A systematic review of outcome measures used in forensic mental health research with consensus panel opinion


March 2010
10.3310/hta14180
How to obtain copies of this and other HTA programme reports
An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public and private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:
– fax (with credit card details)
– post (with credit card details or cheque)
– phone during office hours (credit card only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:
Synergie UK (HTA Department) Email: orders@hta.ac.uk
Digital House, The Loddon Centre Tel: 0845 812 4000 – ask for ‘HTA Payment Services’
Wade Road (out-of-hours answer-phone service)
Basingstoke
Hants RG24 8QW Fax: 0845 812 4001 – put ‘HTA Order’ on the fax header

Payment methods
Paying by cheque
If you pay by cheque, the cheque must be in pounds sterling, made payable to University of Southampton and drawn on a bank with a UK address.

Paying by credit card
You can order using your credit card by phone, fax or post.

Subscriptions
NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of HTA on DVD?
Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). HTA on DVD is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.
A systematic review of outcome measures used in forensic mental health research with consensus panel opinion


1Department of Public Health, University of Oxford, Oxford, UK
2Department of Psychiatry, University of Oxford, Oxford, UK
3Personal Social Services Research Unit, London School of Economics, London, UK
4Division of Psychological Medicine and Psychiatry, Institute of Psychiatry, King’s College London, London, UK

*Corresponding author

Declared competing interests of authors: none

Published March 2010
10.3310/hta14180

This report should be referenced as follows:

Health Technology Assessment is indexed and abstracted in Index Medicus/MEDLINE, Excerpta Medical/EMBASE, Science Citation Index Expanded (SciSearch®) and Current Contents®/Clinical Medicine.
The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. ‘Health technologies’ are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the ‘National Knowledge Service’.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

### Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed ‘systematic’ when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this issue of the journal was commissioned by the National Coordinating Centre for Research Methodology (NCCRM), and was formally transferred to the HTA programme in April 2007 under the newly established NIHR Methodology Panel. The HTA programme project number is 06/91/11. The contractual start date was in January 2006. The draft report began editorial review in July 2009 and was accepted for publication in July 2009. The commissioning brief was devised by the NCCRM who specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors’ report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

**Editor-in-Chief:** Professor Tom Walley CBE  
**Series Editors:** Dr Martin Ashton-Key, Dr Aileen Clarke, Professor Chris Hyde, Dr Tom Marshall, Dr John Powell, Dr Rob Riemsma and Professor Ken Stein

**ISSN 1366-5278**

© 2010 Queen’s Printer and Controller of HMSO

This journal may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to: NETSCC, Health Technology Assessment, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.

Printed on acid-free paper in the UK by the Charlesworth Group.
According to the consensus group, many domains beyond recidivism and mental health were important but under-represented in the review of outcomes. Current instruments that may show future promise in outcome measurement included risk assessment tools. The outcome measure of repeat offending behaviour was by far the most frequently used, occurring in 72% of the studies included in the review. Its measurement varied with position in the criminal justice system, offence specification and method of measurement. The consensus group believed that recidivism is only an indication of the amount of antisocial acts that are committed.

Conclusions: A wide range of domains are relevant to assessing outcomes of interventions in forensic mental health services. Evaluations need to take account of public safety, but also clinical, rehabilitation and humanitarian outcomes. Recidivism is a very high priority; the public expects interventions that will reduce future criminal behaviour. Greater attention needs to be given to validity of measurement, given the enormous variety of approaches to measurement. More research is needed on methods to take account of the heterogeneity of seriousness of forms of recidivism in outcome measurement. Validity of self-report instruments regarding recidivism also needs examination by further research. Mental health is clearly also an important dimension of outcome. The review provides clear support for the view that domains such as quality of life, social function and psychosocial adjustment have not been extensively employed in forensic mental health research, but are relevant and important issues. The role of such instruments needs more consideration.

Abstract

A systematic review of outcome measures used in forensic mental health research with consensus panel opinion

R Fitzpatrick,1* J Chambers,1 T Burns,2 H Doll,1 S Fazel,2 C Jenkinson,1 A Kaur,1 M Knapp,3 L Sutton2 and J Yiend4

1Department of Public Health, University of Oxford, Oxford, UK
2Department of Psychiatry, University of Oxford, Oxford, UK
3Personal Social Services Research Unit, London School of Economics, London, UK
4Division of Psychological Medicine and Psychiatry, Institute of Psychiatry, King’s College London, London, UK

*Corresponding author

Objective: To describe and assess outcome measures in forensic mental health research, through a structured review and a consensus panel. Data sources: A search of eight electronic databases, including CINAHL, EMBASE and MEDLINE, was conducted for the period 1990–2006. Review methods: In the structured review, search and medical subject heading terms focused upon two factors: the use of a forensic participant sample and the experimental designs likely to be used for outcome measurement. Data extraction included general information about the identity of the reference, specific information regarding the study and information pertaining to the outcome measures used. The consensus exercise was implemented in two stages. At the first stage, participants were asked to complete ratings about the importance of various potential areas of outcome measurement in a written consultation. At the second stage, they were asked to attend a consensus meeting to review and agree results relating to the domains, to consider and rate specific outcome instruments identified as commonly used from the structured review and to discuss strengths, weaknesses and future priorities for outcome measurement in forensic mental health research. Results: The final sample of eligible studies for inclusion in the review consisted of 308 separate studies obtained from 302 references. The consensus group agreed on 11 domains of forensic mental health outcome measurement, all of which were considered important. Nine different outcome measure instruments were used in more than four different studies. The most frequently used outcome measure was used in 15 studies.
Contents

<table>
<thead>
<tr>
<th>List of abbreviations</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive summary</td>
<td>ix</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Methods</td>
<td>3</td>
</tr>
<tr>
<td>Research objectives</td>
<td>3</td>
</tr>
<tr>
<td>Strategy</td>
<td>3</td>
</tr>
<tr>
<td>Stage 1: structured</td>
<td>3</td>
</tr>
<tr>
<td>View</td>
<td>5</td>
</tr>
<tr>
<td>3 Results</td>
<td>9</td>
</tr>
<tr>
<td>Analysis of the robustness of the results</td>
<td>9</td>
</tr>
<tr>
<td>Properties of the 308 studies included in the structured review</td>
<td>9</td>
</tr>
<tr>
<td>Outcome measurement</td>
<td>13</td>
</tr>
<tr>
<td>Addiction Severity</td>
<td>17</td>
</tr>
<tr>
<td>Index</td>
<td>18</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>19</td>
</tr>
<tr>
<td>Brief Psychiatric</td>
<td>19</td>
</tr>
<tr>
<td>Rating Scale</td>
<td>21</td>
</tr>
<tr>
<td>Child Behavior</td>
<td>21</td>
</tr>
<tr>
<td>Checklist</td>
<td></td>
</tr>
<tr>
<td>Conflict Tactics</td>
<td>22</td>
</tr>
<tr>
<td>Scale</td>
<td>22</td>
</tr>
<tr>
<td>Family Adaptability</td>
<td>23</td>
</tr>
<tr>
<td>and Cohesion Evaluation Scale</td>
<td>23</td>
</tr>
<tr>
<td>Revised Behavior</td>
<td>24</td>
</tr>
<tr>
<td>Problem Checklist</td>
<td>24</td>
</tr>
<tr>
<td>Symptom-Checklist-90-</td>
<td>25</td>
</tr>
<tr>
<td>Revised</td>
<td>25</td>
</tr>
<tr>
<td>Self-Reported</td>
<td>27</td>
</tr>
<tr>
<td>Delinquency Scale</td>
<td>27</td>
</tr>
<tr>
<td>Summary of most frequently used outcome measures</td>
<td>28</td>
</tr>
<tr>
<td>Consensus opinion about the most frequent outcomes</td>
<td>28</td>
</tr>
<tr>
<td>The outcome measure of recidivism</td>
<td>29</td>
</tr>
<tr>
<td>4 Discussion</td>
<td>33</td>
</tr>
<tr>
<td>5 Conclusions and recommendations</td>
<td>37</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>39</td>
</tr>
<tr>
<td>References</td>
<td>41</td>
</tr>
<tr>
<td>Appendix 1 Search strategies by database</td>
<td>47</td>
</tr>
<tr>
<td>Appendix 2 Data extraction form</td>
<td>51</td>
</tr>
<tr>
<td>Appendix 3 Included studies reference list</td>
<td>57</td>
</tr>
<tr>
<td>Health Technology Assessment reports published to date</td>
<td>69</td>
</tr>
<tr>
<td>Health Technology Assessment programme</td>
<td>91</td>
</tr>
</tbody>
</table>
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASI</td>
<td>Addiction Severity Index</td>
</tr>
<tr>
<td>BDI</td>
<td>Beck Depression Inventory</td>
</tr>
<tr>
<td>BPRS</td>
<td>Brief Psychiatric Rating Scale</td>
</tr>
<tr>
<td>CBCL</td>
<td>Child Behavior Checklist</td>
</tr>
<tr>
<td>CTS</td>
<td>Conflict Tactics Scale</td>
</tr>
<tr>
<td>DSM</td>
<td>Diagnostic and Statistical Manual of Mental Disorders</td>
</tr>
<tr>
<td>FACES</td>
<td>Family Adaptability and Cohesion Evaluation Scales</td>
</tr>
<tr>
<td>HCR-20</td>
<td>Historical, Clinical and Risk Management Scales</td>
</tr>
<tr>
<td>MeSH</td>
<td>medical subject heading</td>
</tr>
<tr>
<td>MMPI</td>
<td>Minnesota Multiphasic Personality Inventory</td>
</tr>
<tr>
<td>RBPC</td>
<td>Revised Behavior Problem Checklist</td>
</tr>
<tr>
<td>RCT</td>
<td>randomised controlled trial</td>
</tr>
<tr>
<td>SCL-90-R</td>
<td>Symptom-Checklist-90-Revised</td>
</tr>
<tr>
<td>SRDS</td>
<td>Self-Reported Delinquency Scale</td>
</tr>
</tbody>
</table>

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.
Introduction

This study examines outcome measurement in forensic mental health research. Forensic mental health services cover many domains such as prisons, community corrections and secure forensic hospitals. Within this complex system each service uses outcome measures for its own specific objectives, with little standardisation between organisations. Outcome measurement is also difficult to standardise as the client population often suffers from multiple problems including mental health disorders and substance abuse, leading to multiple targets for intervention. Research in forensic mental health also suffers problems such as a moving population, priorities to maintain security and duty of care to clients, preventing extensive application of studies based on randomised controlled trials. Overall, there is very little methodological discussion about outcome measurement in forensic mental health research.

Objectives

This research project was exploratory in nature, to describe and assess outcome measures in forensic mental health research. A two-stage study was conducted. Stage one consisted of a structured review of outcome measures used in forensic mental health research. Stage two consisted of a consensus panel that considered the essential domains of outcome measurement in forensic mental health and then assessed the properties of the most frequently used outcome measures against key questions. The panel included experts from within forensic mental health research and services.

Methods

Structured review

A search of eight electronic databases, including CINAHL, EMBASE and MEDLINE, was conducted for the period 1990–2006. Search and medical subject heading terms focused upon two factors: the use of a forensic participant sample and the experimental designs likely to be used for outcome measurement.

Studies eligible for inclusion in the review fulfilled several criteria:

1. Participants were defined as offenders or residents of a forensic mental health institution.
2. The study required examination of an intervention with the use of outcome measurement after the intervention.
3. Study design was required to be either a randomised controlled trial or a quasi-experimental (comparing intervention and control) design with a minimum follow-up of 6 months.
4. A mental health element needed to be present in the participant population, the intervention or an outcome measurement.

Data extraction included general information about the identity of the reference, specific information regarding the study and information pertaining to the outcome measures used.

Data about mental health research outcome measures were extracted from the included references, and were entered into Excel. The outcome measures that occurred most frequently were also calculated.

Consensus group

The consensus exercise was implemented in two stages. At the first stage, participants were asked to complete ratings about the importance of various potential areas of outcome measurement (‘domains’) in a written consultation. At the second stage, they were asked to attend a consensus meeting to review and agree results relating to the domains, to consider and rate specific outcome instruments identified as commonly used from the structured review and to discuss strengths, weaknesses and future priorities for outcome measurement in forensic mental health research. Participants comprised three representatives from psychology, three from psychiatry and one from each of criminology, probation, prison health and nursing.
Results

The final sample of eligible studies for inclusion in the review consisted of 308 separate studies obtained from 302 references.

The consensus group agreed on 11 domains of forensic mental health outcome measurement, all of which were considered important. In the literature review, 1038 distinct variables were identified that were used to assess outcomes in the sample of evaluative studies. Nine different outcome measure instruments were used in more than four different studies. The most frequently used outcome measure was used in 15 studies. A further review of research concerning the psychometric properties of these instruments was carried out. It revealed little evidence specifically to validate their use with forensic populations. The measures that were rated most favourably by the consensus panel were the Beck Depression Inventory, the Brief Psychiatric Rating Scale and the Symptom Checklist-90-Revised. According to the consensus group, many domains beyond recidivism and mental health were important but under-represented in the review of outcomes. Current instruments that may show future promise in outcome measurement included risk assessment tools.

The outcome measure of repeat offending behaviour was by far the most frequently used, occurring in 72% of the studies included in the review. Its measurement varied with position in the criminal justice system, offence specification and method of measurement. The consensus group believed that recidivism is only an indication of the amount of antisocial acts that are committed.

Conclusions and recommendations

A wide range of domains are relevant to assessing outcomes of interventions in forensic mental health services. Evaluations need to take account of public safety, but also clinical, rehabilitation and humanitarian outcomes. To date, research has focused extensively on the first domain, evaluating outcomes in terms of recidivism.

Recidivism is a very high priority; the public expects interventions that will reduce future criminal behaviour. Greater attention needs to be given to validity of measurement, given the enormous variety of approaches to measurement. More research is needed on methods to take account of the heterogeneity of seriousness of forms of recidivism in outcome measurement. Validity of self-report instruments regarding recidivism also needs examination by further research.

Mental health is clearly also an important dimension of outcome. Instruments have been used in forensic mental health research that have been well validated in the context of general mental health research.

The review provides clear support for the view that domains such as quality of life, social function and psychosocial adjustment have not been extensively employed in forensic mental health research, but are relevant and important issues. The role of such instruments needs more consideration. Research is needed in these domains to complement the evidence base of outcomes in terms of public safety and mental health.

The wide array and diversity of measures used in forensic mental health research suggests that there is still substantial scope for standardisation, by further use of consensus-type processes to identify domains and specific measures that are relevant to and familiar in practice and can be more widely used in evaluative research.

The role of instruments assessing dynamic aspects of the risk of violence offer a particular opportunity. They are becoming more widely known in practice and could be more widely used in evaluative research in this field.
The purpose of this study was to examine and assess the range of outcome measures used in forensic mental health research. Currently, there is little agreement about which outcome measures to use in this context. This may reflect the diversity of forensic mental health services as well as reflecting the very broad range of problems experienced by users. Forensic mental health services are very varied, operating in settings as diverse as probation services in the community system and secure forensic hospitals. Services may also be assessed in terms of very diverse goals including clinical, humanitarian, legal and public safety. The specific assessments used by different agencies are also not routine, with agency-specific requirements dictating the use of measures for relevant outcomes. This diversity across agencies means that a standard battery of outcome measures has not developed. In addition, forensic mental health service clients present with multiple problems. For example, personality disorder, mental illness, learning disability, substance abuse and offending behaviour are a few of the possible problems, often occurring together, leading to numerous intervention targets and consequently many combinations of potentially relevant outcomes.

