequately controlled for statistically	2
	Some bias and not adequately controlled for statistically, or not controlled for at all	3
	Not reported	5
b) attrition bias	No/very little (< 10%) attrition	1
	Some attrition but adequately controlled for statistically	2
	Some attrition and not adequately controlled for statistically, or not controlled for at all	3
	Not reported	5
3. Data collection		
a) method	Very appropriate	1
	Appropriate	2
	Not appropriate	3
	Not reported	5
b) timing	Very appropriate	1
	Appropriate	2
	Not appropriate	3
	Not reported	5
c) validation	Very appropriate	1
	Appropriate	2
	Not appropriate	3
	Not reported	5
4. Data analysis		
a) appropriate techniques/ reporting	Very appropriate	1
	Appropriate	2
	Not appropriate	3
	Not reported	5

Guidance notes for completion of Quality Assessment Tool

2a) Response/refusal bias

This score relates to any bias that may have been **introduced once the samples had been selected**. Two examples of potential response/refusal bias:

If a study relied on voluntary take-up of treatment/intervention once the treatment sample had been selected, were those that volunteered to participate comparable to all those chosen to participate in the treatment group?

If a study relied on self-reported data among treatment and control groups (once those groups had been selected), were those in the treatment and control groups who completed the self-report questionnaire/interview comparable to the total populations of the treatment and control groups?

2b) Attrition bias

Were all the participants in the experimental and the control samples accounted for? Were there differences between the study participants (in both treatment and control groups) at the pre- and post-test stages? Were there more “lost-to-follow-ups” in the treatment group compared to the control group (or vice versa)? Is attrition evident but no adequate discussion found in the study, or is it discussed but not controlled for adequately?

3a) Method of data collection

What data collection methods were employed, e.g. self-completion questionnaire, structured interview, analysis of administrative data (crime records)? Were these appropriate in terms of supplying the required data to be able to answer the research question(s) posed?

Studies that rely on the retrospective collection of self-reported pre- and post-intervention data only should be given a maximum score of 2 (given likely recall issues). Studies relying on a single data collection method should be given a maximum score of 2.

3b) Timing of data collection

Was the timing of data collection from the control and comparison groups before and after the treatment appropriate? Was a sufficient length of time left after treatment when collecting recidivism data to adequately determine outcome in terms of reduced offending?

24+ month follow-ups should be rated as 1, 12–24 month follow-ups should be rated as 2 and under-12 month follow-ups should be rated as 3. Those studies where no baseline data are collected should be marked as 3.

For longitudinal studies, were the data collected at appropriate intervals? Was a rationale given for the timing of the data collection, and was it appropriate?

3c) Validation of data

If appropriate, were different sources of data used? Was any triangulation carried out? For example, was self-reported criminality matched to official records?

Studies relying on a single data source should be given a maximum score of 2. Studies that rely on a single measure of recidivism should be given a maximum score of 2.

Data collection – general

Where multiple methods are used, the reviewer must make a judgement regarding the overall standard of the data collection, concentrating on those data deemed most appropriate to answering the research questions.

4a) Appropriate statistics and techniques used

Were appropriate statistics used (e.g. Chi-square, t-test, ANOVA, regression) and reported? Were standard deviations reported as well as differences of means? Were lower and upper quartiles reported (or the range) as well as medians? Were confidence intervals reported as well as odds ratio? Were significance levels reported?

Were repeated measures reported, i.e. were baseline data **and** post-treatment data reported? If post-treatment data only are reported, the maximum score given should be 2.

Appendix 2: OASys scored items

The scored questions within the core OASys assessment are listed below. Unless indicated otherwise, the questions are scored: 0 = **no problems**; 1 = **some problems**; and 2 = **significant problems**.

1 & 2	Offending information
1.3	Total number of separate offences for which convicted at this court (excluding TICS) [0='1'; 1= '2-3'; 2='4+']
1.4	Any current or previous convictions for burglary? [0= 'No'; 1= 'Yes']
1.5	Number of court appearances at which convicted aged under 18 years [0= '0'; 1= '1-2'; 2= '3+']
1.6	Number of court appearances at which convicted aged 18 and over, excluding current appearance [0= '0'; 1= '1-2'; 2= '3+']
1.7	Age at first conviction (record in years) [0= '18+'; 1= '14-17'; 2 = 'under 14']
1.8	Age first in contact with police: first recorded caution, reprimand or final warning (record in years) [0= '18+'; 1= '14-17'; 2= 'under 14']
1.9	Number of previous custodial sentences aged under 21 years [0= '0'; 1= '1-2'; 2= '3+']
1.10	Number of previous custodial sentences aged 21 years or over [0= '0'; 1= '1-2'; 2= '3+']
1.11	Any breaches of probation/parole/licence/bail or community based sentence [0= 'No'; 2= 'Yes']
1.12	Number of different categories of conviction (include previous and current) [0= '0-2'; 1= '3-4'; 2= '5+']
2.14	Are current offences part of an established pattern of similar offending? [0= 'No'; 2= 'Yes']
3	Accommodation
3.3	Currently of no fixed abode or in transient accommodation [0= 'No'; 2= 'Yes']
3.4	Suitability of accommodation
3.5	Permanence of accommodation
3.6	Suitability of location of accommodation [NB: if no fixed abode/transient, automatically score 2 on all four accommodation questions]
4	Employability
4.2	Is the person unemployed, or will be unemployed on release [0= 'No'; 2= 'Yes']
4.3	Employment history
4.4	Work-related skills [2= 'No skills']
4.5	Attitude to employment
4.6	School attendance
4.7	Has problems with reading/writing/numeracy
4.8	Has learning difficulties
4.9	Any educational or formal professional/vocational qualifications [0= 'Any qualification'; 2= 'No qualifications']
4.10	Attitude to education

5	Financial management and income
5.2	What is the offender's financial situation?
5.3	Financial management
5.4	Illegal earnings are a source of income
5.5	Over reliance on family/friends/others for financial support
5.6	Severe impediment to budgeting
6	Relationships
6.1	Current relationship with close family members
6.2	Close family member has criminal record [0= 'No'; 2= 'Yes']
6.3	Experience of childhood
6.4	Current relationship with partner
6.5	Current partner has criminal record [0= 'No'; 2= 'Yes']
6.6	Previous experience of close relationships
7	Lifestyle and associates
7.1	Community integration
7.2	Regular activities encourage offending
7.3	Easily influenced by criminal associates
7.4	Manipulative/predatory lifestyle
7.5	Reckless and risk-taking behaviour
8	Drug misuse
8.4	Current drug noted at 8.1, or drug noted in 8.2 and 8.3 [Score 2 if heroin, methadone, other opiates, cocaine (whether crack or powdered) or misused prescribed drugs; score 0 otherwise.]
8.5	Level of use of main drug [2= 'Weekly or more often'; 0= 'less frequently than weekly']
8.6	Ever injected drugs [0= 'Never'; 1= 'previously'; 2= 'currently']
8.7	Violent behaviour related to drug use [0= 'No'; 2= 'Yes']
8.8	Motivation to tackle drug misuse
8.9	Drug use and obtaining drugs a major activity/occupation [NB: All items are completed only for offenders who have ever used drugs, otherwise automatically score all items as 0]
9	Alcohol misuse
9.1	Is current use a problem?
9.2	Binge drinking or excessive use of alcohol in last six months
9.3	Frequency and level of alcohol misuse in the past
9.4	Violent behaviour related to alcohol use at any time [0= 'No'; 2= 'Yes']
9.5	Motivation to tackle alcohol misuse (if applicable) [NB: 9.4 and 9.5 are only completed if the sum of 9.1, 9.2 and 9.3 is above zero, otherwise score both these items as 0]
10	Emotional wellbeing
10.1	Difficulties coping
10.2	Current psychological problems/depression
10.3	Social isolation
10.4	Offender's attitude to themselves
10.5	Self-harm, attempted suicide, suicidal thoughts or feelings [0= 'No'; 2= 'Yes']
10.6	Current psychiatric problems

11	Thinking and behaviour
11.1	Level of interpersonal skills
11.2	Impulsivity
11.3	Aggressive/controlling behaviour
11.4	Temper control
11.5	Ability to recognise problems
11.6	Problem-solving skills
11.7	Awareness of consequences
11.8	Achieves goals
11.9	Understands other people's views
11.10	Concrete/abstract thinking
12	Attitudes
12.1	Pro-criminal attitudes
12.3	Attitude towards staff
12.4	Attitude towards supervision/licence
12.5	Attitude to community/society
12.6	Does the offender understand their motivation for offending?
12.8	Is the offender motivated to address offending?

Appendix 3: Underlying factors of current scored questions

Factor (Variance explained)	Question	Loading
1. Emotional wellbeing (5.3%)	6.3: Experience of childhood	.356
	10.1: Difficulties coping	.765
	10.2: Current psychological problems/depression	.810
	10.3: Social isolation	.615
	10.4: Offender's attitude to themselves	.689
	10.5: Self-harm, attempted suicide, suicidal thoughts or feelings	.640
	10.6: Current psychiatric problems	.715
2. Alcohol misuse (5.1%)	9.1: Is current use a problem?	.875
	9.2: Binge drinking or excessive use of alcohol in last six months	.874
	9.3: Frequency and level of alcohol misuse in the past	.838
	9.4: Violent behaviour related to alcohol use at any time	.694
	9.5: Motivation to tackle alcohol misuse	.751
3. Adult/ established offending (5.1%)	1.4: Any current or previous convictions for burglary?	.547
	1.6: Number of court appearances at which convicted aged 18 years and over	.803
	1.10: Number of previous custodial sentences aged 21 years and over	.726
	1.11: Any breaches of probation/parole/licence/bail or community-based sentence	.629
	1.12: Number of different categories of conviction	.689
	2.14: Are current offence(s) part of an established pattern of similar offending?	.513
4. Drug misuse (5.0%)	5.4: Illegal earnings are a source of income	.413
	8.4: Current drug noted	.684
	8.5: Level of use of main drug	.764
	8.6: Ever injected drugs	.537
	8.8: Motivation to tackle drug misuse	.681
	8.9: Drug use and obtaining drugs a major activity/ occupation	.782
5. Thinking and behaviour (4.9)	11.2: Impulsivity	.546
	11.5: Ability to recognise problems	.738
	11.6: Problem-solving skills	.741
	11.7: Awareness of consequences	.770
6. Accommodation (4.8%)	11.9: Understands other people's views	.483
	3.3: Currently of no fixed abode or in transient accommodation	.919
	3.4: Suitability of accommodation	.888
	3.5: Permanence of accommodation	.874
	3.6: Suitability of location of accommodation	.848

Factor (Variance explained)	Question	Loading
7. Youth/initial offending (4.8%)	1.5: Number of court appearances at which convicted aged under 18 years	.851
	1.7: Age at first conviction	.895
	1.8: Age first in contact with police	.871
	1.9: Number of previous custodial sentences aged under 21 years	.522
8. Attitudes (4.6%)	11.1: Level of interpersonal skills	.346
	11.10: Concrete/abstract thinking	.465
	12.1: Pro-criminal attitudes	.485
	12.3: Attitude towards staff	.676
	12.4: Attitude towards supervision/licence	.607
	12.5: Attitude to community/society	.612
	12.6: Does the offender understand their motivation for offending?	.470
	12.8: Motivation	.530
9. Employment (4.6%)	4.2: Is the person unemployed, or will be unemployed on release?	.600
	4.3: Employment history	.715
	4.4: Work-related skills	.724
	4.5: Attitude to employment	.671
	4.10: Attitude to education/training	.461
	7.1: Community integration	.386
	11.8: Achieves goals	.427
10. Financial management and income (3.8%)	5.2: What is the offender's financial situation?	.787
	5.3: Financial management	.793
	5.5: Over reliance on family/friends/others for financial support	.599
	5.6: Severe impediment to budgeting	.689
11. Education (3.2%)	4.6: School attendance	.498
	4.7: Has problems with reading, writing and/or numeracy	.785
	4.8: Has learning difficulties	.696
	4.9: Any educational or formal professional/ vocational qualifications?	.578
12. Violence (2.7%)	8.7: Violent behaviour related to drug use	.515
	11.3: Aggressive/controlling behaviour	.747
	11.4: Temper control	.729
13. Relationships (2.5%)	6.1: Current relationship with close family members	.413
	6.4: Current relationship with partner	.631
	6.6: Previous experience of close relationships	.536
	7.4: Manipulative/predatory lifestyle	.421
14. Lifestyle and associates (2.3%)	1.3: Total number of separate offences for which convicted at this court appearance	.458
	7.2: Regular activities encourage offending	.409
	7.3: Easily influenced by criminal associates	.389
	7.5: Recklessness and risk-taking behaviour	.521
15. Family offending (1.8%)	6.2: Close family member has criminal record	.580
	6.5: Current partner has criminal record	.588

Key: Non-matching OASys section/factor.

Appendix 4: Revised underlying factors

Factor (Variance explained)	Question	Loading
1. Alcohol misuse (7.5%)	9.1: Is current use a problem?	.868
	9.2: Binge drinking or excessive use of alcohol in last six months	.871
	9.3: Frequency and level of alcohol misuse in the past	.855
	9.4: Violent behaviour related to alcohol use at any time	.752
	9.5: Motivation to tackle alcohol misuse	.746
2. Emotional wellbeing (7.4%)	10.1: Difficulties coping	.760
	10.2: Current psychological problems/depression	.835
	10.3: Social isolation	.607
	10.4: Offender's attitude to themselves	.677
	10.5: Self-harm, attempted suicide, suicidal thoughts or feelings	.625
	10.6: Current psychiatric problems	.750
3. ETE (7.4%)	4.3: Employment history	.665
	4.4: Work-related skills	.750
	4.5: Attitude to employment	.622
	4.6: School attendance	.648
	4.9: Any educational or formal professional/ vocational qualifications?	.673
	4.10: Attitude to education/training	.676
4. Accommodation (7.3%)	3.3: Currently of no fixed abode or in transient accommodation	.921
	3.4: Suitability of accommodation	.892
	3.5: Permanence of accommodation	.875
	3.6: Suitability of location of accommodation	.854
5. Drug misuse (7.0%)	8.4: Current drug noted	.747
	8.5: Level of use of main drug	.764
	8.6: Ever injected drugs	.642
	8.8: Motivation to tackle drug misuse	.653
	8.9: Drug use and obtaining drugs a major activity/ occupation	.792
6. Thinking and behaviour (6.7%)	11.5: Ability to recognise problems	.774
	11.6: Problem-solving skills	.738
	11.7: Awareness of consequences	.771
	11.8: Achieves goals	.417
	11.9: Understands other people's views	.589
	11.10: Concrete/abstract thinking	.548
7. Attitudes (5.9%)	12.1: Pro-criminal attitudes	.524
	12.3: Attitude towards staff	.720
	12.4: Attitude towards supervision/licence	.638
	12.5: Attitude to community/society	.640
	12.8: Motivation	.486
8. Financial management and income (5.6%)	5.2: What is the offender's financial situation?	.798
	5.3: Financial management	.798
	5.5: Over reliance on family/friends/others for financial support	.609
	5.6: Severe impediment to budgeting	.690
9. Relationships (3.6%)	6.1: Current relationship with close family members	.589
	6.3: Experience of childhood	.696
	6.6: Previous experience of close relationships	.637
10. Lifestyle and associates (3.6%)	7.2: Regular activities encourage offending	.489
	7.3: Easily influenced by criminal associates	.636
	7.5: Recklessness and risk-taking behaviour	.672

Appendix 5: Scored OASys questions included within revised risk factors, criminogenic need scales, OGP and/or OVP (OASys sections 3 to 12)

Scored OASys question	Revised risk factors	Criminogenic need scales	OGP	OVP
3.3: Currently of no fixed abode or in transient accommodation	✓	✓	✓	✓
3.4: Suitability of accommodation	✓	✓	✓	✓
3.5: Permanence of accommodation	✓	✓	✓	✓
3.6: Suitability of location of accommodation	✓	✓	✓	✓
4.2: Is the person unemployed, or will be unemployed on release			✓	✓
4.3: Employment history	✓	✓	✓	✓
4.4: Work-related skills	✓	✓	✓	✓
4.5: Attitude to employment	✓	✓	✓	✓
4.6: School attendance	✓	✓		
4.7: Has problems with reading, writing and/or numeracy				
4.8: Has learning difficulties				
4.9: Any educational or formal professional/vocational qualifications	✓	✓		
4.10: Attitude to education/training	✓	✓		
5.2: What is the offender's financial situation?	✓	✓		
5.3: Financial management	✓	✓		
5.4: Illegal earnings are a source of income				
5.5: Over reliance on family/friends/others for financial support	✓	✓		
5.6: Severe impediment to budgeting	✓	✓		
6.1: Current relationship with close family members	✓			
6.2: Close family member has criminal record				
6.3: Experience of childhood	✓			
6.4: Current relationship with partner				
6.5: Current partner has criminal record				
6.6: Previous experience of close relationships	✓			
7.1: Community integration				
7.2: Regular activities encourage offending	✓		✓	✓
7.3: Easily influenced by criminal associates	✓			