It is not only the large variety of services that forensic mental health encompasses that has affected the standardisation of outcome measurement; there are also significant difficulties inherent in researching forensic mental health populations. Research in forensic mental health has suffered logistical problems, with users often moving through different custodial settings (e.g. remand – prison – probation). Security considerations may have priority over research needs. Practical difficulties with researching forensic populations may be partly responsible for the relative lack of randomised controlled trials (RCTs) in the UK.2

The problems noted above may result in a lack of clear consensus about outcome measures for use in evaluating interventions. A concern expressed in the broader field of mental health research has been that too many different outcome measures have been introduced, with too few receiving proper evaluation, leading to the use of unvalidated outcome measurement.5,6 It has been suggested that, in the broader field of research in mental health, even if time and effort is invested to produce studies of robust design, unvalidated outcome measurement can weaken the value of results.6

A prime example of the difficulties of assessing outcome in forensic mental health research is the commonly used assessment of recidivism. A meta-analysis of recidivism in sexual offenders showed that not only were several different measures used (reconviction, arrest, self-report, parole violation), but that they came from several different sources (national criminal justice records, local records, records from treatment programmes and self-report).7 The diversity of sources for assessing recidivism makes standardisation difficult.8 A method used to increase the validity of reporting recidivism is multisource recording. For example, The MacArthur Violence Risk Assessment Study measured violence from three sources: self-report, collateral informant report and official agency records.9 Also, careful development of self-report instruments can lead to high concurrent validity with court records.10 Thus, whilst the outcome measurement of repeat offending is fraught with problems for valid measurement between different studies, there are strategies that may be employed to make measurement more robust.

The aim of the current study was to conduct a structured review of forensic mental health outcome measures, to identify and, where possible, assess more frequently used outcome measures. This review considered studies that have assessed outcomes within an evaluative study design such as RCTs, or comparative studies with longitudinal observation of groups and a reasonable follow-up period (discussed below).

Elsewhere in health-care research, the body of information about the use of outcome measurement is burgeoning. A good example of a field in which consensus regarding outcome measurement in research and practice has emerged is in relation to musculo-skeletal disorders where a concerted international consensus process, OMERACT (Outcome Measures in Rheumatology),
has delineated methods for agreement about core measures and use of outcomes in practice as well as in research. This consensus process involves:

1. harnessing expert views
2. agreeing key domains of outcome and criteria for assessing evidence
3. applying criteria in systematic reviews
4. identifying sets of preferred outcome measures relevant to a range of applications
5. an ongoing programme of work to test measurement outcome.

OMERACT has focused on three basic criteria for outcome measures: validity, discrimination and feasibility, using consensus processes to interpret emerging research evidence. On a more modest scale, the current study includes the use of consensus processes to launch a process towards consensus in outcome measurement in forensic mental health research. In this study expert opinion about the domains of outcome measurement in forensic mental health were sought. This process was then complemented through assessment of the most frequently used outcome measures from forensic mental health research, gleaned from the structured review, for the criteria of psychometric properties, feasibility and relevance. Thus, we provided expert opinion about all of the different areas of outcome measurement that should be fulfilled and opinion about the measures that have most frequently been used.

In summary, forensic mental health outcome measurement has suffered many barriers to identification of optimal outcome measurement in research, including multi-agency involvement, difficulties for conducting research in a context that may have to prioritise security, and a proliferation of unvalidated measures. In the current study we assessed outcome measures that have been used in previous research in a structured review. Literature pertaining to the most frequently used outcome measures was gathered for information regarding their psychometric properties. We accessed expert opinion about the domains that are pertinent to forensic mental health that require validated outcome measurement. These domains were then assessed according to the outcome measures used in previous research from the studies found in the structured review. Finally, the experts provided opinion about the most frequently used outcome measures from the structured review to assess the validity, feasibility and relevance of measures that have been used in previous research. This assessment of previous use of outcome measurement led to discussion about priorities for future research.
Research objectives and overview of strategy

As has been discussed, this research project was exploratory in nature: (1) to identify outcome measures in use in forensic mental health research, (2) to explore and agree domains of outcome relevant to research in forensic mental health and (3) to assess outcome measures in terms of available evidence and consensus views. To fulfil these aims, a two-stage study was conducted. Stage one consisted of a structured review of outcome measures used in forensic mental health research. Stage two consisted of a consensus panel that considered the essential domains of outcome measurement in forensic mental health and then assessed the properties of the most frequently used outcome measures against key questions. The panel included experts from within forensic mental health research and services. In practice, there was overlap in the timing of aspects of the two stages, with some interdependency of work. For example, practicalities of a 1-year project meant that the consensus panel had to meet before all results of the systematic review were completed. Conversely, an initial consultation with the panel by correspondence provided a classification of domains that was helpful to both the literature review and the consensus meeting. As a result, in parts of this report, description of results moves back and forth between the two pieces of work.

Stage 1: structured review

The methods used in the structured review were considered at length by both the researchers who worked on the project on a day-to-day basis and also by the research team who met on a monthly basis. This process enabled researchers with experience in database searches and those with forensic psychology and psychiatry expertise to contribute to the methodology.

Search inclusion

The structured review was conducted to capture publications from within a set time-frame: 1990–2006 inclusive. The cut-off date for inclusion of emerging articles published during 2006 was November 2006. Only published references were included in the review to ensure that some level of peer review had been undertaken and that studies were available in the public domain; this excluded dissertations.

The databases searched were:

- CINAHL (Cumulative Index to Nursing and Allied Health Literature)
- EMBASE
- MEDLINE (1990 to October 2006)
- National Criminal Justice Reference Service (NCJRS)
- PsycINFO
- Sociological Abstracts
- The Cochrane Database
- The Patient-reported Health Instruments (PHI) website.

Search results from each of the databases were amalgamated into the reference software program Reference Manager. Once all of the references from each of the databases had been uploaded into Reference Manager, a duplicate search was conducted. A duplicate search is necessary as many of the different databases reference the same articles when searched using similar criteria. Once a database that consisted of unique references was constructed, examination of the abstracts began for identification of eligible references for the review. Abstracts were identified as eligible according to the parameters described below. Those abstracts that appeared eligible for the review were marked for collection of a hard copy of the reference. The reference hard copies underwent a final more thorough eligibility analysis and, if eligible, underwent data extraction.

Search strategy

The purpose of this review was to collect data about the most frequently used outcome measures in forensic mental health research. In essence, the purpose of an outcome measure is to measure change after an intervention. The search terms used to fulfil the aim of this project were focused upon two different factors: the use of a forensic
Methods

mental health sample and study designs likely to be used for outcome measurement. Both keyword dictionaries and Medical Subject Heading terms (MeSH) were utilised in the search strategy. The MeSH terms are a pre-designated topic classification system applied to all papers included in each database. Use of MeSH terms for each of the databases widens the scope of only a keyword search by considering categorised references. The keyword search that we employed assessed the presence of a forensic sample population and particular experimental designs.

Forensic participant population
The keyword search strategy was constructed to identify participants involved within the criminal justice system through both the terms used to identify an offender and institutions they might be detained. An earlier strategy identifying the offence types that the participants might have committed (e.g. burglary, rape) was rejected because of a substantial number of irrelevant results. Use of offender and institution terms identified the participant sample descriptions in the abstracts of the references. Search terms included community corrections, parolee, probationer, prisoner, youth custody and forensic unit/hospital (see Appendix 1 for a full listing of search keywords and MeSH terms utilised for each database). By considering the forensic element of sample populations for the search strategy, the researchers were later able to identify specific mental health elements within the population, intervention or outcome measure from the collated reference abstracts.

Experimental design
The search strategy was designed to capture experimental (randomised) or quasi-experimental study designs. The experimental design search included terms such as repeated measures, randomised and longitudinal (see Appendix 1 for a full listing of search keywords utilised for each database).

Studies included in the review
Identification and assessment of studies for inclusion occurred at two stages: first, using the abstracts recovered from the databases; second, with the hard copy of the reference. This two-stage process enabled studies that appeared to be relevant to undergo stringent assessment of eligibility, discarding clearly irrelevant studies at an early stage. There were several elements that a reference had to fulfil to be eligible for inclusion. Firstly, participants in studies must be offenders, defined as: any individual under the supervision of the criminal justice system including community correction clients, parolees, correctional clients, probationers or youthful offender system residents. In addition, references where the participants resided in a forensic mental health institution were also included as they are highly pertinent to forensic mental health research. Residence in a forensic mental health institution did not necessarily mean that the participant had been convicted. They may have been unfit to stand trial, or were sectioned into a forensic mental health institution owing to dangerousness.

Second, the study was required to investigate an intervention, enabling the measurement of change, thus the use of outcome measures. The intervention was required to be an experimental activity that would not usually fall within day-to-day activities. This criterion omitted longitudinal and cohort studies that predicted offending behaviour due to naturalistic social variables such as level of peer/parental support or socioeconomic status. Another essential element of outcome measurement is assessment of change measured both before and after the intervention. Thus, repeated measures were a requirement of study eligibility. This requirement was subsequently relaxed for some variables, assessing outcomes in a retrospective fashion only, e.g. measures of service experience or satisfaction.

The researchers wished to include studies with a robust design. The ‘gold standard’ of experimental design is an RCT. Therefore all studies that were RCTs were included. An RCT was identified by the participants being randomly selected into either the experimental or a control condition. The control condition may have been either another intervention or a control group (i.e. placebo/waitlist). However, it was felt that exclusive focus on RCTs might be too restrictive of the range of outcome measures in common use, given the under-developed nature of experimental research in some areas of forensic mental health. Hence studies with quasi-experimental and similar controlled study designs were included. A requirement was stipulated that quasi-experimental studies had a minimum follow-up of 6 months with longitudinal measurement.

The element of mental health
The final criterion for study eligibility was the presence of mental health as a key issue. The search strategy was focused on forensic participants.
only, with the element of mental health being too theoretically complex to isolate at the search level. The focus on a forensic sample meant that references were excluded that did not possess a mental health element, for example, a sample of prisoners completing a horticultural course in an effort to reduce recidivism. The mental health element was required to be present in one of three ways: (1) the participating study sample or institution, (2) the intervention or (3) the outcome measure. Occurrence of mental health was required in only one of these three for the study to be included, but usually it occurred in more than one. For mental health to be present in the participant sample or institution, the sample required a mental health theme such as diagnosis of a psychiatric or personality disorder, or subjects to be residing in or attending a forensic mental health institution. For forensic mental health to be present in the intervention it must target either a psychological or a psychiatric mechanism; for example, cognitive behavioural therapy to reduce violent behaviour and anger would fulfil this requirement. Finally, for mental health to be present in the outcome measure it must measure a mental health element, such as depression.

Studies excluded from the review

Several types of study were assessed as not eligible for inclusion into the review. There are many areas of forensic mental health research that do not explicitly concern outcome measurement. These alternative areas of research include reduction of risk in populations that would be prone to offending through resilience factors; the assessment of risk within an offender population; and victimology. These types of study also did not tend to use repeated measures or to include identified interventions, thus were often ineligible for several reasons.

The inclusion of references that used only an offender sample meant that victim samples and ‘at risk’ samples that had not yet offended were excluded. For example, a study may examine a community intervention for drug addicts as they are at high risk of offending, and the outcome measure may be offending behaviour. This study would not be included as the participants were not offenders at recruitment into the project. If a reference did report a non-offender sample in addition to an offender sample, then the data were required to be distinctly separate for the two different populations, with only the offender sample included for data extraction.

Interventions that were solely focused on physical health issues were also excluded. Thus, interventions such as those designed to reduce risky behaviours that might lead to contraction of AIDS or hepatitis were excluded if they did not target mental health as a component.

Data extraction

Once a reference had undergone all of the eligibility checks and was considered eligible for inclusion into the structured review, data were extracted. Data extraction included general information about the identity of the reference, specific information regarding the study and information pertaining to the outcome measures used. Reference identity data included the type of report, name of the author, year of publication and country of origin. Information regarding the study included type of study design, study setting, sample size, age of participants, participant psychiatric diagnosis, participant learning disability diagnosis and type of intervention. Outcome measure data consisted of the name of each outcome measure, and the longest follow-up period for outcome measurement for each measure. If the outcome measure was recidivism or criminal behaviour then additional information in the form of the type of recidivism (i.e. arrest, charges, conviction, etc.) was extracted (see Appendix 2 for the data extraction form).

Stage 2: consensus group methodology

Overview

The consensus exercise was implemented in two stages. At the first stage, participants were asked to complete ratings about the importance of various potential areas of outcome measurement (‘domains’). This first stage was carried out by written correspondence. At the second stage, they were asked to attend a consensus meeting to review and agree results relating to the domains, to consider and rate specific outcome instruments identified as commonly used from the structured review, and to discuss strengths, weaknesses and future priorities for outcome measurement in forensic mental health research.
Participants

The target sample was weighted to reflect the relative contributions of different professions within forensic mental health. It comprised three representatives from psychology, three from psychiatry and one from each of the fields of criminology, probation, prison health and nursing. A target sample of three stakeholders or service users was set for participation in the domain rating phase only. Final participant numbers for each part of the exercise are shown in Table 1. The group comprised a mix of senior and experienced practitioners, academics and researchers. Examples of participants' backgrounds and experience included a senior mental health nurse with experience of training and research in risk assessment prediction instruments in forensic contexts; a professor of psychiatry with extensive research experience in interventions and outcomes for dangerous personality disorder in a range of secure settings; a consultant clinical psychologist practising in secure settings with extensive experience of services for men and women with records of violent behaviour; and a professor of forensic clinical psychology with broad research experience of offender behaviour programmes.

Procedure

Domain ratings

Participants were sent a rating form with an inclusive list of domains and asked to rate their relative importance for forensic mental health research. Potential domains for rating were identified based on the results of the structured review. Ratings were made on a Likert scale comprising 1 (not important), 2 (less important), 3 (important), 4 (very important) and 5 (essential). ‘Important’ was specified as ‘how important is it that this domain of measurement is included in future research?’. Participants were asked to list using free text any areas they considered important that had been omitted from the list. Replies were collated and results summarised for presentation at a consensus meeting.

Consensus meeting

For reasons of practicality, some attendees were recruited after the domain rating exercise so that not all experts participated in both stages of the process.

The meeting lasted for a whole working day. It was structured into three discrete sections. First, participants were shown the mean ratings and final rank order of domains obtained at the first stage from written correspondence. This was followed by a period of unstructured discussion of these results.

Next, individual instruments were presented in turn for rating, selected on the basis of their frequency of occurrence (five or more occurrences) within the structured literature review of forensic mental outcome assessment conducted in stage one of the project. The content, format and supporting evidence for each instrument were briefly presented to the group before encouraging participants to consult a more detailed written summary prepared by the team and provided before the meeting. Copies of instruments themselves were made available during the meeting. Each instrument was rated by members of the meeting using a booklet constructed for the purpose and containing a brief synopsis of the available psychometric data.

TABLE 1 Numbers of individuals participating in the consensus exercise

<table>
<thead>
<tr>
<th>Profession</th>
<th>Target number</th>
<th>Number approached</th>
<th>Number completing domain ratings</th>
<th>Number attending consensus meeting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychiatry</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Psychology</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Prison health</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nursing</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Probation</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Criminology</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Stakeholders</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>26</td>
<td>13</td>
<td>8</td>
</tr>
</tbody>
</table>
Participants were asked to identify whether each instrument was familiar to them (yes/no) and then to rate it on three scales:

1. relevance and appropriateness
2. feasibility of use
3. adequacy of measurement properties (e.g., reliability, validity, responsiveness and based on the information provided at the meeting).

All scales comprised 1 (not), 2 (slightly), 3 (fairly), 4 (very) and 5 (extremely).

Finally, the group participated in an unstructured discussion concerning recidivism. Participants were subsequently sent a draft report to confirm whether views had been adequately captured.
Chapter 3

Results

Analysis of the robustness of the results

Figure 1 displays the generation of the eligible studies included in the structured review. As described in the method section there were two levels of inspection of the reference list generated from the electronic database search, including examination of the abstracts and then the hard copies. After duplicate reference removal there were 10,703 references for examination. Ten percent of these references from the electronic database search were deemed suitable for further examination of the hard copies. Every 100th abstract was collected for inter-rater reliability to test agreement about hard copy collection of the reference paper. Agreement between pairs of raters was high at 91.8%. Of the 1272 references marked for hard copy collection, 1075 (84.5%) were retrieved in the time available. Of these hard copy references, fewer than one-third were finally found eligible for inclusion in the review after more rigorous examination of the methodology. The 302 eligible references reported 308 separate studies for inclusion in the review (see Appendix 3 for a complete list of the references included in the review).

Properties of the 308 studies included in the structured review

Owing to the inclusion of published studies in the review, the majority of studies were sourced from journals (Table 2).

The most frequent country of origin of studies included in the review was the USA, producing nearly three-quarters of the studies (Table 3).
This may be due to the trend of more robust experimental design used in the USA, enabled by large correctional facilities.

Nearly half of the studies included in the review were RCTs (Table 4). The rest of the study designs were similarly spread across cohort and comparative study designs. These results reveal that the number of included studies doubled through inclusion of study types other than RCTs. Cross tabulation of the study design and country of origin revealed that significantly more RCTs occurred in the US than in any other region, $\chi^2(12, n = 308) = 39.0$, $p < 0.001$.

Most studies were set in the community (Table 5). The proportion of community-based studies (48.1%) was nearly matched by the proportion of studies set within an institution (43.5%), including prison, secure forensic hospital, juvenile centre and remand. Half of the institutional studies were set in adult prisons, where most of the participants were serving a sentence for a conviction rather than being on remand. Considering the forensic mental health target of the study, it is of interest that only 11% of the included studies were conducted within secure forensic hospitals. This low percentage of forensic hospital location suggests that the majority of forensic mental health research considers the mental health of the general offender population rather than the mentally disordered offender population.

The sample size of the included studies displayed a peak at 101–200, with a curved distribution weighted towards the lower sample size, with a shallow curve for the larger sample sizes (Table 6). The largest sample was 65,390, which was considerably larger than even the next largest sample size of 4072. The largest study sample size distorted the mean sample size considerably to 487, which would have been 276 without its inclusion.

About two-thirds of the included studies consisted of an adult sample, and about one-third an

---

### TABLE 2  Type of report

<table>
<thead>
<tr>
<th>Report type</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal article</td>
<td>293</td>
<td>95.1</td>
</tr>
<tr>
<td>Government report</td>
<td>9</td>
<td>2.9</td>
</tr>
<tr>
<td>Book/chapter</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>Non-profit research institute report</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>Conference proceeding</td>
<td>1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

### TABLE 3  Country of publication

<table>
<thead>
<tr>
<th>Country of publication</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>223</td>
<td>72.4</td>
</tr>
<tr>
<td>UK and Ireland</td>
<td>34</td>
<td>11.0</td>
</tr>
<tr>
<td>Canada</td>
<td>20</td>
<td>6.5</td>
</tr>
<tr>
<td>Other European</td>
<td>15</td>
<td>4.9</td>
</tr>
<tr>
<td>Australia/New Zealand</td>
<td>9</td>
<td>2.9</td>
</tr>
<tr>
<td>Middle East/Asia</td>
<td>6</td>
<td>1.9</td>
</tr>
<tr>
<td>Africa</td>
<td>1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

### TABLE 4  Study design

<table>
<thead>
<tr>
<th>Study design</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomised controlled trial</td>
<td>140</td>
<td>45.5</td>
</tr>
<tr>
<td>Cohort study</td>
<td>85</td>
<td>27.6</td>
</tr>
<tr>
<td>Other comparative design</td>
<td>83</td>
<td>26.9</td>
</tr>
</tbody>
</table>
TABLE 5  Study setting

<table>
<thead>
<tr>
<th>Study setting</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community</td>
<td>148</td>
<td>48.1</td>
</tr>
<tr>
<td>Prison</td>
<td>65</td>
<td>21.1</td>
</tr>
<tr>
<td>Secure forensic hospital</td>
<td>34</td>
<td>11.0</td>
</tr>
<tr>
<td>Juvenile centre</td>
<td>31</td>
<td>10.1</td>
</tr>
<tr>
<td>Other</td>
<td>18</td>
<td>5.8</td>
</tr>
<tr>
<td>Therapeutic community</td>
<td>8</td>
<td>2.6</td>
</tr>
<tr>
<td>Remand</td>
<td>4</td>
<td>1.3</td>
</tr>
</tbody>
</table>

TABLE 6  Sample size

<table>
<thead>
<tr>
<th>Number of subjects</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–50</td>
<td>54</td>
<td>17.5</td>
</tr>
<tr>
<td>51–100</td>
<td>62</td>
<td>20.1</td>
</tr>
<tr>
<td>101–200</td>
<td>72</td>
<td>23.4</td>
</tr>
<tr>
<td>201–300</td>
<td>42</td>
<td>13.6</td>
</tr>
<tr>
<td>301–400</td>
<td>22</td>
<td>7.1</td>
</tr>
<tr>
<td>401–500</td>
<td>20</td>
<td>6.5</td>
</tr>
<tr>
<td>501–750</td>
<td>21</td>
<td>6.8</td>
</tr>
<tr>
<td>750–5000</td>
<td>14</td>
<td>4.5</td>
</tr>
<tr>
<td>&gt;5000</td>
<td>1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

TABLE 7  Participant age

<table>
<thead>
<tr>
<th>Age</th>
<th>Adolescent (n)</th>
<th>Adolescent (%)</th>
<th>Adult (n)</th>
<th>Adult (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>102</td>
<td>33.1</td>
<td>206</td>
<td>66.9</td>
</tr>
<tr>
<td>No</td>
<td>193</td>
<td>62.7</td>
<td>90</td>
<td>29.2</td>
</tr>
<tr>
<td>Not stated</td>
<td>13</td>
<td>4.2</td>
<td>12</td>
<td>3.9</td>
</tr>
</tbody>
</table>

TABLE 8  Participant gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male subjects</td>
<td>132</td>
<td>42.9</td>
</tr>
<tr>
<td>Female subjects</td>
<td>10</td>
<td>3.2</td>
</tr>
<tr>
<td>Mixed</td>
<td>130</td>
<td>42.2</td>
</tr>
<tr>
<td>Not stated</td>
<td>36</td>
<td>11.7</td>
</tr>
</tbody>
</table>

adolescent sample, with the cut-off age of 18 years (Table 7). The inclusion of adult or adolescent samples was not mutually exclusive, with nine studies including both adults and adolescents. The distribution of studies including adolescents and adults reflects the trend of research concerning early intervention with young offenders.

The majority of included studies consisted of a male only sample (Table 8). The prevalence of a male only sample was closely followed by a mixed sample. However, only two of the mixed sample studies consisted of more female than male participants, with 50% of the mixed samples consisting of at least 75% male participants.
TABLE 9  Criminal history of participants

<table>
<thead>
<tr>
<th>Criminal history</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any offence/felony/offence type not stated</td>
<td>204</td>
<td>66.2</td>
</tr>
<tr>
<td>Drug offence/use</td>
<td>66</td>
<td>21.4</td>
</tr>
<tr>
<td>Sexual offence</td>
<td>46</td>
<td>14.9</td>
</tr>
<tr>
<td>Domestic violence</td>
<td>19</td>
<td>6.2</td>
</tr>
<tr>
<td>Violent offence</td>
<td>9</td>
<td>2.9</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
<td>2.3</td>
</tr>
</tbody>
</table>

The criminal history sample characteristics included each offence type identified by a study as being the principle source of data (Table 9). For example, the study title of a paper ‘Long-term treatment and management of violent tendencies of men with intellectual disabilities convicted of assault’ clearly defines the participant sample as having an assault conviction, thus would be coded as a violent offence. Many studies did not specify the offence history of participants for inclusion and used a general offender sample such as prisoners in general. The general offence category consisted of two-thirds of the studies included in the review. Studies that included participants with a specific criminal history included drug offenders, sexual offenders, domestic violence offenders and violent offenders. The criminal history relevant to the participant sample was also not mutually exclusive. For example, some studies used both violent and sexual offenders, which was scored as a presence of both types of offence. Within the specific offence types, drug offenders were the most prevalent, followed by sexual offenders. For violent offences, those that were specified as domestic violence were totalled separately to those that specified general violence such as assault. The two violent offence types are of interest as there were over twice as many studies that specifically considered domestic violence rather than other types of violence. This result reflects the proliferation of domestic abuse programmes that may be assumed to be easily studied.

Similar to the criminal history variables of the sample, the psychiatric diagnosis of the participants was noted wherever there was an explicit description. As can be seen in Table 10, only a minority of studies provided an explicit description of mental health characteristics of the sample. Even in this group, a non-specific description was used such as ‘mental illness’. In this group, six studies specified participants who were not guilty by reason of insanity and six specified participants who were detained under the Mental Health Act 1983. Substance abuse was the other explicit psychiatric specification where participants had been diagnosed, usually by the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria. Finally, four studies explicitly specified disorders including personality, affective, sexual and behavioural.

By far the most frequent type of intervention was cognitive behavioural, designed to reduce offending behaviour (Table 11). Therapeutic communities were the next most frequent intervention explored by the included studies. Other interventions occurred in less than 10% of

TABLE 10  Psychiatric diagnosis of participants

<table>
<thead>
<tr>
<th>Psychiatric history</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental illness</td>
<td>35</td>
<td>15.6</td>
</tr>
<tr>
<td>Of which not guilty by reason of insanity</td>
<td>6</td>
<td>1.9</td>
</tr>
<tr>
<td>Of which detained under the Mental Health Act 1983</td>
<td>6</td>
<td>1.9</td>
</tr>
<tr>
<td>Substance abuse</td>
<td>20</td>
<td>6.5</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>3.6</td>
</tr>
<tr>
<td>Personality disorder</td>
<td>5</td>
<td>1.6</td>
</tr>
<tr>
<td>Affective disorder</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>Sexual disorder</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>Behaviour disorder</td>
<td>2</td>
<td>0.6</td>
</tr>
</tbody>
</table>
TABLE 11 Type of intervention

<table>
<thead>
<tr>
<th>Intervention</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive/behavioural</td>
<td>121</td>
<td>39.3</td>
</tr>
<tr>
<td>Therapeutic community</td>
<td>39</td>
<td>12.7</td>
</tr>
<tr>
<td>Community supervision/aftercare/mental health services</td>
<td>29</td>
<td>9.4</td>
</tr>
<tr>
<td>Multisystemic therapy</td>
<td>20</td>
<td>6.5</td>
</tr>
<tr>
<td>Forensic psychiatric unit/high security hospital</td>
<td>16</td>
<td>5.2</td>
</tr>
<tr>
<td>Drug court</td>
<td>14</td>
<td>4.5</td>
</tr>
<tr>
<td>Alternative therapy</td>
<td>10</td>
<td>3.2</td>
</tr>
<tr>
<td>Family therapy</td>
<td>9</td>
<td>2.9</td>
</tr>
<tr>
<td>Jail diversion</td>
<td>9</td>
<td>2.9</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>2.6</td>
</tr>
<tr>
<td>Case management</td>
<td>6</td>
<td>1.9</td>
</tr>
<tr>
<td>Foster family care</td>
<td>6</td>
<td>1.9</td>
</tr>
<tr>
<td>Mental health court</td>
<td>6</td>
<td>1.9</td>
</tr>
<tr>
<td>Medical drug treatment</td>
<td>6</td>
<td>1.9</td>
</tr>
<tr>
<td>Node link mapping</td>
<td>5</td>
<td>1.6</td>
</tr>
<tr>
<td>Treatment need assessment/assignment</td>
<td>4</td>
<td>1.3</td>
</tr>
</tbody>
</table>

The included studies. Psychiatric interventions such as mental health aftercare/community care and forensic psychiatric units were both under 10% prevalence. Drug court, jail diversion and mental health courts (together consisting of 10.3% of included studies) are all implemented to prevent jail admission through attempts to tackle the core problems of the offenders rather than using incarceration. Family lifestyle interventions such as foster family care, multisystemic therapy and family therapy were interventions targeted at young offenders in an attempt to create resilience from further offending behaviour by altering their home environments. The family-oriented interventions also reflect the trend in sample age where one-third of the studies involved adolescent participants. Treatment provision oriented interventions such as case management and matching treatment need with assessment both occurred in less than 2% of studies. Node link mapping, also occurring in under 2% of included studies, was a new type of therapy strategy. Finally, the use of medical drug interventions was quite low at 1.9% prevalence.

Outcome measurement results

This section considers the subject of outcome measures found in the current survey of forensic mental health research. Information from both the structured review and consensus panel will be used to assess outcomes. First, the issue of domains of outcome measurement for forensic mental health research will be considered. Second, the properties of the most frequently occurring outcome measurement instruments from the review will be examined, with reference to their properties for use in forensic populations. This evidence about most frequently occurring measures was also assessed by the consensus panel, and their ratings and views will be reported. Finally, this section will give particular attention to the outcome measure of recidivism, from the patterns extrapolated from the review to the opinions of the consensus panel.

Domains of outcome measurement

Consensus domain decisions

Domains: ratings

The consensus panel was asked in written correspondence to consider a draft of potential domains of outcome in forensic mental health research. Mean ratings of importance of domains of outcome for forensic mental health research are given in Table 12, listed in rank order. No significant additional domains were identified in comments received and none were dropped as being redundant.
Domains: group discussion

When the consensus panel came together the group was invited to review the quantitative scores previously assigned to domains in written ratings (Table 12). The group did not wish to revise their scores for importance of domains. All 11 domains were considered important (all having received mean ratings of 3 or more). There was universal agreement that recidivism is one of the most important outcomes. Death, suicide and violent recidivism were seen as outcomes consequent upon a failure to identify early warning signals captured within lower ranked domains (e.g. engagement, social functioning). This view suggests that lower domains should be given relatively greater priority for measurement as a way of preventing the most negative outcomes from occurring.

Some more specific points were made. It was noted that employment would be better described as ‘meaningful activity’ as many forensic mental health service users never or rarely work. There was surprise that physical health was ranked the least important domain, given growing concerns around low uptake of health services within forensic mental health service users. Association with criminal peers was noted as a potentially important outcome not captured specifically by any domain. The literature suggests it is one of the best predictors of criminality (in addition to previous convictions), particularly in the case of persistent offenders. The list of domains was endorsed as appropriate.

It was felt that the relative importance of domains is contextual and depends on factors such as:

1. The question being asked by the study.
2. The population being studied.
3. The perspective (e.g. society at large versus the individual service user).