7.4:	Manipulative/predatory lifestyle				
Scored OASys question		Revised risk factors	Criminogenic need scales	OGP	OVP
7.5:	Recklessness and risk-taking behaviour	✓			
8.4:	Current drug noted	✓	✓	✓	
8.5:	Level of use of main drug	✓	✓	✓	
8.6:	Ever injected drugs	✓	✓	✓	
8.7:	Violent behaviour related to drug use				
8.8:	Motivation to tackle drug misuse	✓	✓	✓	
8.9:	Drug use and obtaining drugs a major activity/occupation	✓	✓	✓	
9.1:	Is current use a problem	✓	✓		✓
9.2:	Binge drinking or excessive use of alcohol in last six months	✓	✓		✓
9.3:	Frequency and level of alcohol misuse in the past	✓	✓		✓
9.4:	Violent behaviour related to alcohol use at any time	✓	✓		✓
9.5:	Motivation to tackle alcohol misuse	✓	✓		✓
10.1:	Difficulties coping	✓			
10.2:	Current psychological problems/ depression	✓			
10.3:	Social isolation	✓			
10.4:	Offender's attitude to themselves	✓			
10.5:	Self-harm, attempted suicide, suicidal thoughts or feelings	✓			
10.6:	Current psychiatric problems	✓			
11.1:	Level of interpersonal skills			✓	✓
11.2:	Impulsivity			✓	✓
11.3:	Aggressive/controlling behaviour				✓
11.4:	Temper control				✓
11.5:	Ability to recognise problems	✓	✓	✓	✓
11.6:	Problem-solving skills	✓	✓	✓	✓
11.7:	Awareness of consequences	✓	✓	✓	✓
11.8:	Achieves goals	✓	✓	✓	✓
11.9:	Understands other people's views	✓	✓	✓	✓
11.10:	Concrete/abstract thinking	✓	✓	✓	✓
12.1:	Pro-criminal attitudes	✓	✓	✓	✓
12.3:	Attitude towards staff	✓	✓	✓	✓
12.4:	Attitude towards supervision/licence	✓	✓	✓	✓
12.5:	Attitude to community/society	✓	✓	✓	✓
12.6:	Does the offender understand their motivation for offending?			✓	✓
12.8:	Motivation	✓	✓	✓	✓

Appendix 6: Logistic regression results and 100-point scales for OGP and OVP

This appendix reports the results of logistic regression models of proven reoffending within 24 months of community sentence or discharge from custody, and the process which transformed logistic regression parameters into user-friendly 100-point scales. Tables A6.1 and A6.3 give the logistic regression parameters for OGP and OVP respectively, while Tables A6.2 and A6.4 illustrate the scaling processes. Table A6.5 provides assurance that the transformation process had little effect on predictive validity.

Creating the 100-point scales

To transform the logistic regression parameters into scores out of 100, the minimum and maximum possible scores based on the logistic regression results were calculated. The range between the minimum and maximum was divided into 100, and the range on each risk factor expressed as hundredths of this overall range. For example, on OVP the overall range of logistic regression parameters was 8.33 (from -3.19 to 5.14) and the range for accommodation was 0.17 so the number of points available for accommodation was $100 \times 0.17 / 8.33 = 2$ points. Rounding errors meant that some tweaking of these scores was necessary in order to obtain a total of 100.

These scores were then amended to make them easier for practitioners to understand and to encourage dynamic risk assessment. The static/dynamic balance of OGP was shifted from 59/41 to 60/40 for ease of calculation, and the scoring of OVP was made more dynamic in order to match that of OGP and give more scope for scores to change over time. The maximum scores for each risk factor were adjusted so that they were multiples of five (OGP, and OVP static factors) or two (OVP dynamic factors), to make the scores more user-friendly.

The effects of these amendments on predictive validity were checked, and are displayed in Table A6.5. Compared with the raw results in Tables A6.1 and A6.3, the revisions actually increased the AUCs of both predictors by almost a percentage point, though using the 'original' weights for OVP would give a slightly better AUC than the more dynamic 'revised' weights. It is unclear why these improvements should occur, but it is reasonable to assume that the relative validity of two strongly correlated predictors (i.e. the raw results and weighted scores) will fluctuate randomly.

Transforming the 100-point scales to one-year and two-year predicted rates

Following the transformation of the model results to 100-point scores, further logistic regression models were run to fit the 100-point scores to one-year and two-year reoffending outcomes. The models are as follows:

For all models, the quoted % probability of proven reoffending = $e^z/(1 + e^z)$, rounded down, where e is the exponential constant, approximately 2.718.

$$z = \alpha + \beta$$

For OGP, $\alpha = -3.683$ (1-year prediction) or -3.080 (2-year prediction)
 $\beta = 0.0616 * 100\text{-point score}$

For OVP, $\alpha = -4.522$ (1-year prediction) or -3.877 (2-year prediction)
 $\beta = 0.0622 * 100\text{-point score}$

Table A6.1: Ordinal logistic regression model of proven reoffending within 12 or 24 months: results for OGP offences

OASys question/ other covariate	Parameter estimate	Standard error of estimate	Odds ratio (for scored items, 1 pt. then range)
Constant (12 months)	-3.61	0.057	-
Constant (24 months)	-3.02	0.055	-
OGRS 3 score (range 0–100)	0.037	0.001	1.037/37.83
3.3 to 3.6: Score on four accommodation questions (range 0–8)	0.039	0.007	1.039/1.36
4.2–4.5: Score on four employment questions (range 0–8)	0.038	0.009	1.038/ 1.35
7.2: Regular activities encourage offending (range 0–2)	0.109	0.029	1.12/1.25
8.4–8.6, 8.8, 8.9: Score on five drug misuse questions (range 0–10)	0.115	0.008	1.12/3.16
11.1, 11.2, 11.5–11.10: Score on eight thinking and behaviour questions (range 0–20)	0.017	0.006	1.017/1.31
12.1, 12.3–12.8: Score on six attitudes questions (range 0–12)	0.024	0.011	1.025/1.34

Table A6.2: Scaling OGP logistic regression results to produce a 100-point score

Risk factor	Raw regression parameters			Scores/100	
	Minimum	Maximum	Range	Original (Range* 100/6.23)	Revised
OGRS 3	0	3.67	3.67	59	60
All static factors	0	3.67	3.67	59	60
3.3–3.6 Accommodation	0	0.31	0.31	5	5
4.2–4.5 Employability	0	0.30	0.30	5	5
7.2 Regular activities	0	0.28	0.28	4	5
8.4–8.6, 8.8, 8.9 Drug misuse	0	1.15	1.15	18	15
11.1–11.10: Thinking and behaviour	0	0.27	0.27	4	5
12.1, 12.3–12.8 Attitudes	0	0.29	0.29	5	5
All dynamic factors	0	2.60	2.60	41	40
Total	0	6.23	6.23	100	100

Table A6.3: Ordinal logistic regression model of proven reoffending within 12 or 24 months: results for OVP offences

OASys question/ other covariate	Parameter estimate	Standard error of estimate	Odds ratio (for scored items, 1 pt. then range)
Constant (12 months)	-2.00	0.096	-
Constant (24 months)	-1.35	0.096	
Number of sanctioning occasions for violent-type offences			
None	-1.09	0.08	0.09
1	-0.79	0.07	0.12
2	-0.53	0.07	0.15
3	-0.39	0.08	0.17
4	-0.25	0.08	0.20
5	-0.22	0.10	0.21
6	-0.07	0.11	0.24
7 or 8	0.12	0.11	0.29
9 or 10	0.40	0.15	0.38
11 or 12	0.68	0.22	0.51
13 to 17	0.77	0.26	0.55
18 or more (reference category)	1.37	-	1
Number of sanctioning occasions for other offences			
None, 1 or 2	-0.37	0.04	0.49
3 or 4	-0.14	0.04	0.62
5 to 10	-0.05	0.04	0.68
11 to 20	0.20	0.04	0.87
21 or more (reference category)	0.36	-	1
Is this the offender's first sanction ever?			
Yes	0.32	0.04	1.88
No	-0.32	-	-
Age at date of assessment, grouped			
18–19	0.96	0.05	6.91
20–21	0.68	0.05	5.13
22–23	0.45	0.06	4.08
24–25	0.40	0.06	3.89
26–30	0.19	0.05	3.16
31–35	-0.13	0.05	2.29
36–40	-0.37	0.06	1.80
41–45	-0.58	0.08	1.47
46–50	-0.65	0.12	1.37
51+ (reference category)	-0.96	-	1
Sex			
Female	-0.22	0.04	0.64
Male (reference category)	0.22	-	1
2.6: Recognises impact of offending?			
No (reference category)	0.08	-	1
3.3 to 3.6: Score on four accommodation questions (range 0–8)	0.021	0.008	1.021/1.18
4.2 to 4.5: Score on four employment questions (range 0–8)	0.044	0.009	1.045/1.42
9.1, 9.2: Current alcohol misuse and binge drinking (range 0–4)	0.148	0.013	1.16/1.80

Table A6.3: Ordinal logistic regression model of proven reoffending within 12 or 24 months: results for OVP offences (continued)

OASys question/ other covariate	Parameter estimate	Standard error of estimate	Odds ratio (for scored items, 1 pt. then range)
10.7: Current psychiatric treatment, or treatment pending			
Yes	0.11	0.04	1.25
No (reference category)	-0.11	--	1
11.4: Temper control (range 0–2)	0.193	0.03	1.21/1.46
12.1, 12.3-12.8: Score on six attitudes questions (range 0–12)	0.021	0.010	1.022/1.30

Table A6.4: Scaling OVP logistic regression results to produce a 100-point score

Risk factor	Raw regression parameters			Scores/100	
	Minimum	Maximum	Range	Original (Range* 100/7.54)	Revised
Violent-type sanctions	-1.09	1.36	2.45	29	25
Other sanctions	-0.36	0.36	0.72	9	5
First sanction ever	-0.32	0.32	0.63	8	5
Age	-0.96	0.96	1.92	23	20
Sex	-0.22	0.22	0.44	5	5
All static factors	-2.95	3.22	6.17	74	60
2.6 Recognises impact of offending	-0.08	0.08	0.16	2	4
3.3–3.6 Accommodation	0	0.17	0.17	2	4
4.2-4.5 Employability	0	0.36	0.36	4	6
9.1-9.5: Alcohol misuse	0	0.60	0.60	7	10
10.7 Psychiatric treatment	-0.11	0.11	0.22	3	4
11.4 Temper control	0	0.40	0.40	5	6
12.1, 12.3–12.8 Attitudes	0	0.25	0.25	3	4
All dynamic factors	-0.19	1.97	2.16	26	40
Total	-3.14	5.19	8.33	100	100

Table A6.5: AUC scores for validation sample, comparing raw regression, original and revised score models

Model name	Raw logistic regression parameters	Original score/100	Revised score/100
OGP	0.792	0.799	0.800
OVP	0.733	0.745	0.742

Appendix 7: Distributions and correlations of existing and new predictors

Table A7.1 presents descriptive statistics for OGRS scores, each section of OASys, and total OASys scores. Table A7.1 lists the correlations between each measure.

Table A7.2 shows that OASys sections differ in the ratio of the maximum and average scores. In the ETE and thinking and behaviour sections, the average score is about 30% of the maximum, and in several other sections the average is about 25% of the maximum. The drug misuse section is very different to this – the average of 1.70 is just 14% of the maximum of 12.

These differences feed through to the dynamic portion of the new predictors. In OGP, which weights drug misuse heavily, the average real score is only ten of the maximum of 40 dynamic points (25%). OVP includes the higher-averaging alcohol misuse rather than drug misuse, but also includes the question on psychiatric treatment where only 5% of offenders in this sample score the four available points. (However, in a 2007/08 sample, 10% of a combined prison/probation sample were currently in psychiatric treatment; the change is probably due to both the focus towards only assessing more serious offenders and improved, more thorough assessment.) The net result is that, similarly, only just over ten points are scored on dynamic items. Overall, the average offender's OGP score comprises 75% static and 25% dynamic (10.0/40.1) items, while the average offender's OVP score comprises 72% static and 28% dynamic (10.2/36.8) items.

Correlations within OASys are often weak, especially those involving the accommodation, alcohol misuse and emotional wellbeing sections, highlighting the heterogeneity of the offenders assessed. Alcohol misuse and emotional wellbeing scores are also very weakly correlated with OGRS. OGRS 2 and 3 scores are well correlated with other total scores, with the OGRS 2 predictor of sexual and violent reoffending generally having slightly weaker correlations.

The total OGP score is far more strongly correlated with existing measures than is the total OVP score. The independence of OVP from existing predictors (no correlations above 0.7) helps to explain why it is able to contribute a considerable improvement in predictive validity: it is measuring something quite different to the existing predictors. Comparison of the results for drug and alcohol misuse affirm the extent of the differences between these two forms of substance misuse.

The weak (0.23) correlation between alcohol misuse and the static part of OVP is worth noting, as it contributes to the modest size of the correlation between the total dynamic and static parts of OVP (0.35). A further examination of the component parts of the static part of OVP (not in Table A7.2) reveals that alcohol misuse has near-zero correlations with age and the non-violent sanctions item, but a comparatively strong (also 0.35) correlation with violent sanctions. The difficulty of predicting violent reoffending is made evident by the fact that some of the items which are associated with violent reoffending have little or no correlation with one another. Unless a more advanced form of statistical modelling can be proven to be effective, it is therefore inevitable that predictors of violence will – as Table 6.8 shows – fail to identify many offenders as being almost certain to reoffend violently.

Table A7.1: Descriptive statistics of risk predictors

Predictor	Number of items in predictor	Range of scores	Mean (SD)	Percentiles				
				10th	25th	50th	75th	90th
Existing OASys-based predictors								
OASys total weighted score	73	0–168	55.0 (35.9)	10	25	50	81	106
OASys sections 3–12 weighted score	62	0–18	36.5 (24.3)	7	16	33	54	72
OASys unweighted section scores								
Accommodation	4	0–8	1.62 (2.72)	0	0	0	2	8
Education, training & employability	9	0–18	5.27 (4.40)	0	2	5	8	12
Financial management & income	5	0–10	2.28 (2.50)	0	0	2	4	6
Relationships	6	0–12	2.67 (2.53)	0	0	2	4	6
Lifestyle & associates	5	0–10	2.55 (2.33)	0	1	2	4	6
Drug misuse	6	0–12	1.70 (2.78)	0	0	0	3	6
Alcohol misuse	5	0–10	2.67 (3.12)	0	0	1	5	8
Emotional wellbeing	6	0–12	2.47 (2.99)	0	0	1	4	7
Thinking & behaviour	10	0–20	6.65 (4.64)	1	3	6	10	13
Attitudes	6	0–12	2.37 (2.26)	0	1	2	4	6
Other existing NOMS predictors								
OGRS 2	8	0–100	44.7 (26.6)	12	21	40	68	84
OGRS 2 (sexual/violent)	7	0–100	15.9 (15.1)	3	6	11	21	35
OGRS 3	5/7*	0–100	50.6 (24.6)	14	29	53	72	82
New OASys-based predictors								
OASys General reoffending Predictor (OGP)	OGRS 3 + 6/28**	0–100	40.1 (20.4)	13	23	40	56	68
OGP dynamic component	6/28	0–40	10.0 (5.1)	1	3	8	15	22
OGP dynamic as a proportion of OGP total	n/a	0%–100%	24% (14%)	6	13	23	32	41
OASys Violence Predictor (OVP)	12/23	0–100	36.8 (13.1)	20	28	37	46	54
OVP static component	5/4	0–60	26.5 (8.9)	14	21	27	33	37
OVP dynamic component	7/19	0–40	10.2 (7.1)	2	5	9	15	20
OVP dynamic as a proportion of OVP total	n/a	0%–100%	26% (15%)	7	16	26	36	45

* OGRS 3 has four terms, plus a constant, in its logistic regression algorithm. However, two of these terms (age-by-sex, and modified “Copas rate” [the tempo of offending]) require multiple pieces of information, so in total six data items are required.

** The dynamic part of OGP includes 28 questions in six groups. The static part of OVP includes five scores derived from four individual questions, while the dynamic part includes 19 questions in seven groups. Allowing for overlap between the two dynamic parts, together they require 33 questions in ten groups.

Table A7.2: Intercorrelations of new and existing measures of recidivism risk

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. OASys total weighted score	---	.92	.47	.72	.59	.52	.76	.61	.23	.35	.67	.64	.72	.59	.68	.84	.86	.65	.47	.63
2. OASys sections 3–12 weighted score		---	.55	.76	.67	.57	.81	.62	.26	.46	.73	.67	.54	.45	.54	.76	.91	.62	.36	.70
OASys unweighted section scores																				
3. Accommodation			---	.29	.33	.35	.35	.30	.10	.24	.27	.26	.22	.19	.22	.37	.56	.31	.12	.42
4. Education, training & employability				---	.46	.39	.54	.38	.11	.27	.50	.43	.49	.40	.48	.61	.65	.52	.35	.53
5. Financial management & income					---	.31	.52	.48	.03	.27	.35	.33	.34	.22	.33	.47	.60	.25	.12	.32
6. Relationships						---	.40	.27	.19	.46	.42	.35	.21	.22	.23	.35	.46	.33	.12	.46
7. Lifestyle & associates							---	.49	.18	.28	.59	.56	.50	.40	.48	.68	.77	.50	.33	.51
8. Drug misuse								---	-.06	.19	.25	.29	.45	.30	.43	.63	.81	.23	.18	.21
9. Alcohol misuse									---	.23	.33	.19	.05	.16	.09	.12	.12	.53	.23	.70
10. Emotional wellbeing										---	.35	.22	.03	.05	.06	.18	.32	.20	-.04	.43
11. Thinking & behaviour											---	.68	.36	.38	.38	.51	.60	.62	.37	.69
12. Attitudes												---	.37	.37	.36	.51	.60	.51	.30	.57
13. OGRS 2													---	.74	.92	.88	.57	.60	.67	.29
14. OGRS 2 (sexual/violent offending)														---	.67	.63	.41	.63	.64	.36
15. OGRS 3															---	.94	.56	.66	.73	.32
16. OASys General reoffending Predictor (OGP)																---	.80	.68	.66	.45
17. OGP dynamic component																	---	.52	.32	.56
18. OASys Violence Predictor (OVP)																		---	.86	.78
19. OVP static component																			---	.35
20. OVP dynamic component																				---

Appendix 8: Diversity issues: accuracy and validity of prediction by age, sex and ethnicity

This appendix presents statistics on the applicability of OGP, OVP and other predictors to offenders of varying age, sex and ethnicity.

Design issues

Age and gender

OGRS 3 combines age and gender into a single set of parameters, recognising that the reoffending rates of women are lower but peak at a later age. OGP utilises the OGRS 3 score, so also has this combined age-gender term. OVP does not attempt to combine the two, given the statistical difficulties associated with low base rates of proven violent offending, and the desire to maintain a simple scoring system, but does include age and gender as separate terms. Our results compare OVP age categories.

Of the existing predictors, OGRS 2 includes separate age and gender terms, while the OASys score makes no allowance for either.

Ethnicity

Ethnicity is not a component in offenders' scores – that is, predicted probabilities of proven reoffending – in any of the predictors. The tables below will examine whether these predictions are accurate for each ethnic group. Given the size of the sample, it is only feasible to classify ethnic groups broadly (e.g. Asian) rather than exactly (e.g. Indian, Pakistani, Bangladeshi). The validation sample includes a smaller proportion of non-White offenders than the 2007/08 caseload as the metropolitan probation areas, where many non-White offenders live, were slow to use the electronic version of OASys and are therefore underrepresented.

Methods

There are two possible approaches to empirical comparison of predictive validity between subgroups. One is to use Area Under Curve (AUC) statistics – checking that higher predictor scores represent higher likelihood of proven reoffending – and the other is to compare predicted and actual reoffending rates. Both methods are useful if their limitations are understood, and are best used together: AUCs check whether offenders are ranked correctly, whereas comparisons of predicted and actual rates check whether predictions are, on average, too high or low.

Of all the predictors, only OGP and OVP produce exact percentage predictions for non-violent and violent-type offences. Predicted probabilities of proven reoffending for the other tools are generated by scaling the predictor scores to fit the reoffending rate of interest.⁹⁵ This

⁹⁵ Logistic regression is used to fit the predictor scores to the event of interest, with simple, quadratic and cubic terms fitted to allow the shape of the distribution to work as well as possible; this is an allowance for the use of the predictors to predict outcomes they were not specifically designed for. It does not affect AUC statistics but may improve residuals.

method produces zero overall residuals for all these predictors. For OGP and OVP, variability between the construction and validation samples has an effect, and there are overall residuals of +0.7% for OGP (i.e. actual rates were higher than predicted) and +0.3% for OVP.

Note that AUC scores will naturally be lower for most subgroups than overall. This is because the subgroups are more homogeneous on many risk factors, and therefore the predictors are attempting to distinguish between offenders who are relatively similar. The small size of most ethnic groups means that their residuals are likely to be relatively large, due to the effects of chance variation.

Results

Table A8.1 gives results for the prediction of those offences covered by OGP. As an example of how to read the table, the residual of +2.4% for 18 to 19-year-olds on OGP meant that OGP predicted a mean reoffending rate of $50.0 - 2.4 = 47.6\%$, where 50.0% is the actual reoffending rate. When only comparing the predictions of all 18 to 19-year-olds, OGP attained an AUC of 0.75.

Table A8.1: Predictive validity for offences other than violence against the person, threats/harassment, weapon possession, violent acquisitive, criminal damage, and public order, by age group, gender and ethnicity

Offender characteristic (number of assessments)	Actual 24-month proven reoffending rate	Residual (actual minus predicted rate)				AUC statistics			
		OGRS 2	OGRS 3	OASys score	OGP	OGRS 2	OGRS 3	OASys score	OGP
Age group									
18–19 (1,282)	50.0%	-0.8%	-0.4%	+10.1%	+2.4%	.73	.73	.74	.75
20–21 (1,222)	44.4%	-3.1%	-3.2%	+5.5%	-0.6%	.77	.78	.77	.80
22–23 (1,107)	44.5%	-0.5%	-0.1%	+4.4%	+1.1%	.74	.73	.73	.76
24–25 (968)	44.0%	-1.6%	-1.6%	+2.3%	-0.7%	.77	.78	.76	.80
26–30 (1,825)	43.9%	+1.4%	+0.4%	+1.5%	+0.5%	.78	.78	.77	.80
31–35 (1,556)	40.7%	+3.4%	+2.7%	-1.5%	+2.3%	.77	.77	.77	.80
36–40 (1,176)	32.2%	+1.3%	+1.0%	-7.1%	+0.2%	.76	.80	.77	.81
41–45 (717)	26.5%	+0.2%	+0.5%	-10.0%	+0.1%	.77	.79	.77	.82
46–50 (397)	20.9%	-2.3%	-0.9%	-12.4%	-0.6%	.76	.77	.74	.78
51+ (426)	12.0%	-4.6%	+0.3%	-15.5%	-0.7%	.78	.80	.73	.80
Gender									
Female (1,585)	36.5%	+5.1%	+0.3%	-1.1%	-2.1%	.79	.81	.81	.84
Male (9091)	40.3%	-0.9%	--	+1.0%	+1.1%	.78	.78	.75	.80
Ethnicity									
Asian (286)	37.1%	+5.2%	+7.0%	+7.6%	+8.7%	.73	.75	.70	.77
Black (318)	45.6%	+9.2%	+9.8%	+11.5%	+11.4%	.78	.78	.74	.80
Mixed (180)	48.3%	+3.9%	+3.3%	+6.3%	+4.1%	.75	.76	.75	.78
Other (78)	41.0%	+10.2%	+9.9%	+10.8%	+11.6%	.78	.80	.75	.81
White (9,008)	40.1%	-0.3%	-0.4%	-0.5%	+0.2%	.78	.79	.77	.81
Missing/not stated (806)	31.6%	-3.5%	-3.4%	-3.9%	-2.4%	.75	.77	.72	.77

Table A8.2: Predictive validity for violence against the person, threats/ harassment, weapon possession, violent acquisitive, criminal damage and public order offences, by age group, gender and ethnicity

Offender characteristic (number of assessments)	Actual 24-month proven reoffending rate	Residual (actual minus predicted rate)				AUC statistics			
		OGRS 2	OGRS 3	OASys score	OGP	OGRS 2	OGRS 3	OASys score	OGP
Age group									
18–19 (1282)	41.9%	+9.0%	+9.2%	+15.3%	+2.9%	.65	.66	.66	.71
20–21 (1222)	33.0%	+1.8%	+1.8%	+6.9%	-0.4%	.64	.66	.67	.70
22–23 (1107)	29.6%	-0.1%	+0.1%	+2.9%	+1.1%	.64	.65	.65	.70
24–25 (968)	26.1%	-3.7%	-3.8%	-1.4%	-0.3%	.66	.68	.66	.72
26–30 (1825)	27.1%	-0.9%	-1.6%	-0.9%	+2.6%	.63	.63	.65	.71
31–35 (1556)	23.8%	-1.3%	-1.8%	-4.1%	-0.1%	.67	.68	.66	.71
36–40 (1176)	19.4%	-2.0%	-2.3%	-6.9%	-2.3%	.65	.67	.66	.72
41–45 (717)	16.9%	-1.8%	-1.5%	-7.8%	-2.2%	.63	.65	.64	.69
46–50 (397)	15.1%	-1.9%	-0.7%	-7.9%	-0.1%	.67	.68	.68	.76
51+ (426)	8.0%	-4.8%	-1.2%	-11.8%	-2.9%	.65	.65	.70	.73
Gender									
Female (1585)	17.7%	-4.1%	-6.7%	-8.1%	-0.1%	.69	.70	.72	.79
Male (9091)	28.0%	+0.7%	+1.2%	+1.4%	+0.4%	.66	.68	.65	.72
Ethnicity									
Asian (286)	22.0%	-0.6%	-1.0%	+1.1%	+2.1%	.70	.72	.63	.73
Black (318)	23.0%	-2.1%	-1.6%	-0.4%	+0.8%	.68	.68	.68	.71
Mixed (180)	32.8%	+3.6%	+3.2%	+5.1%	+5.0%	.65	.66	.67	.75
Other (78)	22.7%	+3.8%	+4.0%	+4.3%	+4.2%	.55	.64	.57	.65
White (9008)	25.6%	+0.1%	+0.1%	--	+0.3%	.67	.68	.66	.73
Missing/not stated (806)	27.0%	-1.5%	-1.2%	-1.5%	-1.8%	.65	.68	.67	.73

Appendix 9: Changes in individual scored items

Scored item	% with score change			Mean score		Mean change in score	
	Any	In-crease	De-crease	Initial	Termin-ation	Net	Abso-lute
1.3 Total current offences (R)	16.2%	9.4%	15.6%	0.70	0.69	-0.01	0.20
1.4 Any burglary ^ (R)	1.6%	2.2%	0.9%	0.81	0.83	0.02	0.03
1.5 Court appearances aged <18 (HH)	1.9%	1.6%	1.5%	0.82	0.82	0.00	0.02
1.6 Court appearances aged 18+ ^ (R)	6.5%	14.7%	0.8%	1.40	1.46	0.06	0.07
1.7 Age at first conviction (HH)	1.5%	0.9%	1.4%	0.64	0.64	0.00	0.02
1.8 Age at first police contact (HH)	1.9%	1.3%	1.5%	0.74	0.75	0.00	0.02
1.9 Previous custody aged <21 ^ (R) or (HH)	3.3%	2.7%	2.8%	0.44	0.45	0.02	0.04
1.10 Previous custody aged 21+ ^ (R)	4.3%	4.4%	1.9%	0.55	0.58	0.03	0.05
1.11 Any breach ^ (R)	4.6%	8.8%	1.3%	1.10	1.16	0.07	0.09
1.12 Offending versatility ^ (R)	6.4%	7.7%	2.3%	1.02	1.05	0.04	0.07
2.14 Pattern of similar offending (R)	6.3%	10.3%	3.6%	1.19	1.23	0.04	0.13
3.3 Currently no fixed abode	14.6%	7.9%	39.8%	0.42	0.38	-0.04	0.29
3.4 Suitability of accommodation	22.0%	14.0%	30.3%	0.61	0.59	-0.01	0.33
3.5 Permanence of accom.	21.8%	13.4%	28.5%	0.64	0.63	-0.02	0.31
3.6 Suitability of location of accom.	19.9%	12.6%	25.8%	0.67	0.65	-0.02	0.31
4.2 Employment status	17.3%	17.2%	17.4%	1.26	1.16	-0.09	0.35
4.3 Employment history (H)	7.8%	6.0%	5.3%	1.00	1.01	0.01	0.09
4.4 Work-related skills	8.1%	4.3%	7.8%	0.92	0.90	-0.02	0.09
4.5 Attitude to employment	10.7%	4.8%	14.6%	0.54	0.52	-0.02	0.12
4.6 School attendance (HH)	3.9%	3.3%	3.0%	0.79	0.80	0.01	0.05
4.7 Reading/writing/numeracy	3.8%	1.9%	6.7%	0.42	0.42	0.00	0.04
4.8 Learning difficulties (X)	2.1%	1.0%	8.8%	0.18	0.18	0.00	0.03
4.9 Any qualifications *	4.3%	3.4%	5.4%	0.91	0.90	-0.01	0.09
4.10 Attitude to education/training	6.7%	2.8%	10.1%	0.51	0.49	-0.02	0.08
5.2 Financial situation	17.5%	10.1%	16.2%	0.80	0.78	-0.02	0.21
5.3 Financial management	14.8%	8.5%	14.6%	0.73	0.72	-0.01	0.18
5.4 Illegal earnings	11.2%	5.1%	22.7%	0.45	0.41	-0.04	0.15
5.5 Over reliance on others	11.8%	6.0%	17.8%	0.48	0.46	-0.01	0.14
5.6 Severe impediment to budgeting	11.2%	5.7%	18.3%	0.48	0.46	-0.02	0.14
6.1 Current relationship with family	14.3%	9.2%	12.9%	0.70	0.71	0.01	0.17
6.2 Criminal family member (X)	4.1%	4.3%	4.0%	0.52	0.56	0.05	0.08
6.3 Experience of childhood (H)	5.6%	5.2%	3.5%	0.67	0.70	0.03	0.07
6.4 Current relationship with partner	15.5%	8.7%	22.2%	0.52	0.50	-0.02	0.20
6.5 Criminal partner (X)	3.4%	2.1%	15.2%	0.20	0.20	0.01	0.07
6.6 Previous relationship experience (H)	7.6%	6.7%	4.4%	0.76	0.80	0.04	0.09
7.1 Community integration	14.7%	8.5%	13.4%	0.93	0.90	-0.03	0.17
7.2 Activities encourage offending	16.7%	8.4%	18.1%	0.78	0.74	-0.04	0.20
7.3 Influenced by criminal peers	11.9%	6.2%	13.5%	0.67	0.66	-0.02	0.14