Structured review classification of outcome measures into consensus-specified domains

The 308 studies in this review included 1038 separate variables treated as outcomes, of which 450 were instruments in the sense of scales or

<table>
<thead>
<tr>
<th>Domains</th>
<th>Description</th>
<th>Mean*</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recidivism: violent</td>
<td>Reoffending, violent and sexual</td>
<td>4.6</td>
<td>1</td>
</tr>
<tr>
<td>Suicide</td>
<td></td>
<td>4.5</td>
<td>1</td>
</tr>
<tr>
<td>Substance abuse</td>
<td>Being addicted, e.g. to alcohol, drugs</td>
<td>4.1</td>
<td>1</td>
</tr>
<tr>
<td>Recidivism: non-violent</td>
<td>Reoffending in some way other than violent or sexual</td>
<td>4.0</td>
<td>2</td>
</tr>
<tr>
<td>Mental state</td>
<td>General psychological well-being</td>
<td>4.0</td>
<td>2</td>
</tr>
<tr>
<td>Engagement with treatment</td>
<td></td>
<td>3.9</td>
<td>2</td>
</tr>
<tr>
<td>Relationships</td>
<td>With family, friends, etc.</td>
<td>3.8</td>
<td>2</td>
</tr>
<tr>
<td>Aggression</td>
<td>Verbal or physical</td>
<td>3.8</td>
<td>2</td>
</tr>
<tr>
<td>Cognitive/psychological function</td>
<td>Planning, remembering, problem solving</td>
<td>3.8</td>
<td>1.25</td>
</tr>
<tr>
<td>Death</td>
<td></td>
<td>3.8</td>
<td>2.25</td>
</tr>
<tr>
<td>Self-harm</td>
<td></td>
<td>3.8</td>
<td>2</td>
</tr>
<tr>
<td>Service outcomes</td>
<td>How much someone uses national services</td>
<td>3.7</td>
<td>1</td>
</tr>
<tr>
<td>Compliance</td>
<td>Adherence; concordance</td>
<td>3.5</td>
<td>1</td>
</tr>
<tr>
<td>Stages of change/readiness</td>
<td>Willingness and motivation to change selves/situation</td>
<td>3.5</td>
<td>1</td>
</tr>
<tr>
<td>Economic</td>
<td>Costs/pay back to society (e.g. service use, working)</td>
<td>3.5</td>
<td>1</td>
</tr>
<tr>
<td>Social function</td>
<td>Day-to-day activities that involve contact with people</td>
<td>3.3</td>
<td>1</td>
</tr>
<tr>
<td>Quality of life</td>
<td></td>
<td>3.2</td>
<td>1</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>How good someone feels about themselves</td>
<td>3.2</td>
<td>2</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td>3.2</td>
<td>1</td>
</tr>
<tr>
<td>Satisfaction with treatment</td>
<td></td>
<td>3.1</td>
<td>2</td>
</tr>
<tr>
<td>Physical health</td>
<td></td>
<td>3.0</td>
<td>2</td>
</tr>
</tbody>
</table>

IQR, interquartile range.

a Scale ranges from not important (1) to essential (5), with 3 and above indicating ‘important’.
TABLE 13 Frequency of use of variables and instruments in different domains of outcome in forensic mental health research

<table>
<thead>
<tr>
<th>Domain</th>
<th>Total number of different variables and instruments</th>
<th>Number of instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recidivism: non-violent</td>
<td>314</td>
<td>45</td>
</tr>
<tr>
<td>Substance abuse</td>
<td>133</td>
<td>73</td>
</tr>
<tr>
<td>Service outcomes</td>
<td>99</td>
<td>13</td>
</tr>
<tr>
<td>Recidivism: violent or sexual</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>Mental state</td>
<td>74</td>
<td>65</td>
</tr>
<tr>
<td>Cognitive/psychological function</td>
<td>74</td>
<td>71</td>
</tr>
<tr>
<td>Relationships</td>
<td>41</td>
<td>35</td>
</tr>
<tr>
<td>Compliance</td>
<td>31</td>
<td>13</td>
</tr>
<tr>
<td>Economic</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>Satisfaction with services</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>Social function</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>Physical health</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Employment</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Engagement with treatment</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Aggression</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Stage of change/readiness</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Quality of life</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Suicide</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Death</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>50</td>
<td>27</td>
</tr>
</tbody>
</table>

Multiple item questionnaires. The outcome variables and instruments were classified according to the domains identified as important by the consensus panel (Table 13). As can be seen, the majority of variables and outcome instruments could be adequately described by the classificatory schema for domains.

Non-violent recidivism was assessed through 314 different variables. This variety reflects the diversity of studies, including different stages in the criminal justice process, different types of offence and different types of activity or event considered as recidivism. Examples of non-violent recidivism include illegal activities in the past 90 days, number of months until rearrest, new charge for a property offence and time before first conviction. Violent recidivism displayed less diversity with 80 different methods of measurement and a larger proportion recorded by means of formal instruments (25.0%).

The next largest variety of variables to assess an outcome is in substance abuse, with 133 distinguishable variables, of which just over half comprised instruments (54.9%). Many of the remaining domains comprised variables where 75% or more were formal instruments, for example, mental state, engagement with treatment, relationships, aggression, cognitive/psychological function, stages of change/readiness, quality of life, self-esteem, satisfaction with treatment and physical health. Variables outside the domains identified by the expert panel are grouped together as ‘other’ and include accommodation, sexual behaviour and victimisation.

Properties of the most frequently occurring outcome measure instruments

Further investigation of more commonly used outcome measure instruments was undertaken. Each outcome measure instrument that appeared in more than four separate studies was isolated for further examination. The cut-off of five or more uses was arbitrary, but allowed the research group to focus on a manageable number of outcome measures used with reasonable frequency in
Results

Nine different outcome measures occurred in more than four different studies (Table 14). The most frequent outcome measures were the Addiction Severity Index (ASI) and the various configurations of the Symptom-Checklist-90-Revised (SCL-90-R), both occurring in 15 studies. The next most frequent was the Self-Reported Delinquency Scale (SRDS), which occurred in 11 studies. The remaining six outcome measures appeared in between five and seven different studies.

The focus of interest for these selected outcome measures was the amount of evidence available for their psychometric properties for forensic populations. A simple search for psychometric properties of the selected outcome measures was conducted on MEDLINE and PsycINFO, including the keywords of the instrument name and ‘validity’ and ‘reliability’. The subsequent gathered evidence was assessed according to four criteria:

1. Adequacy of measurement properties for general use.
2. Adequacy of measurement properties for use in forensic mental health research.
3. Feasibility of use in forensic mental health research.\(^{12}\)
4. Relevance and appropriateness for use in forensic mental health research.

The evidence gathered will now be outlined for each outcome measure. As there appear to be many different interpretations of different types of psychometric properties in the literature, in this

---

**TABLE 14 Frequency of occurrence of outcome measures**

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Number of times occurred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addiction Severity Index</td>
<td>15</td>
</tr>
<tr>
<td>Symptom-Checklist-90-Revised (SCL-90-R) and precursor and part measures: SCL-90-R; SCL-90; Hopkins Symptom Checklist; Brief Symptom Inventory; Global Severity Index(^a)</td>
<td>15 (4; 3; 3; 3; 4)</td>
</tr>
<tr>
<td>Self-Reported Delinquency Scale</td>
<td>11</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>7</td>
</tr>
<tr>
<td>Conflict Tactics Scale</td>
<td>7</td>
</tr>
<tr>
<td>Revised Behavior Problem Checklist</td>
<td>6</td>
</tr>
<tr>
<td>Brief Psychiatric Rating Scale</td>
<td>5</td>
</tr>
<tr>
<td>Child Behavior Checklist</td>
<td>5</td>
</tr>
<tr>
<td>Family Adaptability and Cohesion Evaluation Scales III</td>
<td>5</td>
</tr>
</tbody>
</table>

\(^a\) Two studies used both the SCLs (90 and 90-R) in addition to the Global Severity Index measures.

---

**TABLE 15 Psychometric definitions**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reliability</strong></td>
<td></td>
</tr>
<tr>
<td>Internal consistency</td>
<td>The consistency of the measure or subscale items. (Normally Cronbach’s alpha ≥ 0.7)</td>
</tr>
<tr>
<td>Test–retest</td>
<td>The consistency of test scores over two or more administrations with a time lapse. (Normally correlations between scores ≥ 0.7)</td>
</tr>
<tr>
<td>Inter-rater</td>
<td>The consistency of ratings between two separate raters</td>
</tr>
<tr>
<td><strong>Validity</strong></td>
<td></td>
</tr>
<tr>
<td>Concurrent (convergent) validity</td>
<td>Correlation with a measure that has already been validated and measures a similar concept</td>
</tr>
<tr>
<td>Divergent validity</td>
<td>The degree to which a measure does not correlate with other measures that it theoretically should not be similar to</td>
</tr>
<tr>
<td>Discriminant validity</td>
<td>The ability of the measure to discriminate between different populations</td>
</tr>
<tr>
<td>Content validity</td>
<td>The extent to which a measure represents all of the facets of a concept</td>
</tr>
</tbody>
</table>
review the properties were classified as shown in Table 15.

**Addiction Severity Index**

**General information**

The ASI was developed by Thomas McLellan et al.\(^\text{13}\) to evaluate the outcome of an addiction treatment. Although often conducted as a self-administered questionnaire, the ASI was designed and intended to be a semi-structured clinical evaluation interview. McLellan\(^\text{14}\) has stated that interview training is possible for anyone who is able to form rapport, understand the patient, and probe confused answers with clarifying questions, which in their experience has been about 90% of those they have trained. The interview is predicted to last from 50 minutes to 1 hour.\(^\text{14}\)

The ASI consists of 60 items that fall into seven subscales:

1. medical (which refers to lifetime hospitalisations and chronic problems)
2. employment/support (e.g. education and training, skills, employment patterns)
3. drug (history of drug use, treatment for addiction, overdoses)
4. alcohol (history of alcohol use, treatment for addiction)
5. legal (convictions, any current charges, criminal involvement)
6. family/social (e.g. marital conditions, stability, satisfaction, problems, conflicts)
7. psychiatric (hospitalisations and life experiences).

The interviewee answers with reference to their experiences in the past 30 days and also their lifetime in general. Individually, the interviewee and the interviewer give a rating on a five-point scale (0: not at all, 4: extremely) of the perceived severity of the interviewee’s problems (severity scores). Severity scores assess the current severity of the problem area. Composite scores are then developed that consist of a combination of items that are capable of showing change.

The ASI has experienced prolific use, exemplified by nine language translations including: French, Spanish, German, Dutch and Russian.\(^\text{14}\) It has also been utilised in many different populations such as methadone maintenance patients, alcohol treatment patients, cocaine abusers,\(^\text{14}\) prisoners,\(^\text{15,16}\) the homeless,\(^\text{17}\) and the mentally ill.\(^\text{18}\)

**Adequacy of measurement properties – general**

Within general drug abusing samples the ASI has displayed good reliability\(^\text{13}\) and validity.\(^\text{19,20}\) Exploration of the factor structure has replicated the seven subscales with 990 methadone maintenance patients.\(^\text{21}\) However, Alterman et al.\(^\text{22}\) identified only five addiction problem scales (psychiatric, drug, alcohol, family and legal) from a sample of 1008 substance dependent patients. This five-factor structure may also represent problems that have occurred with discriminant validity through correlations between the social and psychiatric subscales.\(^\text{23}\) Thus, the social and psychiatric subscales may not measure distinct areas.

**Adequacy of measurement properties – forensic mental health**

Information regarding the psychometric properties of the ASI within an English speaking forensic sample was found in only one report – a study of 128 inmates by Amoureus.\(^\text{15}\) Also, the French version of the ASI was tested by Brochu,\(^\text{24}\) and found to have sufficient reliability and validity with 304 inmates.

**Reliability – internal consistency**

The internal consistency of the subscale composite scores has been shown to be good in a drug abusing inmate sample.\(^\text{15}\) All of the alpha coefficients were above the recommended 0.6 (medical 0.8; employment 0.63; alcohol 0.65; drugs 0.77; social 0.72; psychiatric 0.76), except for the legal (0.53) scale.\(^\text{25}\)

**Discriminant validity**

To display good discriminant validity, the ASI subscales that denote different problem areas must not correlate, thus displaying measurement of different constructs. Each subscale area must measure unique entities. A prisoner sample of 128 inmates\(^\text{15}\) showed that overall most of the severity and composite scores of one subscale correlated, whilst most of the different subscales did not. However, a strong association was found to exist between the severity ratings and composite scores for employment and psychiatric disorders (0.33, \(p < 0.001\)). Also, the severity ratings and composite scores for social and psychiatric (0.55, \(p < 0.001\)) and social and employment (0.40, \(p < 0.001\)) problem area subscales were seen to correlate with each other.
Concurrent validity
In the prisoner sample the concurrent validity was assessed through correlation of recent (within last 6 months) Diagnostic Interview Schedule/DSM-III diagnoses and the ASI psychiatric severity ratings (low severity, medium severity, high severity). The high severity group were significantly different to the medium and low group for depressive episode, any depressive disorder, any anxiety disorder and any DSM-III axis I disorder.

Feasibility for forensic mental health research
The ASI is quite a long interview (1 hour) and may only be feasible as an intake assessment rather than a repeated outcome measure. The high rate of individuals able to conduct the interview suggests its transferability to researchers rather than it just being administered by clinicians.

Relevance to forensic mental health research
Substance abuse is highly prevalent in forensic populations, and may influence offending behaviour through illegal attempts to fund a drug habit and intoxication influenced behaviour. Thus, a large body of research considers the influence of substance abuse within forensic mental health research, making the ASI highly relevant.

Addiction Severity Index summary
Overall, the ASI has produced positive evidence in favour of its reliability and validity within both a general substance abusing population and a prison sample. Specifically, within a forensic prison sample an area of caution is the legal subscale, which did not produce good internal consistency. This weakness is of concern as the ASI’s legal subscale is often highlighted within forensic mental health outcome research as a self-report criminal behaviour outcome measure. Thus, for assessment of substance abuse severity within a forensic sample the ASI appears valid, yet caution must be paid to its legal scale for reporting offending behaviour outcomes. Overall, the ASI has received much attention for its reliability and validity within general substance abusing populations, yet within forensic samples the evidence is sparse. Although the study by Amoureus provides a comprehensive view of the psychometric properties of the ASI, the sample is small (n = 128). As the ASI was the most frequently used outcome measure in the studies included in this review it is questionable whether its suitability for forensic mental health research has been fully explored.

Beck Depression Inventory
General information
Created by Beck et al. and then revised in 1971, the Beck Depression Inventory (BDI) is a self-report instrument. It consists of 21 items that are considered to be symptoms of depression. The individual rates each item on a 0–3 scale. These scores are then totalled, with higher scores reflecting the most severe depressive symptoms.

The BDI has been used in more than 2000 empirical studies in the years since its introduction in 1961. There is a long form consisting of 21 items and a short form consisting of 13 items. These two forms have been found to correlate strongly (0.89–0.97) in populations such as psychiatric, non-psychiatric and heroin addicts. A later version, the BDI-II was constructed to make the instrument more compatible with DSM criteria; this version was also highly congruent to the previous version long form. However, the short form is thought to represent one cognitive symptom dimension, whereas the long form also represents non-cognitive symptom clusters.