7.4	Manipulative lifestyle	8.4%	4.5%	17.6%	0.31	0.32	0.00	0.10
7.5	Recklessness/risktaking behaviour	14.0%	8.3%	13.1%	0.77	0.76	-0.01	0.17
8.4	Current drug misuse	6.6%	4.4%	12.0%	0.59	0.58	-0.01	0.13
8.5	Level of use of main drug	10.5%	6.1%	18.9%	0.68	0.63	-0.05	0.21
8.6	Injecting drugs #	5.0%	2.9%	9.6%	0.35	0.35	0.01	0.06
8.7	Drug-related violence ^	2.5%	2.0%	6.1%	0.24	0.26	0.02	0.05
8.8	Motivation to tackle drugs	10.3%	5.4%	18.8%	0.34	0.34	0.00	0.12
8.9	Drugs major part of lifestyle	9.7%	4.2%	22.6%	0.41	0.37	-0.04	0.13
9.1	Current alcohol use	13.9%	7.1%	20.0%	0.62	0.58	-0.03	0.17
9.2	Binge drinking	13.5%	7.6%	19.0%	0.65	0.62	-0.03	0.18
9.3	Previous alcohol use (H)	7.2%	7.0%	4.4%	0.83	0.86	0.03	0.09
9.4	Alcohol-related violence ^	4.1%	4.7%	2.9%	0.70	0.74	0.04	0.08
9.5	Motivation to tackle alcohol misuse	9.6%	4.6%	17.0%	0.37	0.36	-0.01	0.11
10.1	Coping/depression	16.8%	10.2%	15.5%	0.79	0.78	-0.01	0.19
10.2	Current psychological problems	10.9%	5.8%	16.0%	0.52	0.51	-0.01	0.13
10.3	Social isolation	10.9%	5.9%	14.7%	0.48	0.48	0.00	0.12
10.4	Attitude to self	11.4%	5.7%	13.1%	0.60	0.59	-0.01	0.13
10.5	Suicide/self-harm ^	5.1%	4.0%	8.6%	0.48	0.50	0.02	0.10
10.6	Current psychiatric problems	4.7%	2.5%	14.6%	0.23	0.23	0.00	0.06
11.1	Interpersonal skills	8.3%	4.3%	13.2%	0.37	0.37	0.00	0.09
11.2	Impulsivity	15.9%	9.5%	12.3%	1.11	1.08	-0.04	0.18
11.3	Temper control	12.0%	7.1%	13.9%	0.64	0.63	0.00	0.14
11.4	Aggressive/controlling behaviour	12.0%	7.1%	12.9%	0.66	0.66	0.00	0.14
11.5	Problem recognition	16.9%	8.4%	13.8%	0.94	0.90	-0.04	0.19
11.6	Problem solving	18.2%	10.2%	13.8%	1.19	1.13	-0.06	0.20
11.7	Awareness of consequences	18.6%	9.0%	15.6%	1.03	0.97	-0.06	0.21
11.8	Achieves goals	13.2%	6.4%	12.2%	0.88	0.85	-0.04	0.15
11.9	Understands others' views	13.0%	6.7%	12.5%	0.72	0.71	-0.02	0.15
11.10	Concrete/abstract thinking	11.1%	6.0%	11.5%	0.63	0.63	0.00	0.13
12.1	Pro-criminal attitudes	11.8%	6.7%	14.6%	0.47	0.47	0.01	0.13
12.2	Discriminatory attitudes	5.7%	4.0%	14.3%	0.18	0.20	0.02	0.07
12.3	Attitude to staff	9.0%	5.3%	20.0%	0.23	0.24	0.02	0.10
12.4	Attitude to supervision	15.5%	10.4%	17.2%	0.46	0.50	0.04	0.19
12.5	Attitude to community/society	8.3%	4.3%	13.2%	0.36	0.36	0.00	0.09
12.6	Understands motivation for offending	13.4%	6.3%	14.1%	0.70	0.67	-0.03	0.15

Key Net reduction in score No net change in score Net increase in score

^ = item score can only rise over time ("has the offender ever [experienced negative event]?"). * = item score can only fall over time ("has the offender ever [experienced positive event]?"). # = item score can rise, and can fall from 2 to 1 ("current" to "past") only over time. (H) = item score is in some other way largely based on historic events, but new events may occasionally occur, or perspectives on the past may change or a sustained change be demonstrated over the long term. (HH) = item score is in some other way wholly based on historic events and will not be changed by new events/behaviour/perspectives. (R) = item score can only change as a result of proven reoffending. [Note: 8.7 and 9.4 can be scored from other sources also.] (X) = item score is factual (about family or partner) or a diagnosis of the offender's innate character – in practice the score may be wholly beyond the offender's control. (It is assumed that other item scores – including skills, relationships, emotional wellbeing, thinking/behaviour and attitudes problems – can at least sometimes be changed, though this may be very difficult and require considerable support. The problems encompassed by some items may be within the offender's control in some cases and beyond his/her control in other cases.)

Appendix 10: The OASys Self-Assessment Questionnaire (paper version)

Do you need any help to complete this form?	<input type="checkbox"/> No	<input type="checkbox"/> Yes
---	-----------------------------	------------------------------

	Are any of these a problem for you? (please tick box)	No	Yes	Is this problem linked to your offending?
1	Finding a good place to live	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Understanding other people's feelings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Keeping to my plans	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Dealing with people in authority	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Gambling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Mixing with bad company	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	Being bored	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Being lonely	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Going to places which cause me trouble	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Taking drugs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	Drinking too much alcohol	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	Losing my temper	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	Doing things on the spur of the moment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	Repeating the same mistakes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	Getting violent when annoyed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	Reading, writing, spelling and numbers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	Getting qualifications	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	Getting a job	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	Keeping a job	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	Managing money, dealing with debts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21	Getting on with my husband/wife/partner	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22	Looking after my children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23	Worrying about things	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24	Making good decisions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25	Feeling depressed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26	Feeling stressed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27	Not having a partner	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28	Do you think you are likely to offend in the future?			

<input type="checkbox"/> Definitely not	<input type="checkbox"/> Unlikely	<input type="checkbox"/> Quite likely	<input type="checkbox"/> Very likely
Why do you think this?			

Appendix 11: A restructured Self-Assessment Questionnaire

The following ordering of questions in the SAQ would establish greater alignment with the core assessment, enabling practitioners to compare views more easily, helping to ensure that potential problems are recognised and that differences in opinion are discussed.

	Closet corresponding question in core OASys assessment
Accommodation	
1. Finding a good place to live	3.4
Education, training and employment	
2. Getting a job	4.3
3. Keeping a job	4.3
4. Reading, writing, spelling and numbers	4.7
5. Getting qualifications	4.9
Financial management and income	
6. Managing money, dealing with debts	5.3
7. Gambling	5.6
Relationships	
8. Not having a partner	6.4
9. Getting on with my husband/wife/partner	6.4
10. Looking after my children	-
Lifestyle and associates	
11. Going to places which cause me trouble	7.2
12. Mixing with bad company	7.3
13. Being bored	7.5
Drug misuse	
14. Taking drugs	8.1
Alcohol misuse	
15. Drinking too much alcohol	9.1
Emotional wellbeing	
16. Worrying about things	10.1
17. Feeling stressed	10.1
18. Feeling depressed	10.2
19. Being lonely	10.3
Thinking and behaviour	
20. Doing things on the spur of the moment	11.2
21. Getting violent when annoyed	11.3
22. Losing my temper	11.4
23. Making good decisions	11.7
24. Keeping to my plans	11.8
25. Understanding other people's feelings	11.9
26. Repeating the same mistakes	11.10
Attitudes	
27. Dealing with people in authority	12.3
Likelihood of reoffending	
28. Do you think you are likely to offend in the future?	-

Appendix 12: OASys representativeness at regional level

The main body of the report compares OASys completion rates for different offender sub-groups within the 2007 probation commencements at the national level. As shown, when using logistic regression to account for the relationships between the variables, there were significant differences in OASys completion rates across all offender characteristics. At the regional level, the differences were more pronounced for some regions than others. Tables A12.1 and A12.2 below set out the OASys completion rates for the North East and London respectively. These two regions were selected as they represented the extremes of OASys coverage – an assessment had been completed in 93% of the North East cases compared to 59% of the London cases. Consequently, London had more marked differences in completion rates between offender sub-groups. As shown by Table A12.2, the odds of OASys completion for offenders managed at OM Tier 4 within London were 4.6 times the odds of OASys completion for those at Tier 1.

Focusing upon those probation commencements with completed assessments, Tables A12.3 and A12.4 set out OASys validity rates for the North East and London respectively. As shown in the main report, the validity rates for these two regions as a whole were 99% and 82% – the highest and lowest of all the regions. Not surprisingly, there were more significant differences in the validity rates between offender sub-groups within London than within the North East. As shown by Table A12.4, the odds of valid OASys completion for offenders managed at OM Tier 4 within London were 7.6 times the odds of valid completion for those at Tier 1.

Table A12.1: North East OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements)

		OASys post-sentence eligibility											
		Non-eligible (n=2,348)				Eligible (n=11,493)				All cases (n=14,410)			
		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio	
Gender	Male	89.3%	^		94.2%	^			92.7%	^			
	Female	88.5%	n/s		94.7%	n/s			93.6%	n/s			
Age groups	18-20	88.7%	^		95.6%	^			93.7%	^			
	21-24	90.3%	n/s		93.7%	*	.603		92.8%	*	.719		
	25-40	90.5%	n/s		94.3%	**	.603		93.1%	*	.693		
	41+	84.9%	*	.520	93.2%	**	.542		90.6%	***	.538		
	White	89.6%	^		94.4%	^			93.1%	^			
Ethnicity	Black	86.4%	n/s		89.2%	*	.326		88.7%	*	.388		
	Asian	87.0%	n/s		92.9%	n/s			89.3%	n/s			
	Mixed	91.7%	n/s		94.0%	n/s			90.7%	n/s			
	Other	96.4%	n/s		95.0%	n/s			93.3%	n/s			
	Violence against the person	90.7%	n/s		94.8%	n/s			93.6%	n/s			
Offence category	Sexual offences	100.0%	n/s		91.8%	n/s			92.2%	n/s			
	Burglary	89.8%	n/s		94.9%	n/s			94.2%	*	1.532		
	Robbery	71.4%	**	.042	88.6%	n/s			88.1%	*	.490		
	Theft and handling	88.5%	n/s		94.3%	n/s			93.1%	n/s			
	Fraud and forgery	86.9%	n/s		92.2%	n/s			89.9%	n/s			
OGRS band	Criminal damage	92.2%	n/s		94.8%	n/s			93.6%	n/s			
	Drug offences	94.2%	n/s		94.8%	n/s			93.8%	*	1.732		
	Other offences	87.9%	^		93.9%	^			91.6%	^			
	Low	94.0%	^		97.5%	^			96.3%	^			
	Medium	93.4%	n/s		96.9%	n/s			96.2%	n/s			
	High	97.5%	*	2.379	97.5%	n/s			97.4%	n/s			

Table A12.1: North East OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements) (continued)

		OASys post-sentence eligibility											
		Non-eligible (n=2,348)				Eligible (n=11,493)				All cases (n=14,410)			
		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio	
OM Tier	1	88.2%	^		-				88.2%	^			
	2	92.3%	n/s		94.2%	^			93.7%	***			1.928
	3	-			95.1%	*	1.341		95.1%	***			2.342
	4	-			92.0%	n/s			92.0%	**			1.730
Sentence	Community sentence	88.0%	^		94.8%	^			92.7%	^			
	Custody/YOI	96.7%			90.2%	***	.376		90.3%	***			.559
	Suspended sentence	95.1%	**	3.288	95.8%	n/s			95.5%	*			1.371
	Other	100.0%	n/s		94.4%	n/s			96.2%	n/s			

^ Used as reference group within logistic regression. Asterisks indicate whether groups differ significantly (confidence levels * < .05, ** < .01, *** < .001).

Table A12.2: London OASys completion rates by post-sentence eligibility and offender characteristics (probation commences)

	OASys post-sentence eligibility											
	Non-eligible (n=8,785)				Eligible (n=21,011)				All cases (n=30,555)			
	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio
Gender												
	Male	^		66.7%	^		59.7%	^				
	Female	n/s		65.7%	n/s		58.3%	n/s				
Age groups												
	18-20	^		66.3%	^		59.6%	^				
	21-24	*	1.385	67.2%	n/s		58.3%	n/s				
	25-40	**	1.489	66.7%	n/s		60.0%	*			1.221	
	41+	**	1.609	65.9%	n/s		59.3%	n/s				
Ethnicity												
	White	^		72.3%	^		65.4%	^				
	Black	n/s		67.7%	*	.810	61.8%	n/s				
	Asian	n/s		68.4%	*	.772	61.1%	n/s				
	Mixed	n/s		70.9%	n/s		65.0%	n/s				
	Other	n/s		51.7%	n/s		48.9%	n/s				
	Violence against the person	n/s		69.5%	***	1.454	63.7%	***			1.337	
	Sexual offences	n/s		79.8%	n/s		79.2%	*			2.341	
	Burglary	*	2.005	69.0%	*	1.484	65.9%	**			1.518	
	Robbery	n/s		70.8%	*	1.552	70.0%	*			1.628	
Offence category												
	Theft and handling	n/s		65.3%	**	1.351	59.0%	**			1.262	
	Fraud and forgery	n/s		59.0%	n/s		52.8%	n/s				
	Criminal damage	n/s		69.0%	n/s		61.4%	n/s				
	Drug offences	n/s		61.4%	*	1.478	57.9%	**			1.483	
	Other offences	^		65.0%	^		54.0%	^				
OGRS band												
	Low	^		94.0%	^		91.3%	^				
	Medium	n/s		93.3%	n/s		91.8%	n/s				
	High	**	2.346	94.6%	n/s		94.5%	*			1.214	

Table A12.2: London OASys completion rates by post-sentence eligibility and offender characteristics (probation commencements) (continued)

		OASys post-sentence eligibility											
		Non-eligible (n=8,785)				Eligible (n=21,011)				All cases (n=30,555)			
		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio		% with OASys	Sig.	Odds ratio	
OM Tier	1	43.1%	^		0.0%				43.1%	^			
	2	46.7%	***	1.812	62.2%	^			55.9%	***			1.976
	3	-			67.5%	***	1.489		67.5%	***			3.120
	4	-			71.8%	***	2.281		71.8%	***			4.631
Sentence	Community sentence	42.9%	^		68.8%	^			58.0%	^			
	Custody/YOI	54.6%	n/s		58.1%	n/s			56.5%	n/s			
	Suspended sentence	50.9%	n/s		72.5%	n/s			66.8%	n/s			
	Other	75.0%	n/s		72.7%	n/s			72.7%	n/s			

^ Used as reference group within logistic regression. Asterisks indicate whether groups differ significantly (confidence levels * < .05, ** < .01, *** < .001).

Table A12.3: North East validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements)

		OASys post-sentence eligibility											
		Non-eligible (n=2,095)				Eligible (n=10,839)				All cases (n=13,378)			
		% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio
Gender	Male	97.5%	^		99.5%	^		99.1%	^				
	Female	97.3%	n/s		99.0%	n/s		98.8%	n/s				
Age groups	18–20	97.1%	^		99.3%	^		98.9%	^				
	21–24	97.3%	n/s		99.6%	n/s		99.1%	n/s				
	25–40	98.2%	n/s		99.5%	n/s		99.2%	n/s				
	41+	96.8%	n/s		99.0%	n/s		98.6%	n/s				
	White	97.4%	^		99.4%	^		99.1%	^				
Ethnicity	Black	100.0%	n/s		100.0%	n/s		98.9%	n/s				
	Asian	97.5%	n/s		98.9%	n/s		98.6%	n/s				
	Mixed	100.0%	n/s		98.4%	n/s		98.7%	n/s				
	Other	100.0%	n/s		92.1%	**	.118	95.2%	*	.229			
	Violence against the person	98.6%	n/s		99.4%	n/s		99.3%	n/s				
	Sexual offences	100.0%	n/s		99.4%	n/s		99.4%	n/s				
Offence category	Burglary	98.2%	n/s		99.3%	n/s		99.2%	n/s				
	Robbery	100.0%	n/s		100.0%	n/s		100.0%	n/s				
	Theft and handling	96.1%	n/s		99.4%	n/s		99.0%	n/s				
	Fraud and forgery	98.8%	n/s		99.7%	n/s		99.4%	n/s				
	Criminal damage	95.0%	n/s		99.5%	n/s		98.8%	n/s				
	Drug offences	99.1%	n/s		99.7%	n/s		99.3%	n/s				
OGRS band	Other offences	97.1%	^		99.2%	^		98.7%	^				
	Low	98.7%	^		99.2%	^		99.0%	^				
	Medium	97.6%	n/s		99.3%	n/s		99.0%	n/s				
	High	98.4%	n/s		99.7%	n/s		99.7%	n/s				

Table A12.4: London validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements)

	OASys post-sentence eligibility											
	Non-eligible (n=3,905)				Eligible (n=13,982)				All cases (n=18,177)			
	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio
Gender												
	Male	^		91.3%	^		83.0%	^				
	Female	n/s		89.5%	n/s		78.0%	n/s				
Age groups												
	18–20	^		92.0%	^		83.7%	^				
	21–24	n/s		91.9%	n/s		82.5%	n/s				
	25–40	n/s		91.1%	n/s	*	82.5%	*			1.169	
	41+	n/s		89.4%	n/s		80.7%	n/s				
Ethnicity												
	White	^		92.1%	^		84.7%	^				
	Black	***	.689	89.3%	***	.701	78.7%	***			.701	
	Asian	n/s		92.5%	n/s		82.2%	n/s				
	Mixed	n/s		92.7%	n/s		85.3%	n/s				
	Other	n/s		86.0%	n/s		75.5%	*			.767	
Offence category												
	Violence against the person	***	2.034	93.9%	***	2.324	86.6%	***			1.926	
	Sexual offences	n/s		96.2%	n/s		95.9%	*			2.284	
	Burglary	**	1.918	94.7%	***	2.153	91.4%	***			1.955	
	Robbery	**	11.140	96.1%	***	2.388	95.9%	***			2.940	
	Theft and handling	*	1.303	88.8%	n/s		82.1%	n/s				
	Fraud and forgery	n/s		87.8%	n/s		67.7%	n/s				
	Criminal damage	***	2.404	91.9%	**	1.722	85.3%	***			1.972	
	Drug offences	***	1.862	91.4%	**	1.646	86.0%	***			1.724	
	Other offences	^		87.0%	^		74.7%	^				
OGRS band												
	Low	^		92.4%	^		82.2%	^				
	Medium	*	1.261	93.1%	*	1.181	87.8%	***			1.335	
	High	***	5.099	95.7%	***	1.924	95.0%	***			2.754	

Table A12.4: London validity rates of completed OASys assessments by post-sentence eligibility and offender characteristics (probation commencements) (continued)

		OASys post-sentence eligibility										
		Non-eligible (n=3,905)			Eligible (n=13,982)			All cases (n=18,177)				
		% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	% with OASys	Sig.	Odds ratio	Odds ratio	
OM Tier	1	54.2%	^		-				54.2%	^		
	2	47.7%	***	.653	87.0%	^		74.9%	***		1.901	
	3	-			91.9%	***	1.671	91.9%	***		6.454	
	4	-			96.1%	***	2.330	96.1%	***		7.629	
Sentence	Community sentence	47.2%	^		89.8%	^		76.9%	^			
	Custody/YOI	94.3%	***	8.398	93.1%	*	1.301	92.6%	***		2.098	
	Suspended sentence	65.9%	***	2.044	91.7%	*	1.245	86.6%	***		1.678	
	Other	66.7%	n/s		87.5%	n/s		82.5%	n/s			

^ Used as reference group within logistic regression. Asterisks indicate whether groups differ significantly (confidence levels * < .05, ** < .01, *** < .001).

Appendix 13: Probation commencements' classification decision tree models

The tables within this appendix set out the importance of the independent variables to each of the classification decision tree models used to predict the risk and need levels of the 2007 probation commencements. The first three levels of each classification model are also provided. The importance of each independent variable was dependent upon how strongly it acted as a primary or surrogate splitter of cases, looking across all the nodes in the tree. A normalised importance statistic ranging from 0 to 100 is also displayed.

As shown, the OGRS score was the most important independent variable in the models predicting the likelihood of reconviction level and seven of the criminogenic needs, while the OM Tier level was the most important variable for predicting the risk of serious harm level and the criminogenic needs relating to relationships and emotional wellbeing. The offence category was the most important variable in the remaining alcohol-misuse model. Ethnicity was the least important variable in seven of the models, the sentence category in three of the models and gender in two of the models. However, ethnicity was the second most important variable in the alcohol-misuse model and gender was the second most important variable in the emotional wellbeing model.

Table A13.1: Likelihood of reconviction classification decision tree model (probation commences)

Independent variable	Importance		Normalised importance									
	Importance	Importance	Variable	Split values								
OGRS	.422	100.0%										
OM tier	.203	48.1%										
Age (years)	.063	15.0%										
Offence category	.060	14.2%										
Sentence category	.037	8.8%										
Ethnicity	.011	2.7%										
Gender	.008	1.9%										
Node	Low		Medium		High		Total		Predicted category	Parent node	Primary independent variable	
	n	%	n	%	n	%	n	%			Variable	Importment
0	47,610	28.4	84,482	50.3	35,771	21.3	167,863	100.0	Medium			
1	37,577	59.0	25,158	39.5	962	1.5	63,697	36.8	Low	0	OGRS	.253 <= 35.5
2	10,033	9.6	59,324	57.0	34,809	33.4	104,166	63.2	Medium	0	OGRS	.253 > 35.5
3	25,702	69.7	10,962	29.7	195	.5	36,859	21.5	Low	1	OM tier	.024 <= 2
4	11,875	44.2	14,196	52.9	767	2.9	26,838	15.4	Medium	1	OM tier	.024 > 2
5	9,464	15.5	40,850	67.1	10,607	17.4	60,921	35.5	Medium	2	OGRS	.105 <= 77.5
6	569	1.3	18,474	42.7	24,202	56.0	43,245	27.7	High	2	OGRS	.105 > 77.5
7	10,656	82.4	2,242	17.3	33	.3	12,931	7.6	Low	3	OM tier	.007 <= 1
8	15,046	62.9	8,720	36.4	162	.7	23,928	13.8	Low	3	OM tier	.007 > 1
9	7,701	55.1	6,116	43.8	150	1.1	13,967	8.0	Low	4	OGRS	.009 <= 18.5
10	4,174	32.4	8,080	62.8	617	4.8	12,871	7.3	Medium	4	OGRS	.009 > 18.5
11	6,461	24.9	17,413	67.2	2,034	7.9	25,908	14.8	Medium	5	OM tier	.014 <= 2
12	3,003	8.6	23,437	66.9	8,573	24.5	35,013	20.7	Medium	5	OM tier	.014 > 2
13	337	3.3	6,395	63.5	3,337	33.1	10,069	6.1	Medium	6	OM tier	.017 <= 2
14	232	.7	12,079	36.4	20,865	62.9	33,176	21.7	High	6	OM tier	.017 > 2

Table 13.2: Risk of serious harm classification decision tree model (probation commencing)

Node	Low		Medium		High		Total		Predicted category	Parent node	Variable	Improvement	Split values
	n	%	n	%	n	%	n	%					
OGRS												100.0%	
OM tier												62.9%	
Age (years)												20.7%	
Offence category												13.0%	
Sentence category												9.6%	
Ethnicity												4.1%	
Gender												1.9%	
0	71,437	41.5	88,664	51.5	12,084	7.0	172,185	100.0	Medium				
1	46,586	60.3	29,825	38.6	786	1.0	77,197	44.4	Low	0	OM tier	.125	<= 2
2	24,851	26.2	58,839	61.9	11,298	11.9	94,988	55.6	Medium	0	OM tier	.125	> 2
3	38,535	69.8	16,349	29.6	342	.6	55,226	31.8	Low	1	Offence category	.039	9; 8; 3; 6; 7; 5
4	8,051	36.6	13,476	61.3	444	2.0	21,971	12.5	Medium	1	Offence category	.039	1; 4; 2
5	22,812	29.1	52,617	67.1	3,036	3.9	78,465	44.9	Medium	2	OM tier	.096	<= 3
6	2,039	12.3	6,222	37.7	8,262	50.0	16,523	10.8	High/V. high	2	OM tier	.096	> 3
7	19,229	62.6	11,254	36.6	254	.8	30,737	17.7	Low	3	Offence category	.008	9; 3; 7
8	19,306	78.8	5,095	20.8	88	.4	24,489	14.2	Low	3	Offence category	.008	8; 6; 5
9	3,194	49.5	3,140	48.7	118	1.8	6,452	3.7	Low	4	OM tier	.003	<= 1
10	4,857	31.3	10,336	66.6	326	2.1	15,519	8.8	Medium	4	OM tier	.003	> 1
11	18,547	41.2	25,501	56.6	985	2.2	45,033	25.7	Medium	5	Offence category	.037	9; 8; 3; 6; 5
12	4,265	12.8	27,116	81.1	2,051	6.1	33,432	19.1	Medium	5	Offence category	.037	1; 4; 7; 2
13	1,425	25.7	2,860	51.6	1,256	22.7	5,541	3.4	Medium	6	Offence category	.016	8; 3; 5
14	614	5.6	3,362	30.6	7,006	63.8	10,982	7.4	High/V. high	6	Offence category	.016	9; 1; 6; 4; 7; 2

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences.

Table 13.3: Accommodation need classification decision tree model (probation commencing)

Independent variable		Importance		Normalised importance	
OGRS		.066		100.0%	
OM tier		.064		97.9%	
Offence category		.044		67.3%	
Age (years)		.004		6.6%	
Sentence category		.004		5.9%	
Gender		.003		4.9%	
Ethnicity		.003		4.8%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	105,120	61.8%	65,000	38.2%	170,120	100.0%	No				
1	54,438	72.8%	20,306	27.2%	74,744	42.3%	No	0	OM tier	.043	<= 2
2	50,682	53.1%	44,694	46.9%	95,376	57.7%	Yes	0	OM tier	.043	> 2
3	40,865	77.2%	12,065	22.8%	52,930	29.4%	No	1	OGRS	.010	<= 56.5
4	13,573	62.2%	8,241	37.8%	21,814	12.8%	No	1	OGRS	.010	> 56.5
5	34,259	59.5%	23,292	40.5%	57,551	34.1%	No	2	OGRS	.014	<= 73.5
6	16,423	43.4%	21,402	56.6%	37,825	23.7%	Yes	2	OGRS	.014	> 73.5
7	14,427	84.2%	2,716	15.8%	17,143	9.3%	No	3	OM tier	.004	<= 1
8	26,438	73.9%	9,349	26.1%	35,787	20.2%	No	3	OM tier	.004	> 1
9	4,716	69.6%	2,061	30.4%	6,777	3.9%	No	4	Offence category	.001	9; 8; 6
10	8,857	58.9%	6,180	41.1%	15,037	8.9%	No	4	Offence category	.001	3; 1; 4; 7; 5; 2
11	30,411	61.6%	18,918	38.4%	49,329	29.0%	No	5	OM tier	.004	<= 3
12	3,848	46.8%	4,374	53.2%	8,222	5.1%	Yes	5	OM tier	.004	> 3
13	13,807	46.9%	15,637	53.1%	29,444	18.2%	Yes	6	OM tier	.003	<= 3
14	2,616	31.2%	5,765	68.8%	8,381	5.4%	Yes	6	OM tier	.003	> 3

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences.

Table A13.4: Education, training and employability need classification decision tree model (probation commences)

Independent variable		Importance		Normalised Importance	
OGRS		.196		100.0%	
Offence category		.069		35.1%	
OM tier		.068		34.6%	
Age (years)		.028		14.1%	
Sentence category		.015		7.8%	
Gender		.011		5.9%	
Ethnicity		.006		3.0%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	70,730	41.7%	98,974	58.3%	169,704	100.0%	Yes				
1	51,297	60.7%	33,202	39.3%	84,499	51.5%	No	0	OGRS	.152	<= 48.5
2	19,433	22.8%	65,772	77.2%	85,205	48.5%	Yes	0	OGRS	.152	> 48.5
3	34,616	68.2%	16,169	31.8%	50,785	31.3%	No	1	OGRS	.016	<= 26.5
4	16,681	49.5%	17,033	50.5%	33,714	20.1%	No	1	OGRS	.016	> 26.5
5	12,687	31.9%	27,080	68.1%	39,767	23.0%	Yes	2	OGRS	.018	<= 76.5
6	6,746	14.8%	38,692	85.2%	45,438	25.5%	Yes	2	OGRS	.018	> 76.5
7	28,751	71.2%	11,635	28.8%	40,386	25.0%	No	3	Gender	.004	1
8	5,865	56.4%	4,534	43.6%	10,399	6.3%	No	3	Gender	.004	2
9	15,308	52.1%	14,079	47.9%	29,387	17.6%	No	4	Gender	.004	1
10	1,373	31.7%	2,954	68.3%	4,327	2.5%	Yes	4	Gender	.004	2
11	4,021	40.6%	5,891	59.4%	9,912	5.8%	Yes	5	Offence category	.003	9; 6
12	8,666	29.0%	21,189	71.0%	29,855	17.2%	Yes	5	Offence category	.003	8; 3; 1; 4; 7; 5; 2
13	5,188	17.9%	23,729	82.1%	28,917	16.3%	Yes	6	OGRS	.002	<= 91.5
14	1,558	9.4%	14,963	90.6%	16,521	9.2%	Yes	6	OGRS	.002	> 91.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Gender categories: 1 Male; 2 Female.

Table A13.5: Financial management and income need classification decision tree model (probation commencing)

Independent variable		Importance		Normalised importance	
OGRS		.114		100.0%	
Offence category		.111		97.4%	
OM tier		.025		22.1%	
Sentence category		.020		18.0%	
Gender		.011		9.2%	
Age (years)		.004		3.9%	
Ethnicity		.001		.6%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	121,615	71.6%	48,152	28.4%	169,767	100.0%	No				
1	37,129	55.2%	30,106	44.8%	67,235	43.1%	Yes	0	Offence category	.095	8; 3; 6; 4; 5
2	84,486	82.4%	18,046	17.6%	102,532	56.9%	No	0	Offence category	.095	9; 1; 7; 2
3	23,226	63.7%	13,221	36.3%	36,447	22.4%	No	1	OGRS	.015	<= 70.5
4	13,903	45.2%	16,885	54.8%	30,788	20.7%	Yes	1	OGRS	.015	> 70.5
5	64,107	87.2%	9,402	12.8%	73,509	39.7%	No	2	OGRS	.022	<= 64.5
6	20,379	70.2%	8,644	29.8%	29,023	17.2%	No	2	OGRS	.022	> 64.5
7	4,948	74.9%	1,661	25.1%	6,609	3.8%	No	3	OM tier	.003	<= 1
8	18,278	61.3%	11,560	38.7%	29,838	18.5%	Yes	3	OM tier	.003	> 1
9	4,246	56.6%	3,257	43.4%	7,503	4.8%	Yes	4	OM tier	.003	<= 2
10	9,657	41.5%	13,628	58.5%	23,285	15.9%	Yes	4	OM tier	.003	> 2
11	35,365	89.9%	3,976	10.1%	39,341	20.9%	No	5	OGRS	.003	<= 28.5
12	28,742	84.1%	5,426	15.9%	34,168	18.8%	No	5	OGRS	.003	> 28.5
13	15,112	74.4%	5,196	25.6%	20,308	11.8%	No	6	OGRS	.004	<= 88.5
14	5,267	60.4%	3,448	39.6%	8,715	5.4%	Yes	6	OGRS	.004	> 88.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences.

Table A13.6: Relationships need classification decision tree model (probation commencing)

Independent variable		Importance		Normalised importance	
OM tier		.068		100.0%	
OGRS		.059		86.6%	
Gender		.050		73.7%	
Offence category		.029		42.3%	
Ethnicity		.019		27.4%	
Age (years)		.013		18.8%	
Sentence category		.003		4.2%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	96,911	57.5%	71,719	42.5%	168,630	100.0%	No				
1	51,821	70.3%	21,890	29.7%	73,711	43.1%	No	0	OM tier	.053	<= 2
2	45,090	47.5%	49,829	52.5%	94,919	56.9%	Yes	0	OM tier	.053	> 2
3	45,531	74.8%	15,355	25.2%	60,886	35.4%	No	1	Gender	.018	1
4	6,290	49.0%	6,535	51.0%	12,825	7.7%	Yes	1	Gender	.018	2
5	42,717	49.9%	42,823	50.1%	85,540	51.2%	Yes	2	Gender	.012	1
6	2,373	25.3%	7,006	74.7%	9,379	5.8%	Yes	2	Gender	.012	2
7	33,908	79.8%	8,575	20.2%	42,483	24.6%	No	3	OGRS	.009	<= 58.5
8	11,623	63.2%	6,780	36.8%	18,403	10.8%	No	3	OGRS	.009	> 58.5
9	4,985	56.2%	3,893	43.8%	8,878	5.3%	No	4	OGRS	.003	<= 35.5
10	1,305	33.1%	2,642	66.9%	3,947	2.4%	Yes	4	OGRS	.003	> 35.5
11	37,200	53.3%	32,654	46.7%	69,854	41.6%	No	5	OM tier	.010	<= 3
12	5,517	35.2%	10,169	64.8%	15,686	9.5%	Yes	5	OM tier	.010	> 3
13	879	37.2%	1,481	62.8%	2,360	1.4%	Yes	6	OGRS	.001	<= 22.5
14	1,494	21.3%	5,525	78.7%	7,019	4.3%	Yes	6	OGRS	.001	> 22.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Gender categories: 1 Male; 2 Female.

Table A 13.7: Lifestyle and associates need classification decision tree model (probation commencing)

Independent variable	Importance		Normalised importance	
	Importance	Importance	Normalised importance	Normalised importance
OGRS	.194	.194	100.0%	100.0%
OM tier	.134	.134	69.2%	69.2%
Offence category	.086	.086	44.5%	44.5%
Sentence category	.082	.082	42.5%	42.5%
Age (years)	.022	.022	11.3%	11.3%
Gender	.004	.004	2.0%	2.0%
Ethnicity	.001	.001	.3%	.3%

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	92,761	54.7%	76,852	45.3%	169,613	100.0%	No				
1	68,041	71.6%	27,047	28.4%	95,088	55.2%	No	0	OGRS	.148	<= 56.5
2	24,720	33.2%	49,805	66.8%	74,525	44.8%	Yes	0	OGRS	.148	> 56.5
3	41,720	79.8%	10,579	20.2%	52,299	30.1%	No	1	OM tier	.020	<= 2
4	26,321	61.5%	16,468	38.5%	42,789	25.1%	No	1	OM tier	.020	> 2
5	16,027	42.5%	21,709	57.5%	37,736	22.5%	Yes	2	OGRS	.015	<= 81.5
6	8,693	23.6%	28,096	76.4%	36,789	22.3%	Yes	2	OGRS	.015	> 81.5
7	31,649	84.9%	5,632	15.1%	37,281	21.3%	No	3	Offence category	.009	9; 1; 6
8	10,071	67.1%	4,947	32.9%	15,018	8.7%	No	3	Offence category	.009	8; 3; 4; 7; 5; 2
9	21,026	67.4%	10,158	32.6%	31,184	18.2%	No	4	Offence category	.010	9; 1; 6; 7
10	5,295	45.6%	6,310	54.4%	11,605	6.9%	Yes	4	Offence category	.010	8; 3; 4; 5; 2
11	7,263	53.5%	6,306	46.5%	13,569	8.0%	No	5	OM tier	.006	<= 2
12	8,764	36.3%	15,403	63.7%	24,167	14.5%	Yes	5	OM tier	.006	> 2
13	2,928	35.6%	5,303	64.4%	8,231	4.9%	Yes	6	OM tier	.003	<= 2
14	5,765	20.2%	22,793	79.8%	28,558	17.4%	Yes	6	OM tier	.003	> 2

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences.