Adequacy of measurement properties – general
The longevity of the BDI has made several meta-analyses on its psychometric properties available. Beck reviewed studies from the BDI’s inception in 1961 to 25 years later, 1986, whilst Richter et al. extended a review from 1961 to 1998. These meta-analyses have provided ample evidence for the strong validity and reliability of the BDI in both psychiatric and non-psychiatric populations. The only problems reported with the BDI concern divergent validity for associations with anxiety and questionable test–retest validity, although it has been argued that these reflect the sensitivity of the BDI to change. In addition, results concerning the factor structure of the BDI have varied from three to seven factors. Of these studies, a three-factor structure including negative attitudes towards self, performance impairment and somatic disturbance was found by Beck and Lester, and then later replicated by Tanaka and Huba.
Adequacy of measurement properties – forensic mental health

Compared with the extensive evidence for the BDI from psychiatric and non-psychiatric populations, few studies actually examined a forensic population. Giambra34 included 20 male prisoners in a sample also consisting of 91 college students, and Scott et al.35 tested a sample of 65 female prisoners. A larger study of 1494 prisoners considered only the discriminant validity and the factor structure of the BDI.36

Reliability – internal consistency

High internal consistency was displayed in a sample including 29 male prisoners with 91 college students (Spearman–Brown coefficient of 0.87)34 and a sample of 65 female prisoners (alpha coefficient of 0.9).35

Discriminant validity

In a prisoner population the BDI was able to discriminate between those in close custody and those in medium or minimum custody, as they were significantly more depressed.36 Also, first-time prison inmates displayed significantly more depression on the BDI.

Concurrent validity

Within a forensic population of 29 male prisoners, included with 91 college students, the BDI correlated with the Zung SRDS with a correlation coefficient of 0.66.34 Also, the BDI displayed a correlation coefficient of 0.63 with the Minnesota Multiphasic Personality Inventory (MMPI) Depression Scale in a sample of 65 female prisoners.35 Thus, the BDI has displayed good concurrent validity within forensic samples.

Factor structure

In a prisoner sample of 1494, four distinct factors were found: cognitive symptoms, vegetative symptoms, emotional symptoms and feelings of punishment.36

Feasibility for forensic mental health research

The BDI is a self-report measure which is available in both a 21-item and 13-item form. The short duration of completing the BDI means that it would be suitable for repeated measures outcome measurement, possibly as part of a testing battery. The limitation of self-report instruments in a forensic population is the low levels of literacy. However, assistance with understanding the items would still take limited time, especially with the short form.

Relevance to forensic mental health research

Depression is clearly a large problem associated with incarceration,36 and is thus highly relevant to forensic mental health research. Feelings of depression may lead to suicide attempts,32 which pose a public health issue. Depression may also affect the impact of interventions as the participant may not be susceptible to behaviour change whilst unable to conceive of a future.37

Beck Depression Inventory summary

The BDI has displayed good psychometric properties in psychiatric and non-psychiatric samples. In forensic samples the high internal consistency and satisfactory concurrent validity have been replicated, but test–retest reliability and divergent validity have not been explored. Further, the varying factor structures displayed by non-forensic samples are sustained by the four-factor solution displayed by a prisoner sample, where a factor specific to incarceration appears to have developed: feelings of punishment. Thus, whilst the BDI has displayed robust qualities in psychiatric and non-psychiatric populations, its direct applicability to forensic samples requires more examination as it may measure alternative themes.

Brief Psychiatric Rating Scale

General information

The Brief Psychiatric Rating Scale (BPRS) was created by Overall and Gorgam38 for use with individuals with psychiatric disorders such as schizophrenia. It evaluates treatment change whilst also describing major symptom characteristics.38 The BPRS usually takes the form of an interview. Self-administered forms of the scale are not encouraged as the interview allows disorganised speech and unusual thoughts to be more easily observed. The interview should be conducted only by clinicians or other trained raters such as social workers, as an understanding of the symptoms and their scores is required. However, high levels of
training are not absolutely necessary for reliable administration of the BPRS. Ventura et al. found that both an advance trained and a postdoctoral group of administrators were able to produce an excellent inter-rater reliability intraclass correlation coefficient for 22 of the 24 items. Further, the excellent levels of reliability were maintained over 6 and 12 months after initial training.

There are several versions of the BPRS. Originally it was a 16-item measure based on principal symptom factors from a large set of items taken from the Inpatient Multidimensional Psychiatric Scale. In 1974 Overall added two new symptom items, and in 1986 Lukoff et al. added six more for better evaluation of patients with schizophrenia. The 18-item version has also been anchored. A global symptom score can be calculated by adding the points for each item. Even though there are many different versions, many researchers refer to all of them as the BPRS, so it can be difficult to establish which variant has been used.

The BPRS consists of 16–24 items that are rated on a seven-point scale (1: not present; 7: extremely severe) that measures positive and emotional symptoms, along with general psychopathology. Some items require self-reporting by the patient (e.g. anxiety, hallucinations, etc.), whereas others can be observed (e.g. mannerisms). The BPRS items produce four subscales including thinking disturbance, withdrawal retardation, anxious depression and hostile suspiciousness. The interview is specified to last 18 minutes, but in practice it can vary according to the patient.

The BPRS has been cited in over 1000 medical studies as the main outcome measure for psychopharmacological and psychotropic medications. Its popularity is also expressed by its presence in the list of outcome measures identified for use in the assessment of psychiatric symptom change by the Joint Commission of Accreditation of Healthcare Organizations, which evaluates and accredits nearly 19,000 health-care programmes and organisations in the USA.

**Adequacy of measurement properties – general**

The BPRS has displayed good internal consistency, inter-rater reliability, discriminant validity, and a relatively consistent factor structure. However, these studies of reliability and validity were conducted on psychiatric patients, with little information about the psychometric properties of the BPRS conducted on forensic mental health populations.

**Adequacy of measurement properties – forensic mental health**

Few studies were available to provide the psychometric properties of the BPRS in a forensic sample, with the most comprehensive study describing the concurrent validity of the BPRS for 192 prisoners.

**Reliability – inter-rater**

Concordance of ratings with a ‘gold standard’ training level for 21 mental health professionals from a forensic psychiatric hospital displayed an average 0.83 concordance rate for all the items combined, ranging from 0.60 to 0.98 for each item.

**Discriminant validity**

The BPRS total score was able to predict violence in 34 mentally disordered offenders.

**Concurrent validity**

In a sample of 192 prisoners the BPRS identified 33% defined broadly as having a disorder and 15% defined narrowly as having a disorder. These results were compared with 14% broad and 11% narrow from the diagnostic profile, and 80% broad and 16% narrow from the diagnostic interview schedule, version III-A. The BPRS and diagnostic profile had moderate agreement \( (k = 0.45, p < 0.001) \), but nearly no agreement with the diagnostic interview schedule on broad disorders. For narrow disorders the BPRS and the diagnostic profile agreed \( (k = 0.57, p < 0.001) \), but again agreement with the diagnostic interview schedule was only just significant. Thus, when compared with the diagnostic profile the BPRS displays good concurrent validity, but when compared with the diagnostic interview schedule it does not.

**Feasibility for forensic mental health research**

Although the BPRS is recommended for administration by mental health professionals, it appears that anyone with sufficient interview skills can administer it making it feasible for forensic
mental health research where an interviewer is available. The BPRS’s popularity is further reflected by its quick administration time, making it feasible for outcome measurement.

**Relevance to forensic mental health research**

The BPRS appears to be a convenient instrument for quick measurement of psychiatric symptoms. Thus, it is relevant to assessment of psychiatric symptoms within a forensic mental health population.

**Brief Psychiatric Rating Scale summary**

Information concerning the psychometric properties of the BPRS for use in a forensic mental health population is sparse. It appears to have good discriminant validity and moderate concurrent validity with the diagnostic profile, but not the diagnostic interview schedule. Research about the internal consistency of the BPRS with a forensic mental health population is required, as well as factor structure exploration to determine if it measures the same properties within a forensic mental health population as within a psychiatric one.

**Child Behavior Checklist**

**General information**

The Child Behavior Checklist (CBCL) is a parent report questionnaire designed to assess the behavioural problems and social competencies of children aged 4–18 years. Developed by Achenbach, the CBCL consists of two sections. The first section consists of 20 competence items grouped into four competence subscales. The second section consists of 120 items concerned with problematic behaviour or emotions during the past 6 months, which are grouped into 11 problem subscales (including eight syndrome scales). There are also two higher order scales, internalising and externalising. The items are rated on a seven-point Likert scale. The CBCL is completed by the child’s parents (or other adults who know the child well), and the child’s problem behaviours and competencies are rated. There is also a teacher’s report form using 118 items and a youth self-report form, which shares 89 of the problem items. It is recommended that only trained professionals should examine the results of the CBCL.

The CBCL has displayed much popularity by its use in over 1000 published studies between 1983 and 1993. It has also been validated across 12 countries.

**Adequacy of measurement properties – general**

The CBCL displays acceptable internal consistency and satisfactory test–retest reliability. It also differentiates well between different populations, has good divergent validity and concurrent validity. However, the complexity of the measure appears to affect the independence of the subscales, with much shared variance. There are also problems concerning different factor structures across age groups, perhaps making the CBCL unsuitable for children under 5 years old. Finally, the dubious practice of creating new subscales from the items is a concern, as their presence has not been apparent in any of the previous factor analysis studies.

**Adequacy of measurement properties – forensic mental health**

Unfortunately, no studies assessing the CBCL within a forensic mental health sample were accessed.

**Feasibility for forensic mental health research**

The CBCL is a substantial instrument consisting of 140 items. There is a reliance on the caregiver or teacher to fill the questionnaire, thus their perceptions are the focus of the results. However, this method of administration would present literacy and understanding problems in a young sample.

**Relevance to forensic mental health research**

The subscales and problem areas covered within the CBCL seem highly relevant to delinquent participants, especially the delinquent behaviour scale. The disorders that are outlined such as attention deficit hyperactivity disorder, oppositional defiant disorder and conduct disorder are also relevant for forensic mental health research in young populations. It is peculiar that the CBCL has not been psychometrically evaluated within a forensic mental health population.
Child Behavior Checklist summary

The CBCL assesses many disorders that are associated with delinquent behaviour, such as attention deficit hyperactivity disorder, oppositional defiant disorder and conduct disorder. Considering the properties of the CBCL, it is surprising that it has not specifically been validated with a forensic mental health sample such as forensic hospitals and participants involved within the criminal justice system. Unfortunately, psychometric properties are available only from psychiatric populations.

Conflict Tactics Scale

General information

The Conflict Tactics Scale (CTS) was devised by Straus and is used to identify and evaluate domestic violence within families and other relationships. The CTS measures the extent to which partners who are either dating, cohabiting or married engage in physical and psychological attacks on each other. It also gauges their use of reasoning or negotiation to deal with conflicts. It can be completed by one partner or both partners separately, although it is best to collect information from both partners in the relationship. The CTS is usually self-administered but can be conducted as an interview. It takes approximately 10 minutes to complete. Anyone can complete the CTS (including children), although training is recommended for professionals who assess the reports.

The CTS consists of 78 items, half of which refer to the respondent’s behaviour and the other half to the partner’s behaviour. The respondent then indicates how often the behaviour has occurred on an eight-point scale. These scores make up the ‘self’ and ‘partner’ scores for the following dimensions: negotiation, physical assault, injury, psychological aggression and sexual coercion.

The original version of the CTS consisted of three scales including reasoning, psychological aggression and physical assault. The number of items in these scales was increased in the second version of the CTS, the CTS2, to increase reliability. The CTS2 also consisted of two new scales – sexual coercion and physical injury from assaults by a partner. The increase in instrument size meant that the CTS2 took 10–15 minutes to administer, whereas the CTS took 7–10 minutes. Scores are created by calculating the mean for each set of variables for each subscale and then finding the subscale means across all observations.

The popularity of the CTS was reflected by the publication of 10 studies per month using the CTS in 1994.

Adequacy of measurement properties – general

The CTS displays good internal consistency for the subscales, but varied internal consistency for the items in a student sample. When considering the inter-rater reliability of the CTS, the main concern is the consistency of reports from both the male and female respondents in a couple. Studies report that the male perpetrators of violence under-report their levels of violence compared with their female partner’s report describing their victimisation. Further, a factor analysis of the violence subscale reveals that it produces a different factor structure for males and females. Schafer reported that factor analysis studies show that the violence subscale consists of one factor for female respondents, but not for male respondents. Using the same measure for both males and females may lead to measuring different concepts; a more unitary measure may be required. These conceptual problems cause concern for the validity of the CTS for measuring violence within couples.

Adequacy of measurement properties – forensic mental health

Although the CTS measures behaviour that may be classed as criminal, only two studies have assessed its psychometric properties within forensic samples. First, Browning used a sample of 30 couples where the males were in treatment for domestic abuse, and Jones et al. used a sample of 264 incarcerated females.

Reliability – inter-rater

Browning found that in 30 couples, in which the male was in treatment for domestic violence, the wives rated significantly more violence for their husbands than their husbands rated for themselves [F(3, 87) = 26.045; p < 0.001]. However, there was no difference between the husband rating the wife with the wife’s own rating. Overall, the
correlation of agreement for the husband and wife violence increased for more severe forms of violence. The difference between partner ratings of violent behaviour is likely to represent the minimising of violent behaviour on the part of the perpetrator. When using the violence scale of the CTS, researchers must be aware of the difference in ratings likely to occur between the perpetrator and the victim.

**Concurrent validity**

A sample of 264 incarcerated females completed the CTS2 and the Abusive Behavior Checklist. All of the items from both the self as victim and self as aggressor subscales were positively and significantly correlated with the Abusive Behavior Checklist items.

**Factor structure**

There has been criticism of the violence subscale due to the occurrence of different factors for male and female reporters. In a sample of 264 incarcerated female participants, where prevalence and severity of domestic abuse was high, exploratory factor analysis was conducted combining both the self as victim and the self as aggressor subscales. The factors found were negotiation, sexual coercion, injury and general assault. Whilst these four factors represent three of the CTS’s subscale dimensions, the psychological and physical aggression subscales had combined into one factor: general assault.

**Feasibility for forensic mental health research**

The CTS is quite flexible in its administration with the ability to use it as a self-report or an interview. The instrument is also of moderate length, thus may be acceptable as part of an assessment battery. A problem with feasibility for forensic mental health research is the acceptability of the item content. The CTS asks about highly sensitive information, thus, high refusal rates and distorted answers have been observed, which can result in invalid data.

**Relevance to forensic mental health research**

Domestic violence is a large issue within forensic mental health, thus the CTS is highly relevant. The relevance of domestic abuse as a standalone offence class for study in forensic mental health research was displayed by 6.2% of the studies identified by this review considering participants who had committed this offence. Further, the propensity for ratings from both partners provides more than one view of the violent behaviour.

**Conflict Tactics Scale summary**

The CTS measures the commission of violence in couples. The CTS has been validated using couple groups from the general population, such as students and military personnel. However, only two studies were found that specifically assessed the CTS within forensic samples, a female incarcerated sample and a couples sample where the male was receiving treatment for domestic abuse. A significant concern related to the CTS is the apparent under-reporting of their own violent behaviour in male participants compared with female participants. Perhaps the under-reporting of the males specifically taps the cognitive distortion of minimising violent behaviour, but this issue still causes problems for inter-rater reliability. Further, the four-factor structure of the CTS for incarcerated females did not correspond with the five CTS subscales. The varying factor solutions may provide further evidence that the CTS measures different factors between male and female respondents, although the difference provided by Jones et al. may be due to the forensic sample compared with a non-forensic sample.

**Family Adaptability and Cohesion Evaluation Scale**

**General information**

The Family Adaptability and Cohesion Evaluation Scale (FACES) was compiled from the constructs of cohesion and adaptability from the circumplex model. Family cohesion assesses the degree of separation or connection of family members to the family. There are four levels of family cohesion ranging from extreme low cohesion to extreme high cohesion, these are: disengaged, separated, connected and enmeshed. In addition there are four levels of adaptability: rigid, structured, flexible and chaotic. From the four levels of family cohesion and four levels of adaptability, a classification of 16 family types is constructed, with three more general types: balanced, mid-range and extreme. Family members answer 20
Adequacy of measurement properties – general

For a measure that reached eligibility for inclusion in this review, information about its psychometric properties are sparse. Most studies examining the FACES have considered the structure of the circumplex model rather than indices of reliability or validity. A study that has considered reliability displayed good internal consistency for the subscales in a sample of 243 university students. The FACES is also able to discriminate well between different types of family. However, low consistency of scores between family members and therapists does cause concern.

Adequacy of measurement properties – forensic mental health

Similar to general information about the psychometric properties of the FACES, there is little information concerning forensic samples.

Discriminant validity

Amongst a delinquent child father-absent family (n = 29) and father-absent families without history of arrest or psychiatric referral (n = 29), the FACES was able to differentiate the delinquent from the non-delinquent families.

Factor structure

Some researchers have not found the three subscales of cohesion, adaptability and social desirability to be easily differentiated through factor analysis. This result was partially replicated in a sample of 95 male juvenile offenders, where only two factors that resembled the cohesion and adaptability factors were found. The lack of differentiation of a social desirability factor within a forensic sample causes concern for two reasons: the ability of the FACES to detect lies from a forensic population and an alternate measurement structure between populations.

Feasibility for forensic mental health research

The FACES is a self-report instrument that is of a reasonable length. However, the instrument may be slightly conceptually complex with the respondent requiring to answer both how his or her family currently is and how he or she would like their family to be. The scoring of the instrument and placement into family types may also be complex.

Relevance to forensic mental health research

Within the wider sphere of forensic mental health the FACES may be useful to assess a participant’s home circumstances. However, as an outcome measurement family cohesion does not appear a strong candidate as a main measure of outcome in forensic mental health.

Family Adaptability and Cohesion Evaluation Scales summary

The FACES appears to distinguish between delinquent and non-delinquent families, thus displaying good discriminant validity. In comparison with non-forensic samples, forensic samples appear to report results that consist of two factors rather than three. This may reflect a difference between the two populations regarding what the instrument measures. Again, the information for the psychometric properties of the FACES for a forensic mental health sample is sparse.

Revised Behavior Problem Checklist

General information

The Revised Behavior Problem Checklist (RBPC) is a teacher and parent rating instrument for the major broad categories of child psychopathology. It takes approximately 20 minutes to complete, and consists of 89 items, which the individual then rates on a three-point Likert scale: 0 = not a problem; 1 = a mild problem; 2 = a severe problem. These scores refer to four major subscales including conduct disorder, socialised aggression, attention problems immaturity and anxiety withdrawal, and two minor subscales: psychotic behaviour and motor tension excess.

It is recommended that examiners should have at least a bachelor’s degree in psychology, counselling
or a related field and relevant coursework in psychological measurements and tests.

Adequacy of measurement properties – general

The RBPC has been evaluated in general populations consisting of samples of young school children. The four major and minor subscales have produced good internal consistency.\(^7\)\(^9\),\(^8\)\(^0\) However, inter-rater reliability was not as robust with only mild to modest agreement between parent and teacher ratings.\(^7\)\(^9\) Also, test-retest reliability was not good for longer time periods such as 17 months.\(^8\)\(^1\) Good concurrent and divergent validity has been proven through relevant associations of the attention problems and inattention subscales with DSM-III criteria for attention deficit hyperactivity disorder and interaction/aggression.\(^7\)\(^9\) The original factor structure of four major and two minor scales\(^8\)\(^0\),\(^8\)\(^2\) was replicated with similar factors from a sample of 284 kindergarten children, who were at risk of psychopathology and also with 299 who were not at risk.\(^7\)\(^9\)

Adequacy of measurement properties – forensic mental health

Locating studies examining the psychometric properties of the RBPC within forensic samples was difficult, with only one assessing the discriminant validity.

Discriminant validity

A sample of 24 incarcerated juvenile offenders was compared with 24 non-offending adolescents on the subscales of the RBPC.\(^8\)\(^3\) The offender sample scored higher than the non-offending sample on all subscales of the RBPC, thus displaying more psychopathology, psychoticism, externalising problems and also internalising problems. Therefore the RBPC distinguished between the offenders and non-offenders.

Feasibility for forensic mental health research

Similar to the CBCL, the RBPC is administered to either the teacher or caregiver of the participant. Thus, the results are of the perceptions of the adult who completes the form rather than directly from the individual being assessed. Third-party form completion also means an extra complication for gathering repeated measures, where the most valid results would occur from administration to the same individual. In addition, scoring of the instrument is recommended to be conducted by individuals with psychological qualifications. A positive aspect of the feasibility of the RBPC is the short 20-minute administration time, making it suitable for repeated use within an assessment battery.

Relevance to forensic mental health research

Again, similar to the CBCL, the RBPC measures child psychopathology that may be relevant to offending behaviour, such as conduct disorder and aggression. However, there are no subscales to assess the development of delinquent offending behaviour as there are in the CBCL.

Revised Behavior Problem Checklist summary

Overall, there is little psychometric evidence for use of the RBPC within a forensic mental health population. The evidence that does exist displays that it can distinguish between delinquent and non-delinquent samples. However, although the RBPC measures child psychopathology that would be relevant to offending behaviour, a subscale measuring offending behaviour does not exist.

Symptom-Checklist-90-Revised

General information

Derogatis\(^8\)\(^4\) devised the SCL to evaluate a range of psychological problems and symptoms of psychopathology. The SCL consists of 90 items that the participant rates on a five-point scale. The items reflect nine primary symptom dimensions: anxiety, depression, hostility, interpersonal sensitivity, obsessive compulsive, paranoid ideation, phobic anxiety, psychoticism and somatisation. The SCL is used by professionals in mental health as well as medical and educational settings in addition to research purposes. It is generally administered only to individuals aged 13 years and older. The instrument should take between 10 and 20 minutes to administer.

The SCL is often used as an initial evaluation of patients for symptom assessment, measuring a patient’s progress during and after treatment, as an outcome measurement, and in clinical trials to measure change.
The SCL-90-R\textsuperscript{84} has several precursors: the Hopkins Symptom Checklist-58 (HSCL-58),\textsuperscript{85} the SCL-90\textsuperscript{86} and the HSCL-90.\textsuperscript{87} The SCL-90 expanded on the previous HSCL-58, with the HSCL-90 and the SCL-90 being almost identical. Finally, the SCL-90-R consisted of the same nine dimensions as the SCL-90, but with modification of seven items and replacement of two items. An overall score known as the Global Severity Index may be constructed from the total of the items from the SCL-90-R.

### Adequacy of measurement properties – general

The SCL-90 has displayed good internal consistency for all of the subscales\textsuperscript{88} and good test–retest reliability over a week.\textsuperscript{84} For concurrent validity the SCL-90-R subscales displayed acceptable associations with DSM-III-R diagnoses for anxiety and depression with 408 primary care outpatients.\textsuperscript{89} Further, the relevant subscales of the SCL-90 displayed good concurrent validity with associated subscales from the MMPI\textsuperscript{88} with 209 symptomatic volunteers. However, divergent validity was not as strong with many of the MMPI subscales correlating with many of the SCL-90-R subscales. Results from factor structure investigation have varied between eight\textsuperscript{90} and nine\textsuperscript{86} meaningful factors. Evidence for use of the SCL-90-R as a set of nine subscales is limited to its internal consistency, inter-rater reliability and convergent validity. Evidence against use of the nine subscales is their shared variance, and the fact that many of the subscales correlate to depression and anxiety. Consequently, some researchers consider the SCL-90 to measure one construct of general distress rather than distinct dimensions of psychopathology.\textsuperscript{91} In support of one dimension, the variance accounted for by one factor has been up to 9.25 more than that of the second factor.\textsuperscript{92} As Cyr et al.\textsuperscript{92} state, the SCL has been plagued with problems for defining consistently independent dimensions of symptom distress. Perhaps the best use of the SCL-90-R is for its Global Severity Index, which as a total score adheres to a single dimension.

### Adequacy of measurement – forensic mental health

Only one study was available assessing the psychometric properties of the SCL-90 in a forensic population.

### Concurrent validity

Wilson et al.\textsuperscript{93} examined a sample of 89 men remanded in prison awaiting trial in the hospital area. The SCL-90 displayed associations with items on the Comprehensive Psychopathological Rating Scale (CPRS), including moderate correlation between obsessional subscales (0.41) and good agreement with the depression subscales (0.62). Also, the CPRS schizophrenic subscale displayed associations with the SCL-90 subscales of psychotism (0.63), paranoid ideation (0.53) and interpersonal sensitivity (0.44). The Present State Examination displayed appropriate correlations between its 38 syndrome diagnoses and the relevant SCL-90 subscales. However, the only subscale on the SCL-90 that distinguished psychotic from non-psychotic participants was the paranoid ideation subscale ($t = 2.74, p < 0.01$), which was surprising considering the three subscales that had correlated with the CPRS schizophrenic subscale.

### Feasibility for forensic mental health research

The SCL-90-R takes a short time to administer for such a comprehensive range of dimensions. However, it is recommended to be administered by professionals, limiting its use for untrained researchers. The SCL-90 is feasible for use with a forensic population as prisoners understood the words used.\textsuperscript{93}

### Relevance to forensic mental health research

The SCL-90-R is relevant to forensic mental health research as it can assess the psychopathology of participants. Thus, its usefulness in general mental health research is readily transferable to forensic mental health.

### Symptom-Checklist-90 summary

Other than concurrent validity, the psychometric properties of the SCL-90 and its predecessors have not been evaluated using a forensic population. Although it is informative to know that the SCL-90 displays good concurrent validity in forensic samples it is the other areas of validity that have been shown to be low with non-forensic samples, such as the factor structure and the feasibility of using the subscales as different dimensions of
measurement rather than as a Global Severity Scale. Therefore, it is imperative that these areas of psychometric evaluation are conducted for a forensic mental health population.

**Self-Reported Delinquency Scale**

**General information**

The SRDS was created for use in the National Youth Survey in 1977. Participants report their delinquent activities including property damage, theft, assault and substance use. An interview was thought to produce more reliable data than self-report. The instrument consists of 47 items for which the respondent is first asked if they have committed the offence over 10 times in the last year (from the past Christmas to the previous Christmas), if so then they can choose how often (i.e. 2–3 times a day to once a month). For each type of delinquent act, the participant is also asked if other people were involved and if the participant was under the influence of alcohol or drugs at the time of the act. The means of the items are calculated to create two types of scales: offence general category scales (which refer to status offences and interpersonal violence) and summary scales (which refer to index offences and general delinquency).

The SRDS was used in the 1977 National Youth Survey, which consisted of a probability sample of households in the USA, producing a sample of 1726 youths. After 5 years of panel data, the fifth National Youth survey sample consisted of 1494 youths from the original 1726, thus displaying a high retention rate.

**Adequacy of measurement properties – general**

Huizinga and Elliott believe that using internal consistency as a measure of reliability for the SRDS is inappropriate as there have no expected links between different types or frequencies of delinquent behaviour. Test–retest reliability was conducted with 177 youths at a 4-week period. The correlations were 0.75 for the frequency score and 0.84 for the variety score. For the crime-type subscales, the reliability correlations were from 0.52 to 0.93. Content validity was considered at instrument construction where several steps were taken to ensure it was high, including offences listed by the Uniform Crime Reports and those considered relevant in the literature. Finally, concurrent validity was assessed through social trends from the SRDS with official arrest data. They found that the SRDS race/class categories displayed similar trends to the official arrest data.

**Adequacy of measurement properties – forensic mental health**

All of the psychometric properties of the SRDS were assessed using the National Youth Panel data, which consisted of a section of all youths in a community that included some delinquents. Data have not been collected and evaluated using a purely forensic mental health sample.

**Feasibility for forensic mental health research**

The SRDS is thought to be best administered as an interview rather than self-report, thus requires an interviewer. Reporting also requires a participant with a good memory, for identification of the amount of times the offence was committed over the last year. Thornberry and Krohn report a substantial amount of concealing or forgetting past criminal behaviour, producing considerable under-reporting. For example, in self-reported substance abuse, validity may be less for more serious offences involving hard drugs such as heroin than for those involving soft drugs such as tobacco and marijuana.

**Relevance to forensic mental health research**

The SRDS is purely a measure of offending behaviour, rather than an index of mental health. Thus, from a purely forensic perspective, it would be useful to assess criminal activity, yet it would not provide information about the participant’s mental health.

**Self-Reported Delinquency Scale summary**

Thornberry and Krohn believe that the SRDS appears to have acceptable content validity, and construct validity appears high, with concurrent validity being from moderate to strong. Reliability also appears quite high, although there is no evidence of differential reliability. Overall, the
SRDS appears to be better than previous self-report measures of delinquency that had questionable content validity with few items to assess the full range of criminal activity, which erred towards the trivial end.

**Summary of most frequently used outcome measures**

Several instruments were found to be used with some frequency, having been extensively used as assessments of key aspects of mental health in other more general populations: the BDI, the BPRS and the SCL-90.

Overall, not one of the outcome measures that occurred in over four studies in the review has been substantially psychometrically tested with a forensic mental health population. Further, the most frequently occurring outcome measures displayed in this review were not formulated for use in forensic mental health populations. For instance, the ASI was formulated for use with substance abusing samples, and the SCL-90 for use with psychiatric samples.

Mental health is the focus of most of the measures including the BDI, BPRS, CBCL, RBPC and SCL-90. Only the CBCL also includes a delinquency subscale to assess offending behaviour. Alternatively, the CTS and the SRDS focus solely on offending behaviour, thus not fulfilling any mental health measurement. Some measures consider both aspects of forensic mental health, with the ASI including both a legal and a psychiatric subscale, amongst five others assessing life circumstances and substance use. The wide scope of the ASI would make it the instrument that considers the most domains within forensic mental health, a topic discussed by the consensus group that is reported in the next section. Finally, the FACES considers family cohesion, which would appear to be a component of forensic mental health, although not a focal issue. In summary, the foci of mental health, offending behaviour and addiction are elements that are changeable through intervention, and thus are of major interest in forensic mental health outcome research. Instruments exist that can be considered sound assessments of these domains.

The prevalence of four measures that are aimed at youth samples including the CBC, FACES, SRDS and RBPC reflect the high proportion of studies assessing interventions aimed at young offenders. These measures may also be useful for research concerning youths that are at risk of offending. Both the CBC and the RBPC assess child psychopathology, yet neither has any significant evaluation of its psychometric properties within a forensic mental health sample, thus reliance remains with results from psychiatric samples for psychometric evaluation. It is questionable whether these psychometric results are transferable to a forensic context.

The next results section reports the findings from the panel of experts regarding their views of the most frequent outcome measures found in the structured review.

**Consensus opinion about the most frequent outcomes**

The nine most frequently occurring instruments from this structured review that have just been reported in the previous section were rated and discussed by the consensus panel.