Table A13.8: Drug misuse need classification decision tree model (probation commencing)

Independent variable		Importance		Normalised importance	
OGRS		.197		100.0%	
Offence category		.157		80.0%	
OM tier		.059		30.2%	
Age (years)		.033		16.6%	
Sentence category		.016		8.2%	
Gender		.005		2.5%	
Ethnicity		.003		1.6%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	123,168	70.6%	51,338	29.4%	174,506	100.0%	No				
1	29,390	48.1%	31,726	51.9%	61,116	39.0%	Yes	0	Offence category	.139	8; 3; 4; 5
2	93,778	82.7%	19,612	17.3%	113,390	61.0%	No	0	Offence category	.139	9; 1; 6; 7; 2
3	15,511	64.1%	8,685	35.9%	24,196	14.3%	No	1	OM tier	.023	<= 2
4	13,879	37.6%	23,041	62.4%	36,920	24.6%	Yes	1	OM tier	.023	> 2
5	66,885	90.4%	7,079	9.6%	73,964	38.2%	No	2	OGRS	.045	<= 54.5
6	26,893	68.2%	12,533	31.8%	39,426	22.9%	No	2	OGRS	.045	> 54.5
7	7,283	75.6%	2,352	24.4%	9,635	5.4%	No	3	OGRS	.006	<= 39.5
8	8,228	56.5%	6,333	43.5%	14,561	8.9%	Yes	3	OGRS	.006	> 39.5
9	5,873	52.1%	5,407	47.9%	11,280	7.1%	Yes	4	Age (years)	.007	<= 23.6
10	8,006	31.2%	17,634	68.8%	25,640	17.6%	Yes	4	Age (years)	.007	> 23.6
11	41,366	94.3%	2,511	5.7%	43,877	22.2%	No	5	OGRS	.007	<= 26.5
12	25,519	84.8%	4,568	15.2%	30,087	16.0%	No	5	OGRS	.007	> 26.5
13	9,790	78.6%	2,663	21.4%	12,453	6.8%	No	6	OM tier	.006	<= 2
14	17,103	63.4%	9,870	36.6%	26,973	16.0%	No	6	OM tier	.006	> 2

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences.

Table A13.9: Alcohol misuse need classification decision tree model (probation commences)

Independent variable		Importance		Normalised importance	
Offence category		.068		100.0%	
Ethnicity		.037		54.1%	
OGRS		.028		41.3%	
OM tier		.027		40.3%	
Age (years)		.009		13.7%	
Gender		.007		10.0%	
Sentence category		.003		3.9%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	90,406	53.0%	80,222	47.0%	170,628	100.0%	No				
1	45,098	43.4%	58,701	56.6%	103,799	60.4%	Yes	0	Offence category	.056	9; 1; 4; 7
2	45,308	67.8%	21,521	32.2%	66,829	39.6%	No	0	Offence category	.056	8; 3; 6; 5; 2
3	35,423	39.4%	54,541	60.6%	89,964	52.1%	Yes	1	Ethnicity	.027	1
4	96,75	69.9%	4,160	30.1%	13,835	8.2%	No	1	Ethnicity	.027	3; 4; 2; 5
5	16,928	80.7%	4,050	19.3%	20,978	12.6%	No	2	OGRS	.011	<= 39.5
6	28,380	61.9%	17,471	38.1%	45,851	27.1%	No	2	OGRS	.011	> 39.5
7	7,014	60.3%	4,621	39.7%	11,635	6.9%	No	3	OM tier	.014	<= 1
8	28,409	36.3%	49,920	63.7%	78,329	45.3%	Yes	3	OM tier	.014	> 1
9	1,516	56.2%	1,181	43.8%	2,697	1.6%	No	4	Ethnicity	.001	4
10	8,159	73.3%	2,979	26.7%	11,138	6.6%	No	4	Ethnicity	.001	3; 2; 5
11	1,961	66.4%	991	33.6%	2,952	1.7%	No	5	Offence category	.002	3
12	14,967	83.0%	3,059	17.0%	18,026	10.8%	No	5	Offence category	.002	8; 6; 5; 2
13	24,259	59.7%	16,357	40.3%	40,616	23.9%	No	6	Ethnicity	.004	1
14	4,121	78.7%	1,114	21.3%	5,235	3.1%	No	6	Ethnicity	.004	3; 4; 2; 5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Ethnicity categories: 1 White; 2 Black; 3 Asian; 4 Mixed; 5 Other

Table A13.10: Emotional wellbeing need classification decision tree model (probation commences)

Independent variable		Importance		Normalised importance	
OM tier		.046		100.0%	
Gender		.030		65.5%	
Age (years)		.022		48.4%	
Ethnicity		.021		44.4%	
Offence category		.011		24.8%	
OGRS		.008		18.0%	
Sentence category		.002		3.3%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	102,097	58.4%	72,596	41.6%	174,693	100.0%	No				
1	93,730	61.8%	57,865	38.2%	151,595	86.2%	No	0	Gender	.030	1
2	8,367	36.2%	14,731	63.8%	23,098	13.8%	Yes	0	Gender	.030	2
3	46,628	71.8%	18,315	28.2%	64,943	36.1%	No	1	OM tier	.029	<= 2
4	47,102	54.4%	39,550	45.6%	86,652	50.0%	Yes	1	OM tier	.029	> 2
5	2,244	59.6%	1,519	40.4%	3,763	2.1%	No	2	OM tier	.006	<= 1
6	6,123	31.7%	13,212	68.3%	19,335	11.7%	Yes	2	OM tier	.006	> 1
7	16,307	82.4%	3,475	17.6%	19,782	10.8%	No	3	OM tier	.009	<= 1
8	30,321	67.1%	14,840	32.9%	45,161	25.4%	No	3	OM tier	.009	> 1
9	25,661	61.0%	16,394	39.0%	42,055	23.9%	No	4	Age (years)	.009	<= 27.4
10	21,441	48.1%	23,156	51.9%	44,597	26.1%	Yes	4	Age (years)	.009	> 27.4
11	1,073	69.8%	464	30.2%	1,537	.9%	No	5	Age (years)	.001	<= 28.8
12	1,171	52.6%	1,055	47.4%	2,226	1.3%	Yes	5	Age (years)	.001	> 28.8
13	5,449	30.3%	12,544	69.7%	17,993	10.9%	Yes	6	Ethnicity	.001	1; 4
14	674	50.2%	668	49.8%	1,342	.8%	Yes	6	Ethnicity	.001	3; 2; 5

Key to Ethnicity categories: 1 White; 2 Black; 3 Asian; 4 Mixed; 5 Other. Key to Gender categories: 1 Male; 2 Female.

Table A13.11: Thinking and behaviour need classification decision tree model (probation commencing)

Independent variable		Importance		Normalised importance	
OGRS		.146		100.0%	
OM tier		.127		86.6%	
Offence category		.051		35.1%	
Age (years)		.019		12.9%	
Sentence category		.014		9.8%	
Gender		.007		5.1%	
Ethnicity		.007		4.5%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	70,399	40.3%	104,482	59.7%	174,881	100.0%	Yes				
1	45,750	58.0%	33,164	42.0%	78,914	47.7%	No	0	OM tier	.111	<= 2
2	24,649	25.7%	71,318	74.3%	95,967	52.3%	Yes	0	OM tier	.111	> 2
3	32,824	68.3%	15,207	31.7%	48,031	29.9%	No	1	OGRS	.028	<= 43.5
4	12,926	41.9%	17,957	58.1%	30,883	17.8%	Yes	1	OGRS	.028	> 43.5
5	13,183	35.4%	24,042	64.6%	37,225	21.0%	Yes	2	OGRS	.017	<= 49.5
6	11,466	19.5%	47,276	80.5%	58,742	31.3%	Yes	2	OGRS	.017	> 49.5
7	13,549	79.6%	3,480	20.4%	17,029	11.0%	No	3	OM tier	.006	<= 1
8	19,275	62.2%	11,727	37.8%	31,002	19.0%	No	3	OM tier	.006	> 1
9	9,545	47.2%	10,680	52.8%	20,225	11.8%	No	4	OGRS	.004	<= 76.5
10	3,381	31.7%	7,277	68.3%	10,658	5.9%	Yes	4	OGRS	.004	> 76.5
11	6,605	29.4%	15,876	70.6%	22,481	12.4%	Yes	5	Offence category	.005	3; 1; 4; 7
12	6,578	44.6%	8,166	55.4%	14,744	8.5%	No	5	Offence category	.005	9; 8; 6; 5; 2
13	10,403	21.9%	37,143	78.1%	47,546	25.6%	Yes	6	OM tier	.004	<= 3
14	1,063	9.5%	10,133	90.5%	11,196	5.8%	Yes	6	OM tier	.004	> 3

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences.

Table A13.12: Attitudes need classification decision tree model (probation commencing)

Independent variable		Importance		Normalised importance	
OGRS		.160		100.0%	
OM tier		.109		68.0%	
Sentence category		.042		26.1%	
Offence category		.024		15.3%	
Age (years)		.012		7.7%	
Ethnicity		.001		.4%	
Gender		.000		.3%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%			Variable	Improvement	Split values
0	111,630	63.9%	63,066	36.1%	174,696	100.0%	No				
1	77,760	78.4%	21,481	21.6%	99,241	55.1%	No	0	OGRS	.121	<= 56.5
2	33,870	44.9%	41,585	55.1%	75,455	44.9%	Yes	0	OGRS	.121	> 56.5
3	47,469	84.9%	8,425	15.1%	55,894	30.6%	No	1	OM tier	.016	<= 2
4	30,291	69.9%	13,056	30.1%	43,347	24.5%	No	1	OM tier	.016	> 2
5	18,380	55.4%	14,799	44.6%	33,179	19.3%	No	2	OGRS	.015	<= 78.5
6	15,490	36.6%	26,786	63.4%	42,276	25.6%	Yes	2	OGRS	.015	> 78.5
7	30,170	89.3%	3,627	10.7%	33,797	18.3%	No	3	OGRS	.005	<= 28.5
8	17,299	78.3%	4,798	21.7%	22,097	12.3%	No	3	OGRS	.005	> 28.5
9	27,204	73.2%	9,978	26.8%	37,182	20.9%	No	4	OM tier	.007	<= 3
10	3,087	50.1%	3,078	49.9%	6,165	3.6%	Yes	4	OM tier	.007	> 3
11	8,239	64.9%	4,463	35.1%	12,702	7.3%	No	5	OM tier	.005	<= 2
12	10,141	49.5%	10,336	50.5%	20,477	12.1%	Yes	5	OM tier	.005	> 2
13	13,873	39.9%	20,877	60.1%	34,750	20.9%	Yes	6	OM tier	.004	<= 3
14	1,617	21.5%	5,909	78.5%	7,526	4.7%	Yes	6	OM tier	.004	> 3

Appendix 14: Risk and need residual values and percentages correctly predicted for classification decision tree models (probation commencements)

Commentary regarding the goodness-of-fit/accuracy of the classification models for all the probation commencements is provided in the main body of the report. Table A14.1 compares the model fit/accuracy for those cases which met the post-sentence eligibility for OASys completion and those cases which did not. As shown, the percentages correctly predicted were higher in the non-eligible cases for all the models except education, training and employability. The best performing model in the non-eligible cases was the drug-misuse model with a correct classification in 88% of the cases. However, the residuals were much greater in the non-eligible cases with under-predictions of all criminogenic need rates, indicating that the use of OASys in these cases was unlikely to be random. Consequently, practitioners were not differing as much in their assessments of criminogenic needs between the non-eligible and eligible cases as suggested by the basic offender characteristic variables. For thinking and behaviour (which had the largest residuals), the predicted levels of need in the non-eligible and eligible cases were 14% and 72% respectively, whereas the actual recorded levels were much closer at 32% and 68%.

Table A14.1: Goodness-of-fit and accuracy of classification models by OASys post-sentence eligibility (probation commencements)

Risk and need level	OASys post-sentence eligibility					
	Non-eligible (n=27,133)		Eligible (n=130,797)		All cases (n=165,830)	
	Residual (actual minus predicted rate)	Per cent correctly predicted	Residual (actual minus predicted rate)	Per cent correctly predicted	Residual (actual minus predicted rate)	Per cent correctly predicted
Likelihood of reconviction:						
Low	-3.9%	75.7%	1.0%	68.2%	0.1%	69.4%
Medium	2.2%		-0.7%		-0.3%	
High	1.7%		-0.3%		0.2%	
Risk of serious harm:						
Low	-12.2%	69.7%	0.5%	67.8%	-1.2%	67.9%
Medium	11.5%		-1.4%		0.2%	
High/Very high	0.7%		0.9%		1.0%	
Criminogenic need:						
Accommodation	15.3%	80.0%	-3.5%	61.8%	-0.3%	64.9%
ETE	10.6%	70.8%	-2.4%	71.9%	-0.2%	71.7%
Finance	10.2%	83.2%	-1.1%	71.4%	0.9%	73.2%
Relationships	12.7%	79.2%	-3.8%	64.0%	-0.4%	66.6%
Lifestyle & associates	10.5%	81.4%	-3.0%	70.6%	-0.6%	72.2%
Drug misuse	5.5%	88.1%	-0.7%	75.3%	0.5%	77.4%
Alcohol misuse	13.1%	69.1%	-2.3%	66.1%	0.3%	66.7%
Emotional wellbeing	11.2%	77.0%	-2.2%	61.1%	0.6%	63.9%
Thinking & behaviour	18.1%	70.8%	-4.6%	70.6%	0.0%	70.3%
Attitudes	8.5%	81.8%	-2.5%	68.5%	-0.5%	70.8%

Appendix 15: Prison receptions' classification decision tree models

The tables in this appendix set out the importance of the independent variables to each of the classification decision tree models used to predict the risk and need levels of the 2007 prison receptions. The first three levels of each classification model are also provided. The importance of each independent variable was dependent upon how strongly it acted as a primary or surrogate splitter of cases, looking across all the nodes in the tree. A normalised importance statistic ranging from 0 to 100 is also displayed.

As shown, the offence category was the most important independent variable in all of the models except the emotional wellbeing model, for which age was most important. Gender was the least important variable in seven of the models, although it was the second most important variable in the emotional wellbeing model. While ethnicity was the least important variable in the risk of serious harm and financial-management models, it was the second most important variable in the relationships and alcohol misuse models.

Table A15.1: Likelihood of reconviction classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.121		100.0%	
No. previous convictions		.116		96.3%	
No. previous custodial sentences		.105		86.8%	
Sentence length (days)		.055		45.5%	
Security category		.050		41.5%	
Age (years)		.043		35.6%	
Ethnicity		.031		25.7%	
Gender		.004		3.4%	

Node	Low		Medium		High		Total		Predicted category	Parent node	Primary independent variable	
	n	%	n	%	n	%	n	%			Variable	Improve-ment
0	7,865	14.8%	24,825	46.7%	20,439	38.5%	53,129	100.0%	Medium			
1	740	4.6%	5,877	36.8%	9,371	58.6%	15,988	29.0%	High	0	Offence category	.080
2	7,125	19.2%	18,948	51.0%	11,068	29.8%	37,141	71.0%	Medium	0	Offence category	.080
3	544	3.6%	5,395	35.9%	9,097	60.5%	15,036	27.2%	High	1	Security category	.012
4	196	20.6%	482	50.6%	274	28.8%	952	1.8%	Medium	1	Security category	.012
5	7,095	21.3%	17,564	52.6%	8,717	26.1%	33,376	64.3%	Medium	2	No. previous convictions	.039
6	30	.8%	1,384	36.8%	2,351	62.4%	3,765	6.7%	High	2	No. previous convictions	.039
7	539	4.3%	4,863	38.9%	7,093	56.8%	12,495	22.6%	High	3	No. previous convictions	.008
8	5	.2%	532	20.9%	2,004	78.9%	2,541	4.6%	High	3	No. previous convictions	.008
9	196	23.3%	444	52.7%	202	24.0%	842	1.6%	Medium	4	No. previous custodial sentences	.001
10	0	.0%	38	34.5%	72	65.5%	110	.2%	High	4	No. previous custodial sentences	.001
11	1,370	43.4%	1,484	47.1%	300	9.5%	3,154	6.7%	Low	5	Offence category	.020
12	5,725	18.9%	16,080	53.2%	8,417	27.9%	30,222	57.5%	High	5	Offence category	.020
13	23	1.4%	787	49.3%	786	49.2%	1,596	2.8%	High	6	No. previous custodial sentences	.003
14	7	.3%	597	27.5%	1,565	72.2%	2,169	3.9%	High	6	No. previous custodial sentences	.003

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Security categories: 1 Category A; 2 Category B; 3 Category C; 4 Category D; 5 YOI open/short; 6 YOI closed; 7 uncatagorised.