**Outcome measures: ratings**

Only two instruments were known to the whole group, the BDI and the BPRS. The SCL-90-R was known to all but one expert. The remaining measures were not widely known, with only three people having heard of the SRDS, FACES, CTS and CBCL, and only one person knowing of the ASI and the RBPC.

Table 16 shows the group mean ratings for each instrument on the three scales of relevance, feasibility and adequacy of measurement properties. Instruments are listed in rank order for each rating. As can be seen, ratings were generally low, with few instruments receiving ratings approaching ‘very good’ on any scale.

Three instruments emerged as consistently the best in terms of all three ratings made by the group: the BDI, BPRS and SCL-90-R. Whilst all three were considered fairly or very relevant and feasible, their measurement properties in relation to forensic mental health were only considered adequate. Some instruments, such as the ASI and SRDS, were considered relevant or feasible, but were rated less favourably in terms of their measurement properties, especially in the specific forensic mental health context. Most of the remaining instruments were considered only marginally relevant and...
TABLE 16  Group mean ratings from consensus meeting (scales listed in rank order for each dimension separately)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Group mean rating(^a)</th>
<th>Scale</th>
<th>Group mean rating(^a)</th>
<th>Scale</th>
<th>Group mean rating(^a)</th>
<th>Scale</th>
<th>Group mean rating(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPRS</td>
<td>3.6</td>
<td>BDI</td>
<td>3.9</td>
<td>BPRS</td>
<td>2.5</td>
<td>BDI</td>
<td>3.6</td>
</tr>
<tr>
<td>SCL-90-R</td>
<td>3.5</td>
<td>BPRS</td>
<td>3.9</td>
<td>SCL-90-R</td>
<td>1.9</td>
<td>ASI</td>
<td>3.0</td>
</tr>
<tr>
<td>BDI</td>
<td>3.3</td>
<td>SCL-90-R</td>
<td>3.3</td>
<td>SCL-90-R</td>
<td>1.9</td>
<td>SRDS</td>
<td>3.3</td>
</tr>
<tr>
<td>ASI</td>
<td>3.0</td>
<td>SRDS</td>
<td>1.9</td>
<td>SRDS</td>
<td>1.6</td>
<td>ASI</td>
<td>3.0</td>
</tr>
<tr>
<td>CTS</td>
<td>2.6</td>
<td>ASI</td>
<td>2.5</td>
<td>RBPC</td>
<td>1.3</td>
<td>CTS</td>
<td>1.3</td>
</tr>
<tr>
<td>SRD</td>
<td>2.6</td>
<td>CTS</td>
<td>2.0</td>
<td>ASI</td>
<td>1.3</td>
<td>RBPC</td>
<td>1.6</td>
</tr>
<tr>
<td>CBCL</td>
<td>2.3</td>
<td>CBCL</td>
<td>1.9</td>
<td>ASI</td>
<td>1.3</td>
<td>CTS</td>
<td>1.4</td>
</tr>
<tr>
<td>FACES</td>
<td>1.6</td>
<td>FACES</td>
<td>1.9</td>
<td>FACES</td>
<td>1.3</td>
<td>RBPC</td>
<td>1.6</td>
</tr>
<tr>
<td>RBPC</td>
<td>1.6</td>
<td>RBPC</td>
<td>1.6</td>
<td>CBCL</td>
<td>1.2</td>
<td>FACES</td>
<td>1.3</td>
</tr>
</tbody>
</table>

\(a\) Scale ranged from 1 (not), 2 (slightly), 3 (fairly) and 4 (very) to 5 (extremely).  
\(b\) This scale was added during the meeting at the recommendation of the group.

feasible with adequate to poor measurement properties.

The final column of Table 16 was added on the recommendation of the group. It reflects the view that the scales should also be rated in terms of their general measurement properties, to acknowledge the possibility that some may not have been assessed in specific relation to forensic mental health populations. This is captured by the somewhat higher ratings (better properties) in this general column, particularly for those scales ranked towards the top.

Outcome measures: group discussion

The group agreed that most of the domains previously identified and confirmed as important were not represented by the current selection of instruments. There was agreement that there were many potential outcome measures that to date have been used only as predictors or measures of process and that future work should recognise their potential as outcome measures in trials and evaluative studies. It was felt that candidate measures of outcome could be found in existing measures of impulse control, antisocial attitudes, aggression, emotional control, impulsivity, socialisation, self-awareness, severity of opiate dependence and alcohol use. Specific instruments suggested include the Psychological Inventory of Criminal Thinking Styles,\(^{98}\) the Anti-Social Activities Attitude Scale;\(^{99}\) the Criminal Sentiments Scale;\(^{100}\) the Barratt Impulsivity Scale;\(^{101}\) the Novaco Anger Scale;\(^{102}\) and the Self-Appraisal Questionnaire.\(^{103}\)

The group felt that risk assessment tools such as Historical, Clinical, and Risk Management Scales (HCR-20)\(^{104}\) and the Violence Risk Scale\(^{105}\) offered a particularly promising source for outcome measures. However, there was little formal evidence of their use in this context other than as predictive tools.

The discussion highlighted the need for more research to establish validity and relevance both for instruments reviewed here and for other suggested outcomes of interest. There is a need for this to focus specifically on forensic mental health populations, as extrapolating from general population psychometrics may be invalid. There was a notable absence of ‘positive’ measures able to reflect desirable rather than undesirable outcomes.

The outcome measure of recidivism

By far the most prevalent outcome variable used in the eligible studies in this structured review was some form of offending behaviour or recidivism, occurring in 72% (\(n=223\)). The domain of
Recidivism is considered further in terms of the data collected in the structured review and discussion on the topic in the consensus panel.

**Recidivism measures from the structured review**

*Table 17* displays how the different studies recorded offending behaviour. In the true sense of the term recidivism, the legal indices of reoffending included the legal process from contact with the police to time spent incarcerated and violations whilst on parole. The most frequent of these measurements was arrest, followed by conviction. Most frequently, the type of offence measured was unspecified. About a third of the arrest measures and just under half of the conviction measures specified a violent or sexual offence, thus determining a more specific mode of reoffending. The presence of specific violent or sexual offending behaviour also occurred only for criminal behaviour, arrest, charge, conviction and offence, displaying a narrower range of measurement than for any unspecified type of offence. This pattern was similar for other specified types of offence that were not violent or sexual with only the addition of parole violation as the index of criminal behaviour. A very non-specific measure of recidivism was displayed in 42 measures of reoffending behaviour in studies that used only the definition of reoffence, not specifying at what point during the legal process data were collected. Participant reports of offending behaviour that did not reach legal attention were only the fifth most prevalent measurement type for unspecified offence types, yet were the fourth most prevalent in specifically violent or sexual offences and second with specified offences that were not violent or sexual. These results show that for specific types of offences, forensic mental health researchers are interested in actual criminal behaviour, whereas more general offending that is not specified invites more legally defined measures. Use of measures of general offending may reflect use of databases where classification of specific offence types may not be available, or conceptually difficult.

The source of the recidivism data was most frequently official records; for example, the Home Office Offender Index, or state/national records in the USA (*Table 18*). Sixteen studies used both official records and self-reports, which would provide a method of validation for both sources: crimes committed and not detected for official records, and crimes not admitted to or forgotten in self-report.

**TABLE 17** Different forms of recidivism measurement by offence type

<table>
<thead>
<tr>
<th>Index of criminal behaviour</th>
<th>Any type of offence/ unspecified</th>
<th>Specifically a violent or sexual offence</th>
<th>Other type of offence categorisation (specified, not violent or sexual offence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criminal behaviour</td>
<td>25</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Contact with police</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Arrest</td>
<td>59</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>Charge</td>
<td>19</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Court</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Conviction</td>
<td>40</td>
<td>25</td>
<td>9</td>
</tr>
<tr>
<td>Parole violation</td>
<td>30</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Community</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Institutional</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Time incarcerated</td>
<td>26</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Time until incarceration</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Offence – general (no specific stage of criminal process)</td>
<td>24</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### TABLE 18  Source of offending behaviour data

<table>
<thead>
<tr>
<th>Source of data</th>
<th>Official records</th>
<th>Self-report</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of times used</td>
<td>172</td>
<td>46</td>
<td>21</td>
</tr>
</tbody>
</table>

### Consensus group discussion about recidivism

Finally, the group returned to general discussion about the use of recidivism as an outcome measure. An enormous range of indicators have been used including convictions, arrest, court appearances and revocation of parole. Different indicators reflect levels of severity of the recidivist behaviour. Which indicator to use depends on the precise study aims, population and context. It was felt that there was promising evidence that self-reported offending behaviour can be accurate.¹⁰³

Recidivism itself was noted as being a proxy measure. Reconviction, arrests and other indicators will only be a sampling of the true frequency of antisocial acts. In addition these indicators are unlikely to be pure. For example, incidents in addition to violent acts may lead to revocation of parole.

An inherent problem was noted for outcomes involving severe offending behaviours, namely that these are usually rare. Use of rare behaviours (such as homicide) as outcomes can be problematic owing to statistical issues of power. Lower level crimes could be considered precursors of more serious offences, suggesting that the former could be useful as proxy outcomes for more serious behaviours.

More objective measures such as recidivism were contrasted to intervening variables, such as aggressive interpersonal style and other psychological and social measures. Both were considered important for future research.
This report has consisted of two distinct but related components. Firstly, a structured review of forensic mental health outcome research was carried out to assess the use of outcome measures. Secondly, the literature review was supported and supplemented by a multidisciplinary consensus process that identified and rated the importance of different domains that might be assessed in terms of outcomes in forensic mental health research and then judged the most frequently used outcome measures identified from the structured review.

There are some limitations to the study. First, only references in the English language were examined in the review, which may have caused biases due to sampling. In a similar vein the consensus group consisted of participants from the UK only. It is not possible to estimate to what extent an international consensus group would have reached similar conclusions. Second, the review considered only references gleaned from electronic databases, and owing to time constraints did not include dissertations. There is a body of research conducted by international justice departments such as the Home Office and New Zealand, Australian and Canadian corrections departments that would also provide eligible studies. These justice sources were also not included in the review because of time constraints, although many of them would have appeared in our electronic database search owing to subsequent journal publication. Finally, not all of the references marked as relevant at the abstract stage were accessible as hard copies. It is nevertheless difficult to believe that, even with these acknowledged methodological limitations, outcome measures were omitted from the review that might have proved more robust and more frequently used than those that were identified and assessed in the study. In support of this speculation, none of the consensus panel was able to identify any outcome measures that were considered important but overlooked by the review process.

As was found in the review of trials on aggressive and violent people by Cure et al., there was a large presence of studies from the USA in the current review.

The majority of studies forming the basis for the current review used an RCT methodology. However, the sample of eligible studies was significantly increased by the inclusion of cohort and other comparative designs as long as they consisted of at least a 6-month follow-up and a comparative (intervention versus control) design. This more inclusive approach was adopted to ensure that the review would assess all outcome measures commonly used by or familiar to the forensic mental health community. The common occurrence of non-RCT studies reflects the difficulties of adhering to ‘gold-standard’ methodological approaches to evaluation in the forensic mental health context. Overall, a typical forensic mental health outcome study would be conducted in the community, with a male adult sample of between 101 and 200 participants who had committed an offence and received cognitive behavioural therapy.

Considering that this review focused on forensic mental health, details of mental health diagnosis in study samples were uncommon. Most attention in studies was given to the offending behaviours. However, even with offence, details were often not specified. Given that this lack of details was very common in the literature, it seemed sensible not to exclude them from analysis, given the focus on quality of outcome measurement rather than details of study samples per se.

In this review of 308 different studies of forensic mental health outcome research, the number of different variables used to assess outcome was very large at 744. A previous review of trials for seriously mentally ill violent offenders identified 345 different measures from 300 trials. The large number of variables used to assess outcomes in forensic mental health research must create problems in terms of comparing results of interventions. It may indeed impede the development of common understanding of the scale and nature of benefits of interventions if there is so little shared and commonly used measures.
A useful typology of outcomes in mental health research in general was produced by Atkisson et al. They argued that mental health research needs to be multidimensional in perspective on outcomes, and proposed a fourfold typology of domains as a framework:

1. The clinical domain. Here outcomes are concerned with signs and symptoms of mental illness and health status more broadly including mortality and morbidity.
2. The rehabilitation domain. Here outcomes are focused on adaptation and function, especially in terms of social function (e.g. interpersonal relations, social integration) and instrumental functioning (e.g. problem solving, work, education).
3. The humanitarian domain. This domain would include assessment of outcomes in terms of quality of life and well-being, and experiences of and satisfaction with services.
4. The public safety domain. This domain is concerned with societal rights to public safety and the balance between individual rights and community perceptions of safety.

In an overview of policy and research in forensic mental health, Cohen and Eastman expressed the view that the majority of forensic mental health research focused on outcomes in the fourth domain, described by Atkisson et al. as public safety. In particular, they argued that there was substantial focus on outcomes in terms of recidivism, especially with regard to re-arrest and recollection rates.

The current review provides clear evidence to support the view expressed by Cohen and Eastman. Recidivism is by far the most commonly measured domain used to assess outcomes in forensic mental health research; either explicitly identified as recidivism (non-violent, violent or sexual) or implicitly focused on recidivism in terms of substance abuse. Cohen and Eastman argued that the focus on recidivism results in neglect of the clinical, rehabilitation and humanitarian domains of forensic mental health. Indeed they argue more controversially that the degree of emphasis upon public safety reinforces a ‘separatist’ tendency in forensic mental health research in relation to general mental health research.

The vast array of recidivism outcome measures displayed in this review illustrates the problems posed for comparison of results between different studies. Clearly, the term recidivism is not sharply defined and operationalised, ranging from offending behaviour, through parole violations to incarceration. Falshaw et al. provided a practical example of the problems for comparison of different measures of recidivism. They found that the rate of recidivism increased by a factor of 5.3 when measured by any offence-related behaviour in treatment programme files in sexual offenders compared with measuring reconviction using the Home Office Offenders Index database. Similarly, a study in the USA showed major differences in the estimated rate of violent sexual offence depending on which official criminal record was used. Grann et al. expressed concern about the continued practice in forensic mental health research of ‘lumping’ together behaviours of extremely different levels of seriousness into outcome measures of recidivism.

The next most frequently used variables to assess outcomes were a variety of measures of mental health and cognitive or psychological function, reflecting the distinctive needs and forms of intervention most likely to be encountered in a mental health population. As will be discussed below, only a small number of such measures were used with any frequency. The wide array of scales to assess mental health has already been commented upon and was noted by Cure et al. This extreme diversity of instruments in use to assess mental health does not facilitate the emergence of shared understanding of the effectiveness of forensic mental health interventions.

Few studies in the database were found to assess broader aspects of health status, well-being, social function and quality of life. The consensus panel stated that many potentially important domains of outcome appear neglected in evaluative studies. Although a number of such measures have been developed and validated and provide multidimensional measures of outcome from the respondent’s perspective, they have not been taken...
up in forensic mental health research. This is not entirely surprising as the same is true for mental health research generally as well as for routine practice where patient-reported outcome measures have not been widely adopted.111

Studies involving cost-effectiveness methods and outcomes were also uncommon. Given potential opportunities for cost savings from, for example, prevention of institutionalised custody or care of mentally offenders, it is surprising that health economic methods have not been more widely adopted in research in this field.