Table A15.2: Risk of serious harm classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.199		100.0%	
Sentence length (days)		.179		90.0%	
Security category		.080		40.0%	
Age (years)		.030		15.2%	
No. previous convictions		.019		9.7%	
Gender		.015		7.4%	
No. previous custodial sentences		.015		7.4%	
Ethnicity		.014		7.2%	

Node	Low		Medium		High		Total		Predicted category	Parent node	Primary independent variable		
	n	%	n	%	n	%	n	%			Variable	Improve-ment	Split values
0	1,856	10.3%	10,959	60.6%	5,272	29.1%	18,087	36.1%	Medium	0	Offence category	.120	1; 2; 4; 7
1	15,086	42.6%	17,848	50.4%	2,487	7.0%	35,421	63.9%	Low	0	Offence category	.120	3; 5; 6; 8; 9
2	1,720	11.8%	9,747	66.6%	3,165	21.6%	14,632	27.6%	Medium	1	Sentence length (days)	.043	<= 1365.5
3	136	3.9%	1,212	35.1%	2,107	61.0%	3,455	8.5%	High/V. high	1	Sentence length (days)	.043	> 1365.5
4	8,612	54.5%	6,709	42.4%	488	3.1%	15,809	28.4%	Low	2	Offence category	.035	5; 6; 8
5	6,474	33.0%	11,139	56.8%	1,999	10.2%	19,612	35.5%	Medium	2	Offence category	.035	3; 9
6	1,487	13.2%	7,657	67.9%	2,128	18.9%	11,272	20.9%	Medium	3	Sentence length (days)	.005	<= 567
7	233	6.9%	2,090	62.2%	1,037	30.9%	3,360	6.7%	Medium	3	Sentence length (days)	.005	> 567
8	119	5.6%	1,000	47.2%	999	47.2%	2,118	4.8%	High/V. high	4	Sentence length (days)	.006	<= 4810.5
9	17	1.3%	212	15.9%	1,108	82.9%	1,337	3.7%	High/V. high	4	Sentence length (days)	.006	> 4810.5
10	1,134	80.8%	260	18.5%	9	.6%	1,403	2.6%	Low	5	Offence category	.007	6
11	7,478	51.9%	6,449	44.8%	479	3.3%	14,406	25.8%	Low	5	Offence category	.007	5; 8
12	6,423	33.5%	10,980	57.3%	1,753	9.2%	19,156	34.4%	Medium	6	Sentence length (days)	.012	<= 1826.5
13	51	11.2%	159	34.9%	246	53.9%	456	1.1%	High/V. high	6	Sentence length (days)	.012	> 1826.5
14	1,856	10.3%	10,959	60.6%	5,272	29.1%	18,087	36.1%	Medium	0	Offence category	.120	1; 2; 4; 7

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and forgery; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences.

Table A15.3: Accommodation need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.024		100.0%	
Security category		.020		83.7%	
Sentence length (days)		.017		70.3%	
No. previous convictions		.013		54.3%	
No. previous custodial sentences		.012		50.8%	
Ethnicity		.009		37.1%	
Age (years)		.006		24.7%	
Gender		.002		6.6%	

Node	No		Yes		Total	Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%				Variable	Improvement	
0	26,587	48.7%	28,048	51.3%	54,635	100.0%	Yes			
1	6,431	39.1%	10,005	60.9%	16,436	29.9%	Yes	0	Offence category	.016
2	20,156	52.8%	18,043	47.2%	38,199	70.1%	No	0	Offence category	.016
3	508	60.7%	329	39.3%	837	1.5%	No	1	Security category	.004
4	5,923	38.0%	9,676	62.0%	15,599	28.4%	Yes	1	Security category	.004
5	18,291	51.4%	17,300	48.6%	35,591	65.2%	No	2	Security category	.009
6	1,865	71.5%	743	28.5%	2,608	4.8%	No	2	Security category	.009
7	195	50.5%	191	49.5%	386	.7%	No	3	Sentence length (days)	.001
8	313	69.4%	138	30.6%	451	.8%	No	3	Sentence length (days)	.001
9	263	54.5%	220	45.5%	483	.9%	No	4	Ethnicity	.001
10	5,660	37.4%	9,456	62.6%	15,116	27.5%	Yes	4	Ethnicity	.001
11	1,402	69.2%	624	30.8%	2,026	3.7%	No	5	Ethnicity	.005
12	16,889	50.3%	16,676	49.7%	33,565	61.5%	No	5	Ethnicity	.005
13	907	67.3%	441	32.7%	1,348	2.5%	No	6	Sentence length (days)	.000
14	958	76.0%	302	24.0%	1,260	2.3%	No	6	Sentence length (days)	.000

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Ethnicity categories: 1 White; 2 Black; 3 Asian; 4 Mixed; 5 Other. Key to Security categories: 1 Category A; 2 Category B; 3 Category C; 4 Category D; 5 YOI closed; 6 YOI open/short; 7 uncatagorised

Table A15.4: Education, training and employability need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.056		100.0%	
Age (years)		.047		83.2%	
No. previous convictions		.047		83.0%	
No. previous custodial sentences		.043		76.4%	
Security category		.039		70.3%	
Sentence length (days)		.031		55.9%	
Ethnicity		.010		17.7%	
Gender		.002		3.7%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%	n	%			Variable	Improvement	
0	16,415	30.2%	38,006	69.8%	54,421	100.0%	Yes				
1	3,561	18.4%	15,812	81.6%	19,373	33.8%	Yes	0	Offence category	.040	3; 4; 5; 7
2	12,854	36.7%	22,194	63.3%	35,048	66.2%	Yes	0	Offence category	.040	1; 2; 6; 8; 9
3	352	41.8%	490	58.2%	842	1.6%	No	1	Security category	.006	4
4	3,209	17.3%	15,322	82.7%	18,531	32.2%	Yes	1	Security category	.006	5; 3; 2; 7; 1; 6
5	4,529	28.5%	11,351	71.5%	15,880	29.0%	Yes	2	Sentence length (days)	.017	<= 182.5
6	8,325	43.4%	10,843	56.6%	19,168	37.2%	No	2	Sentence length (days)	.017	> 182.5
7	49	21.9%	175	78.1%	224	.4%	Yes	3	Offence category	.001	3
8	303	49.0%	315	51.0%	618	1.2%	No	3	Offence category	.001	4; 5; 7
9	2,638	15.9%	13,992	84.1%	16,630	28.7%	Yes	4	Ethnicity	.004	1; 4
10	571	30.0%	1,330	70.0%	1,901	3.5%	Yes	4	Ethnicity	.004	5; 3; 2
11	823	18.0%	3,755	82.0%	4,578	8.0%	Yes	5	Age (years)	.007	<= 21.5
12	3,706	32.8%	7,596	67.2%	11,302	21.0%	Yes	5	Age (years)	.007	> 21.5
13	7,936	45.7%	9,435	54.3%	17,371	34.0%	No	6	No. previous convictions	.008	<= 12.5
14	389	21.6%	1,408	78.4%	1,797	3.2%	Yes	6	No. previous convictions	.008	> 12.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Ethnicity categories: 1 White; 2 Black; 3 Asian; 4 Mixed; 5 Other. Key to Security categories: 1 Category A; 2 Category B; 3 Category C; 4 Category D; 5 YOI closed; 6 YOI open/short; 7 Uncategorised

Table A15.5: Financial management and income need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.094		100.0%	
No. previous convictions		.020		21.1%	
Sentence length (days)		.019		20.6%	
No. previous custodial sentences		.019		19.7%	
Age (years)		.013		13.9%	
Gender		.011		12.0%	
Security category		.011		11.6%	
Ethnicity		.006		6.6%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%	n	%			Variable	Improvement	
0	32,541	59.7%	21,924	40.3%	54,465	100.0%	No				
1	10,964	44.9%	13,460	55.1%	24,424	46.0%	Yes	0	Offence category	.075	3; 4; 5; 6; 8
2	21,577	71.8%	8,464	28.2%	30,041	54.0%	No	0	Offence category	.075	1; 2; 7; 9
3	5,972	39.0%	9,356	61.0%	15,328	29.2%	Yes	1	Offence category	.010	3; 5
4	4,992	54.9%	4,104	45.1%	9,096	16.8%	No	1	Offence category	.010	4; 6; 8
5	12,417	67.3%	6,040	32.7%	18,457	33.4%	No	2	Sentence length (days)	.008	<= 272.5
6	9,160	79.1%	2,424	20.9%	11,584	20.5%	No	2	Sentence length (days)	.008	> 272.5
7	4,330	36.3%	7,590	63.7%	11,920	22.8%	Yes	3	Security category	.002	3; 2; 7
8	1,642	48.2%	1,766	51.8%	3,408	6.4%	Yes	3	Security category	.002	5; 1; 4; 6
9	4,455	56.4%	3,442	43.6%	7,897	14.6%	No	4	No. previous convictions	.001	<= 7.5
10	537	44.8%	662	55.2%	1,199	2.3%	Yes	4	No. previous convictions	.001	> 7.5
11	1,247	82.2%	270	17.8%	1,517	2.7%	No	5	Security category	.003	4
12	11,170	65.9%	5,770	34.1%	16,940	30.8%	No	5	Security category	.003	5; 3; 2; 7; 1; 6
13	8,673	80.5%	2,099	19.5%	10,772	19.1%	No	6	No. previous custodial sentences	.003	<= 3.5
14	487	60.0%	325	40.0%	812	1.5%	No	6	No. previous custodial sentences	.003	> 3.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Security categories: 1 Category A; 2 Category A; 3 Category B; 4 Category C; 5 YOI closed; 6 YOI open/short; 7 Uncategorised

Table A15.6: Relationships need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.024		100.0%	
Ethnicity		.023		97.2%	
Security category		.018		75.4%	
Gender		.018		73.6%	
Sentence length (days)		.016		68.1%	
Age (years)		.016		67.9%	
No. previous convictions		.014		59.2%	
No. previous custodial sentences		.011		44.4%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%	n	%			Variable	Improvement	
0	26,688	49.6%	27,157	50.4%	53,845	100.0%	No				
1	21,891	46.7%	24,975	53.3%	46,866	86.7%	Yes	0	Ethnicity	.021	1; 4
2	4,797	68.7%	2,182	31.3%	6,979	13.3%	No	0	Ethnicity	.021	5; 3; 2
3	20,839	48.8%	21,853	51.2%	42,692	79.2%	No	1	Gender	.016	1
4	1,052	25.2%	3,122	74.8%	4,174	7.5%	Yes	1	Gender	.016	2
5	1,665	59.9%	1,115	40.1%	2,780	5.2%	No	2	Offence category	.003	1; 3; 5; 7
6	3,132	74.6%	1,067	25.4%	4,199	8.0%	No	2	Offence category	.003	2; 4; 6; 8; 9
7	14,065	45.4%	16,937	54.6%	31,002	57.3%	Yes	3	Security category	.011	3; 2; 7; 1
8	6,774	57.9%	4,916	42.1%	11,690	21.9%	No	3	Security category	.011	5; 4; 6
9	93	52.0%	86	48.0%	179	.3%	No	4	Offence category	.001	6
10	959	24.0%	3,036	76.0%	3,995	7.2%	Yes	4	Offence category	.001	1; 2; 3; 4; 5; 7; 8; 9
11	1,059	67.8%	503	32.2%	1,562	3.0%	No	5	Age (years)	.002	<= 29.5
12	606	49.8%	612	50.2%	1,218	2.3%	No	5	Age (years)	.002	> 29.5
13	2,978	76.1%	936	23.9%	3,914	7.5%	No	6	No. previous convictions	.001	<= 11.5
14	154	54.0%	131	46.0%	285	.5%	No	6	No. previous convictions	.001	> 11.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Ethnicity categories: 1 White; 2 Black; 3 Asian; 4 Mixed; 5 Other. Key to Security categories: 1 Category A; 2 Category B; 3 Category C; 4 Category D; 5 YOI closed; 6 YOI open/short; 7 Uncategorised. Key to Gender categories: 1 Male; 2 Female.

Table A15.7: Lifestyle and associates need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.054		100.0%	
Security category		.032		58.9%	
No. previous convictions		.030		54.2%	
No. previous custodial sentences		.026		47.0%	
Sentence length (days)		.019		35.7%	
Age (years)		.018		33.1%	
Ethnicity		.005		8.9%	
Gender		.001		1.6%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%	n	%			Variable	Improvement	
0	19,379	35.6%	35,079	64.4%	54,458	100.0%	Yes				
1	4,370	22.6%	15,005	77.4%	19,375	33.9%	Yes	0	Offence category	.044	3; 4; 5; 7
2	15,009	42.8%	20,074	57.2%	35,083	66.1%	No	0	Offence category	.044	1; 2; 6; 8; 9
3	406	48.3%	435	51.7%	841	1.6%	No	1	Security category	.007	4
4	3,964	21.4%	14,570	78.6%	18,534	32.3%	Yes	1	Security category	.007	5; 3; 2; 7; 1; 6
5	1,599	68.0%	754	32.0%	2,353	4.8%	No	2	Security category	.014	4
6	13,410	41.0%	19,320	59.0%	32,730	61.3%	Yes	2	Security category	.014	5; 3; 2; 7; 1; 6
7	77	34.2%	148	65.8%	225	.4%	Yes	3	Offence category	.000	3
8	329	53.4%	287	46.6%	616	1.2%	No	3	Offence category	.000	4; 5; 7
9	3,461	22.6%	11,823	77.4%	15,284	26.7%	Yes	4	No. previous convictions	.001	<= 13.5
10	503	15.5%	2,747	84.5%	3,250	5.5%	Yes	4	No. previous convictions	.001	> 13.5
11	1,387	71.6%	550	28.4%	1,937	4.0%	No	5	No. previous convictions	.001	<= 5.5
12	212	51.0%	204	49.0%	416	.8%	No	5	No. previous convictions	.001	> 5.5
13	12,605	42.8%	16,865	57.2%	29,470	55.5%	No	6	No. previous convictions	.008	<= 14.5
14	805	24.7%	2,455	75.3%	3,260	5.7%	Yes	6	No. previous convictions	.008	> 14.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Security categories: 1 Category A; 2 Category A; 3 Category B; 4 Category C; 5 YOI closed; 6 YOI open/short; 7 Uncategorised.

Table A15.8: Drug misuse need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.113		100.0%	
Age (years)		.047		41.4%	
No. previous convictions		.041		36.4%	
No. previous custodial sentences		.036		32.0%	
Security category		.024		21.6%	
Sentence length (days)		.018		16.2%	
Ethnicity		.015		13.1%	
Gender		.009		7.8%	

Node	No		Yes		Total	Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%				Variable	Improvement	
0	31,792	57.4%	23604	42.6%	55396	No	0			
1	9,212	39.3%	14,207	60.7%	23,419	Yes	0	Offence category	.099	3; 4; 5; 8
2	22,580	70.6%	9,397	29.4%	31,977	No	0	Offence category	.099	1; 2; 6; 7; 9
3	6,098	34.0%	11,842	66.0%	17,940	Yes	1	Security category	.014	3; 2; 7
4	3,114	56.8%	2,365	43.2%	5,479	No	1	Security category	.014	5; 1; 4; 6
5	18,175	67.4%	8,796	32.6%	26,971	No	2	Age (years)	.016	<= 41.5
6	4,405	88.0%	601	12.0%	5,006	No	2	Age (years)	.016	> 41.5
7	2,646	27.7%	6,923	72.3%	9,569	Yes	3	Sentence length (days)	.005	<= 364.5
8	3,452	41.2%	4,919	58.8%	8,371	Yes	3	Sentence length (days)	.005	> 364.5
9	2,910	55.8%	2,302	44.2%	5,212	Yes	4	Age (years)	.001	<= 37.5
10	204	76.4%	63	23.6%	267	No	4	Age (years)	.001	> 37.5
11	17,382	69.2%	7,748	30.8%	25,130	No	5	No. previous convictions	.009	<= 18.5
12	793	43.1%	1,048	56.9%	1,841	Yes	5	No. previous convictions	.009	> 18.5
13	2,118	82.2%	459	17.8%	2,577	No	6	Age (years)	.002	<= 47.5
14	2,287	94.2%	142	5.8%	2,429	No	6	Age (years)	.002	> 47.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Security categories: 1 Category A; 2 Category A; 3 Category B; 4 Category B; 5 Category C; 6 Category C; 7 Category C; 8 Category C; 9 Category C.

Table A15.9: Alcohol misuse need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.097		100.0%	
Ethnicity		.049		50.9%	
Age (years)		.021		21.3%	
Sentence length (days)		.017		17.9%	
Gender		.011		10.9%	
Security category		.009		9.4%	
No. previous convictions		.005		5.6%	
No. previous custodial sentences		.003		3.4%	

Node	No		Yes		Total	Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%				Variable	Improvement	
0	28,355	51.8%	26,371	48.2%	54,726	No	0			
1	15,844	67.1%	7,778	32.9%	23,622	No	0	Offence category	.070	2; 3; 5; 6; 8
2	12,511	40.2%	18,593	59.8%	31,104	Yes	0	Offence category	.070	1; 4; 7; 9
3	10,845	62.2%	6,595	37.8%	17,440	No	1	Offence category	.011	2; 3; 5
4	4,999	80.9%	1,183	19.1%	6,182	No	1	Offence category	.011	6; 8
5	9,451	35.6%	17,120	64.4%	26,571	Yes	2	Ethnicity	.029	1
6	3,060	67.5%	1,473	32.5%	4,533	No	2	Ethnicity	.029	5; 4; 3; 2
7	2,193	50.9%	2,112	49.1%	4,305	No	3	Age (years)	.005	<= 22.5
8	8,652	65.9%	4,483	34.1%	13,135	No	3	Age (years)	.005	> 22.5
9	3,077	75.6%	993	24.4%	4,070	No	4	Ethnicity	.002	1
10	1,922	91.0%	190	9.0%	2,112	No	4	Ethnicity	.002	5; 4; 3; 2
11	3,149	26.8%	8,594	73.2%	11,743	Yes	5	Offence category	.012	1; 7
12	6,302	42.5%	8,526	57.5%	14,828	Yes	5	Offence category	.012	4; 9
13	435	54.6%	362	45.4%	797	No	6	Ethnicity	.001	4
14	2,625	70.3%	1,111	29.7%	3,736	No	6	Ethnicity	.001	5; 3; 2

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Ethnicity categories: 1 White; 2 Black; 3 Asian; 4 Mixed; 5 Other.

Table A15.10: Emotional wellbeing need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Age (years)		.025		100.0%	
Gender		.022		87.9%	
Ethnicity		.017		66.1%	
Offence category		.016		65.7%	
Security category		.016		65.0%	
Sentence length (days)		.009		36.8%	
No. previous convictions		.005		20.2%	
No. previous custodial sentences		.003		10.5%	

Node	No		Yes		Total	Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%				n	%	
0	30,112	54.2%	25,445	45.8%	55,557	100.0%	No			
1	28,691	56.5%	22,108	43.5%	50,799	91.2%	No	0	Gender	.022
2	1,421	29.9%	3,337	70.1%	4,758	8.8%	Yes	0	Gender	.022
3	13,260	64.8%	7,210	35.2%	20,470	36.4%	No	1	Age (years)	.017
4	15,431	50.9%	14,898	49.1%	30,329	54.8%	Yes	1	Age (years)	.017
5	1,205	27.6%	3,154	72.4%	4,359	8.1%	Yes	2	Ethnicity	.002
6	216	54.1%	183	45.9%	399	.7%	No	2	Ethnicity	.002
7	10,506	61.9%	6,471	38.1%	16,977	30.3%	No	3	Ethnicity	.006
8	2,754	78.8%	739	21.2%	3,493	6.1%	No	3	Ethnicity	.006
9	2,522	68.3%	1,171	31.7%	3,693	6.5%	No	4	Offence category	.009
10	12,909	48.5%	13,727	51.5%	26,636	48.2%	Yes	4	Offence category	.009
11	243	37.7%	402	62.3%	645	1.2%	Yes	5	Offence category	.001
12	962	25.9%	2,752	74.1%	3,714	6.9%	Yes	5	Offence category	.001
13	134	64.4%	74	35.6%	208	.4%	No	6	Offence category	.000
14	82	42.9%	109	57.1%	191	.3%	Yes	6	Offence category	.000

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Ethnicity categories: 1 White; 2 Black; 3 Asian; 4 Mixed; 5 Other. Key to Gender categories: 1 Male; 2 Female.

Table A15.11: Thinking and behaviour need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
Offence category		.058		100.0%	
Security category		.027		46.3%	
Sentence length (days)		.022		38.4%	
Age (years)		.017		28.5%	
No. previous convictions		.015		26.3%	
No. previous custodial sentences		.011		19.5%	
Ethnicity		.010		17.5%	
Gender		.002		3.9%	

Node	No		Yes		Total	Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%				n	%	
0	15,059	27.1%	40,589	72.9%	55,648	100.0%	Yes			
1	3,313	52.6%	2,990	47.4%	6,303	14.0%	No	0	Offence category	.044
2	11,746	23.8%	37,599	76.2%	49,345	86.0%	Yes	0	Offence category	.044
3	3,039	55.7%	2,421	44.3%	5,460	12.4%	No	1	No. previous convictions	.003
4	274	32.5%	569	67.5%	843	1.6%	No	1	No. previous convictions	.003
5	1,212	41.7%	1,694	58.3%	2,906	5.9%	No	2	Security category	.014
6	10,534	22.7%	35,905	77.3%	46,439	80.1%	Yes	2	Security category	.014
7	912	68.9%	411	31.1%	1,323	3.3%	No	3	Offence category	.002
8	2,127	51.4%	2,010	48.6%	4,137	9.1%	No	3	Offence category	.002
9	219	29.7%	518	70.3%	737	1.4%	Yes	4	Age (years)	.000
10	55	51.9%	51	48.1%	106	.2%	No	4	Age (years)	.000
11	1,071	45.1%	1,306	54.9%	2,377	5.0%	No	5	No. previous convictions	.001
12	141	26.7%	388	73.3%	529	.9%	Yes	5	No. previous convictions	.001
13	9,404	23.9%	29,926	76.1%	39,330	68.7%	Yes	6	No. previous convictions	.005
14	1,130	15.9%	5,979	84.1%	7,109	11.5%	Yes	6	No. previous convictions	.005

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Security categories: 1 Category A; 2 Category A; 3 Category B; 4 Category C; 5 YOI closed; 6 YOI open/short; 7 Uncategorised.

Table A15.12: Attitudes need classification decision tree model (prison receptions)

Independent variable		Importance		Normalised importance	
		Importance		Normalised importance	
Offence category		.033		100.0%	
Security category		.023		69.8%	
No. previous convictions		.020		61.9%	
Sentence length (days)		.018		56.0%	
No. previous custodial sentences		.015		46.5%	
Age (years)		.012		35.7%	
Ethnicity		.009		26.0%	
Gender		.001		2.0%	

Node	No		Yes		Total		Predicted category	Parent node	Primary independent variable		Split values
	n	%	n	%	n	%			Variable	Improvement	
0	25,722	46.3%	29,831	53.7%	55,553	100.0%	No	0	Offence category	.021	6; 8
1	4,187	67.2%	2,043	32.8%	6,230	11.9%	No	0	Offence category	.021	1; 2; 3; 4; 5; 7; 9
2	21,535	43.7%	27,788	56.3%	49,323	88.1%	No	1	Sentence length (days)	.002	<= 180.5
3	479	50.3%	473	49.7%	952	1.7%	No	1	Sentence length (days)	.002	> 180.5
4	3,708	70.3%	1,570	29.7%	5,278	10.2%	No	2	Security category	.013	4
5	1,902	65.5%	1,001	34.5%	2,903	5.5%	No	2	Security category	.013	5; 3; 2; 7; 1; 6
6	19,633	42.3%	26,787	57.7%	46,420	82.5%	Yes	3	Offence category	.000	6
7	200	61.3%	126	38.7%	326	.6%	No	3	Offence category	.000	8
8	279	44.6%	347	55.4%	626	1.1%	No	4	No. previous convictions	.002	<= 7.5
9	3,396	72.8%	1,266	27.2%	4,662	9.1%	No	4	No. previous convictions	.002	> 7.5
10	312	50.6%	304	49.4%	616	1.1%	No	5	No. previous convictions	.001	<= 6.5
11	1,647	69.3%	728	30.7%	2,375	4.6%	No	5	No. previous convictions	.001	> 6.5
12	255	48.3%	273	51.7%	528	1.0%	No	6	No. previous convictions	.008	<= 16.5
13	18,096	44.0%	23,000	56.0%	41,096	73.5%	No	6	No. previous convictions	.008	> 16.5
14	1,537	28.9%	3,787	71.1%	5,324	9.1%	Yes	6	No. previous convictions	.008	> 16.5

Key to Offence categories: 1 Violence against the person; 2 Sexual offences; 3 Burglary; 4 Robbery; 5 Theft and handling; 6 Fraud and forgery; 7 Criminal damage; 8 Drug offences; 9 Other offences. Key to Security categories: 1 Category A; 2 Category A; 3 Category B; 4 Category C; 5 YOI closed; 6 YOI open/short; 7 Uncategorised.

Appendix 16: Risk and need residual values and percentages correctly predicted for classification decision tree models (prison receptions)

Commentary regarding the goodness-of-fit/accuracy of the classification models for all the prison receptions is provided in the main body of the report. Table A16.1 compares the model fit/accuracy for those cases which met the post-sentence eligibility for OASys completion and those cases which did not. As shown, there was a greater variance in the percentages correctly predicted in the non-eligible cases, ranging from 58% in the risk of serious harm model to 74% in the education, training and employability model. For some of the criminogenic needs, large over-predictions of criminogenic need rates in the non-eligible cases were offset by large under-predictions of criminogenic need rates in the eligible cases. At the extreme, an over-prediction of 19% for relationships in the non-eligible cases was offset by an under-prediction of 17% in the eligible cases. Such offsetting demonstrates that practitioners were not differing as much in their assessments of criminogenic needs between the non-eligible and eligible cases as suggested by the basic offender characteristic variables. For relationships, the predicted levels of need in the non-eligible and eligible cases were 75% and 30% respectively, whereas the actual recorded levels were much closer at 56% and 46%.

Table A16.1: Goodness-of-fit and accuracy of classification models by OASys post-sentence eligibility (prison receptions)

Risk and need level	OASys post-sentence eligibility					
	Non-eligible (n=24,040)		Eligible (n=27,285)		All cases (n=51,325)	
	Residual (actual minus predicted rate)	Per cent correctly predicted	Residual (actual minus predicted rate)	Per cent correctly predicted	Residual (actual minus predicted rate)	Per cent correctly predicted
Likelihood of reconviction:						
Low	2.3%	61.1%	-2.3%	58.7%	-0.1%	59.8%
Medium	-0.7%		1.1%		0.2%	
High	-1.6%		1.2%		-0.1%	
Risk of serious harm:						
Low	-6.4%	58.5%	3.9%	63.6%	-0.9%	61.2%
Medium	2.6%		-1.6%		0.4%	
High/Very high	3.8%		-2.2%		0.6%	
Criminogenic need:						
Accommodation	-7.1%	61.2%	7.1%	61.4%	0.5%	61.3%
ETE	-8.5%	73.6%	4.7%	70.5%	-1.5%	72.0%
Finance	0.4%	67.6%	-2.7%	66.6%	-1.2%	67.1%
Relationships	-18.8%	62.9%	16.6%	62.9%	0.0%	62.9%
Lifestyle & associates	-0.3%	66.9%	1.0%	68.0%	0.4%	67.5%
Drug misuse	1.1%	71.3%	3.2%	68.8%	2.2%	70.0%
Alcohol misuse	1.3%	65.7%	-2.6%	69.2%	-0.8%	67.5%
Emotional wellbeing	-11.4%	60.5%	10.2%	64.6%	0.1%	62.7%
Thinking & behaviour	-4.2%	72.7%	4.0%	71.1%	0.1%	71.8%
Attitudes	-11.3%	61.9%	8.8%	63.6%	-0.6%	62.8%

Appendix 17: Example of the relationships textual analysis matrix

Respondent's Assessment ID	3112133
Demographic & Offence information	Male, White, Aged 24; Suspended sentence
Section score, Assessor judgment	Need but no link
Risk of reconviction & of serious harm	Recon: Medium Harm: Medium
Current relationship (family members)	Strained relationship with mother; No contact with father and siblings
Criminal record (family member/s)	No record
Childhood Experience	Had a 'happy' childhood
Current relationship (partner)	In a relationship with partner for over two years; Frequent arguments about his drinking and being unfaithful; Willing to seek counselling
Criminal record (partner)	Partner on unpaid work order
Past experience of relationships	No record of previous relationships
Domestic violence	Denies any incidence relating to DV
Other (additional themes)	None
Respondent's Assessment ID	3113886
Demographic & Offence information	Female, White Aged 27; Community sentence
Section score, Assessor judgment	Need but no link
Risk of reconviction & of serious harm	Recon: Medium Harm: Medium
Current relationship (family members)	Good relationship with mother and one of her sisters
Criminal record (family member/s)	Brother has criminal conviction
Childhood Experience	In foster care between the ages of 13 to 15
Current relationship (partner)	Supportive partner; Good influence on her; Offending has reduced; Partner also has drug problem
Criminal record (partner)	Partner remanded in custody
Past experience of relationships	No other record other than incidences of DV
Domestic violence	Has been a victim of DV in previous relationships
Other (additional themes)	Has a son aged 4; Wants to give up drugs and offending because of him

Appendix 18: Example of the lifestyle and associates matrix

Respondent's Assessment ID	3321197
Demographic & Offence information	Male, White, aged 21, DTTO, Violence offence
Section score, Assessor judgment	Need but no link
Risk of reconviction & of serious harm	Recon: Medium Harm: High
Community integration	Used to play rugby; Plays in a local pool team
Activities encourage offending	Spends most of his time gambling with boss or at the pub with friends; Funds his lifestyle through defrauding shops
Easily influenced by criminal associates	Does not feel that he is influenced by others
Manipulative/predatory lifestyle	Enjoys defrauding shops;
Recklessness/risk-taking behaviour	Demonstrates risk taking behaviour in gambling to release boredom; Enjoys not getting caught when defrauding shops;
Other (additional themes)	None
Respondent's Assessment ID	3331647
Demographic & Offence information	Female, White, aged 34; DTTO, Drugs
Section score, Assessor judgment	Need but no link
Risk of reconviction & of serious harm	Recon: Medium Harm: Medium
Community integration	No record
Activities encourage offending	No record
Easily influenced by criminal associates	Still associates with negative peers and is influenced by them
Manipulative/predatory lifestyle	No record
Recklessness/risk-taking behaviour	No record
Other (additional themes)	None

Ministry of Justice Research Series 16/09

A compendium of research and analysis on the Offender Assessment System (OASys) 2006-2009

The Offender Assessment System (OASys) is a national risk/need assessment tool used by the prison and probation services in England and Wales. The tool combines actuarial methods of prediction with structured professional judgement to provide standardised assessments of offenders' risks and needs, as well as linking these risks and needs to individualised sentence plans and risk management plans. OASys data is collated centrally within the OASys Data Evaluation and Analysis Team (O-DEAT) database, and this report presents OASys research and analysis conducted by O-DEAT from 2006 to 2009. The report demonstrates the extent to which OASys meets the criteria for good offender assessment. It begins with an introduction to OASys, describing its development, content and uses, followed by a chapter updating the literature underpinning its development and several chapters presenting findings on coverage, completion and various aspects of reliability and validity. It then includes findings from exploratory qualitative analysis of OASys content, followed by a presentation of offender profiles using OASys data, and concludes with information about other research and usage of OASys, and the actions which followed the O-DEAT research findings. Importantly, the research has contributed to important modifications to OASys, streamlining its content through the introduction of layered OASys while improving its measurement of offender risk and need. Further improvements will be sought through continuing analysis of OASys data, with current projects focusing upon specific types of reoffending and the practitioners' risk of serious harm ratings.

ISBN 978 1 84099 328 8

© Crown copyright

Produced by the Ministry of Justice

Alternative format versions of this report are available on request.

E-mail: research@justice.gsi.gov.uk

<http://www.justice.gov.uk/publications/research.htm>