The review set out, where feasible, to assess the measurement properties of outcome measures that were used with any degree of frequency in forensic mental health research. With regard to specific outcome instruments, a cut-off was set that the research team would examine in more detail any instrument that emerged from the review as having been used in at least five separate studies in the database. Only nine instruments were found that fulfilled this requirement. This is evidence in another form of the extent to which the field of forensic mental health research lacks outcome instruments that are commonly enough used that the forensic mental health community would be familiar with the instruments and be readily able to interpret their results.

Three measures assessing broad aspects of mental health were found that were used with a reasonable level of frequency in the sample of studies – the BDI, the BPRS and the SCL-90-R. These are widely used measures of dimensions of mental health in the broader area of mental health research and have been satisfactorily assessed (especially the BDI and the BPRS) for measurement properties in that wider context. It is not surprising that these were the three outcome measures that were known to the consensus group, and were the only measures that were scored consistently positively by the consensus panel for relevance, feasibility and adequacy of measurement properties for forensic mental health research.

In addition, the ASI emerged as a commonly used measure of outcome for addiction interventions with significant supportive evidence for its measurement properties and moderately consistent support from the consensus panel’s ratings. It is reported to require a trained interviewer and nearly an hour to administer, so may only marginally qualify as an instrument to be thought of as feasible for large-scale use in pragmatic trials.

Some other instruments have good supportive evidence for use in very specialist contexts, e.g. the CBCL to assess delinquent behaviours in children. Otherwise, instruments that emerged from the literature review were very poorly supported for measurement properties (e.g. the RBPC) or very specialist in their range of application, e.g. the FACES to assess family cohesion. The panel were not very supportive of the role of such instruments in forensic mental health research.

In view of the paucity of robust outcome measures emerging from the literature review, the database of studies was re-examined with a lower threshold; in this second review, outcome measures were now examined if at least three studies were found to use an instrument. This reanalysis did not yield a single instrument that had been highlighted as promising in the discussion of the consensus meeting. This check on the database strengthened the confidence of the research group that no major source of evidence had been omitted.

It is in the nature of reviews to look backwards. What is largely missing from the review is any substantial evidence of recent debates in forensic mental health research about risk assessment tools. They have not featured as outcomes in trials or evaluative research to any significant degree. Some commentators are beginning to raise questions as to whether approaches to assessing risk of violence have a greater role in evaluative research.110,112 They were cited by members of the consensus panel as promising for use in evaluative research.

Partly in response to public disquiet and clamour for better decisions about potentially dangerous mental health clients, a large amount of effort has gone into research to better predict individuals who are more at risk of future violent behaviour in forensic mental health services. Risk models partly comprise static variables that may predict violence, for example, demographic or socioeconomic variables. They also include more dynamic variables, for example, attitudes, orientation and treatment engagement, that may also be predictive of violence. These dynamic variables, because they can and do change, are of particular interest because they may not only be predictive of violence, but may also be responsive to interventions. Crucially, they would be of greatest interest if they respond to interventions and are causally associated with subsequent reduction in violent behaviour in a causal chain.113 The evidence informing this area of forensic mental health research is complex and is still a work in progress.
There are competing instruments the relative merits of which are as yet unclear, including the HCR-20, the Level of Service Inventory-Revised, the Violence Risk Scale, the Structured Outcome Assessment and Community Risk Monitoring. Similarly, it is far from clear which specific dynamic risk factors are critical in the causal chain; candidates include impulsiveness, antisocial attitudes, substance abuse and treatment provider alliance.

This array of constructs and measures offers significant potential for targeting of interventions, for monitoring changes over time in key causal variables and ultimately having a positive impact on violent and serious criminal behaviours. There is growing evidence of their translation across from US to UK settings and of their validity to UK populations in terms of observational predictive applications. However, there is a dearth of research using such instruments as outcome measures in trials, and the design and conduct of such studies are likely to be highly challenging. Small numbers of eligible subjects and logistical difficulties of mounting multicentre RCTs in settings focused on security are among more obvious difficulties.
A wide range of domains are relevant to assessing outcomes of interventions in forensic mental health services. Evaluations need to take account of public safety, but also clinical, rehabilitation and humanitarian outcomes. To date, research has focused extensively on the first domain, evaluating outcomes in terms of recidivism.

Recidivism is a very high priority; the public expects interventions that will reduce future criminal behaviour. The very wide range of variables used to assess recidivism makes it difficult to draw conclusions across studies using different variables. It is difficult to see how complete standardisation of measures can be achieved given the enormous variation over time and across countries in systems of criminal justice. However, it is conceivable that more research could be productive to address the heterogeneity of seriousness of forms of recidivism in outcome measurement. Research to assess the validity of self-report measures of recidivism is another priority.

Mental health is clearly also an important dimension of outcome. Instruments have been used in forensic mental health research that have been well validated in the context of general mental health research. Much of the evidence of their use is based on studies carried out in the USA, so that it is not always clear that evidence of performance can be translated automatically to apply to their use in the context of the UK.

The review provides clear support for the view that domains such as quality of life, social function and psychosocial adjustment have not been extensively employed in forensic mental health research but are relevant and important issues. These are important domains for forensic mental health research, and the role of such instruments needs more consideration. Research is needed in these domains to complement the evidence base of outcomes in terms of public safety and mental health.

The wide array and diversity of measures used in forensic mental health research suggest that there is still substantial scope for standardisation, by further use of consensus-type processes to identify domains and specific measures that are relevant and familiar in practice and can be more widely used in evaluative research.

The role of instruments assessing dynamic aspects of risk of violence offer a particular opportunity. They are becoming more widely known in practice. There is growing confidence in their role in predicting the risk of subsequent offending and other key outcomes. There is a lack of any evidence to recommend that any particular measures of risk could also be used as outcome measures, but it should be a priority for the field to apply and assess their potential in a longitudinal context for the purposes of evaluative research.
Acknowledgements

The authors wish to thank the contributors to the consensus process: Richard Benson, Danny Clarke, Simon Draycott, Connor Duggan, Cathriona Hynes, James McGuire, Maria Leitner and Caroline Logan.

Financial support from the NHS Research & Development Methodology programme is also gratefully acknowledged.

Contributions of authors

Ray Fitzpatrick is Professor of Public Health and Primary Care, University of Oxford; he had overall responsibility for the conduct and delivery of the project and writing-up of results. Jemma Chambers is Research Officer, Department of Public Health, University of Oxford; she oversaw the day-to-day running of structured review and forensic psychology research issues and played the lead role in drafting the results. Tom Burns is Professor of Social Psychiatry, University of Oxford; he oversaw aspects of review and consensus relating to general mental health research issues. Helen Doll is Statistician, Health Service Research Unit, University of Oxford; she oversaw statistical issues in the analysis of results and write-up of the report. Seena Fazel is Clinical Senior Lecturer in Forensic Psychiatry, University of Oxford, and Honorary Consultant Psychiatrist at the Oxford Clinic; she oversaw forensic and general mental health research issues. Crispin Jenkinson is Professor of Health Service Research, University of Oxford; he took the lead regarding psychometric aspects of outcomes assessment. Asha Kaur is Research Assistant, Department of Public Health, University of Oxford; she conducted database maintenance and reference hard copy retrieval and was responsible for the checking of all references. Martin Knapp is Director of the Personal Social Services Research Unit at the London School of Economics, Professor of Social Policy, LSE, and Professor of Health Economics and Director of the Centre for the Economics of Mental Health, the Institute of Psychiatry, King’s College London; he led the interpretation of health economic studies. Lesley Sutton is Research Assistant, Department of Public Health, University of Oxford; she conducted the structured review and database maintenance and wrote up the research strategy. Jenny Yiend is Senior Research Fellow, Department of Psychiatry, University of Oxford; she ran the consensus group and drafted its results, and oversaw the interpretation of cognitive and psychological outcomes.
References


## Appendix I

### Search strategies by database

#### CINAHL

| Search Strategy |
|-----------------|-----------------|
| 1               | **CONVICT$** or **CRIMIN$** or **DELIQUEN$** or **FELON$** or **INCARCERAT$** or **INMATE$** or **OFFEND$** or **PAROLE$**. **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 2               | **BORSTAL$** or **GAOL$** or **JAIL$** or **PENAL$** or **PRISON$** or **PROBATION$** or **REMAND$**. **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 3               | **BOOT CAMP$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 4               | **COMMUNIT$** adj2 **CORRECTION$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 5               | **CORRECTION$** adj3 **PROGRAM$** or **FACILITY$** or **SERVICE$**. **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 6               | **CORRECTIONAL$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 7               | **FORENSIC** adj3 **UNIT$** or **HOSPITAL$**. **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 8               | **DETECTION CENT$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 9               | **SECURE** adj2 **HOSPITAL$** or **INSTITUT$** or **TRAINING CENT$** or **FACILITY$$. **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 10              | **THERAPEUTIC COMMUNIT$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 11              | **YOUTH CUSTOD$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 12              | **YOUTH OFFEND$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 13              | **EXP PRISONERS/** |
| 14              | **EXP PUBLIC OFFENDERS/** |
| 15              | **EXP CORRECTIONAL FACILITIES/** |
| 16              | or/1–15 |
| 17              | ((cohort or follow up or follow?up or longitudinal or prospective or retrospective or control or case?control) adj stud$.** **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 18              | **SINGLE$** or **DOUBL$** or **TRIPLE$** adj **BLIND$** or **MASK$**. **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 19              | **RANDOMI$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 20              | **RANDOM$** adj **ALLOCAT$** or **ASSIGN$** or **MASK$**. **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 21              | ((cross over or cross?over) adj stud$.** **mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 22              | **PLACEBO$$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 23              | **REPEAT$** measure$.mp.** [mp=title, subject heading word, abstract, instrumentation] |
| 24              | **EXP CROSSOVER DESIGN/** |
| 25              | **EXP CLINICAL TRIALS/** |
| 26              | **EXP PROSPECTIVE STUDIES/** |
| 27              | **EXP REPEATED MEASURES/** |
| 28              | **EXP NONRANDOMIZED TRIALS/EXP PRETEST-POSTTEST DESIGN/** |
| 29              | **EXP META ANALYSIS/** |
| 30              | **EXP “SYSTEMATIC REVIEW”/** |
| 31              | or/17–30 |
| 32              | 31 and 16 |

### Cochrane

| Search Strategy |
|-----------------|-----------------|
| 1               | **CONVICT$** or **CRIMIN$** or **DELIQUEN$** or **FELON$** or **INCARCERAT$** or **INMATE$** or **OFFEND$** or **PAROLE$**:**ti,ab,kw** or **BORSTAL$** or **GAOL$** or **JAIL$** or **PENAL$** or **PRISON$** or **PROBATION$** or **REMAND$**:ti,ab,kw |
| 2               | **FORENSIC** near **UNIT$** or **HOSPITAL$**:ti,ab,kw |
| 3               | **SECU RE** near **HOSPITAL$** or **INSTITUT$** or **UNIT$****:ti,ab,kw |
| 4               | **YOUTH CUSTOD$**:ti,ab,kw |
| 5               | **YOUTH OFFEND$**:ti,ab,kw |

### MeSH descriptor Prisoners, this term only

| Search Strategy |
|-----------------|-----------------|
| 1               | (1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10) |
EMBASE
1 (convict$ or crimin$ or delinquen$ or felon$ or incarcerat$ or inmate$ or offend$ or parole$). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
2 (borstal$ or gaol$ or jail$ or penal or penol$ or penitentia$ or prison$ or probation$ or remand$). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
3 boot camp$. mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
4 (communit$adj2 correction$). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
5 (correction$adj3 (program$ or facilit$ or service$)). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
6 correctional$. mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
7 (forensic adj3 (unit$ or hospital$)). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
8 detention cent$. mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
9 (secure adj2 (hospital$ or institut$ or unit$ or training cent$ or facilit$)). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
10 therapeut$communit$. mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
11 youth custod$. mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
12 young offen$. mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
13 exp offender/
14 exp prisoner/
15 exp custody/or exp detention/or exp prison/or exp probation/
16 exp Criminal Justice/or exp custody/
17 or/1–16
18 ((cohort or follow up or follow?up or longitudinal or prospective or retrospective or case control or case?control) adj stud$). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
19 ((singl$or doubl$or trebl$or tripl$) adj (blind$ or mask$)). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
20 randomi$. mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
21 (random$adj (assign$ or mask$ or allocat$)). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
22 ((cross over or cross?over) adj stud$). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
23 (placebo$ or repeat$ measure$). mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
24 exp cohort analysis/
25 exp case control study/or exp longitudinal study/or exp prospective study/or exp retrospective study/
26 exp crossover procedure/or exp double blind procedure/or exp single blind procedure/or exp randomized controlled trial/
27 exp meta analysis/or exp “systematic review”/
28 exp clinical trial/or exp multicenter study/or exp phase 1 clinical trial/or exp phase 2 clinical trial/or exp phase 3 clinical trial/or exp phase 4 clinical trial/
29 exp randomized controlled trial/
30 or/18–29
31 17 and 30

MEDLINE
1 (convict$ or crimin$ or delinquen$ or felon$ or incarcerat$ or inmate$ or offend$ or parole$). mp. [mp=title, abstract, name of substance word, subject heading word]
2 (borstal$ or gaol$ or jail$ or penal or penol$ or penitentia$ or prison$ or probation$ or remand$). mp. [mp=title, original title,
abstract, name of substance word, subject heading word]
3 boot camp$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
4 (communit$adj3 correction$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
5 (correction$adj3 (program$or facilit$or service$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
6 correctional$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
7 (forensic adj3 (unit$or hospital$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
8 detention cent$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
9 (secure adj2 (hospital$or institut$or unit$or training cent$or facilit$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
10 therapeut$communit$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
11 youth custod$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
12 young offen$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
13 exp Prisoners/
14 Prisons/
15 or/1–14
16 ((cohort or follow up or follow?up or longitudinal or prospective or retrospective or case control or case?control) adj stud$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
17 (singl$or doubl$or tripl$or trip$ adj (blind$or mask$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
18 randomi$.mp. [mp=title, original title, abstract, name of substance word, subject heading word]
19 (random$adj (assign$or mask$or allocat$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
20 ((cross over or cross?over) adj stud$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
21 (placebo$or repeat$measure$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
22 exp case-control studies/or exp cohort studies/
23 exp Clinical Trials/
24 exp intervention studies/
25 exp cross-over studies/or exp double-blind method/or exp matched-pair analysis/or exp meta-analysis/or exp random allocation/or exp single-blind method/
26 or/16–25
27 26 and 15

NCIRS

((KW=(randomi* or placebo*)) or (KW=(repeat* measure*)) or (KW=((singl* or doubl* or trebl* or tripl*) within 2 (blind* or mask*))) or (KW=(cross over or cross-over or crossover) within 2 stud*)) or (KW=(random* within 2 (assign* or allocat*))) or (KW=(clinical* within 2 (trial* or study*))) or (KW=((follow up or follow-up or followup) within 1 stud*)) or (KW=((longitudinal or prospective or retrospective) within 1 stud*)) or (KW=((case control or case-control or casecontrol) within 1 stud*) or (cohort stud*)) or (KW=(systematic review*)) or (KW=(meta analys* or meta-analys* or metanalys*)) and ((KW=((convict* or crimin* or delinquen*)) or (felon* or incarcerat* or inmate*) or (offend* or parole*)) or (KW=((boot camp* or borstal*) or (correctional* or detention cent*) or (gaol* or jail*)) or KW=((penal or penitentia*)) or (penol* or prison* or probation*) or (remand* or youth custod*)) or KW=(therapeut* communit*)) or (KW=(communit* within 2 correction*)) or (KW=(correction* within 3 (program* or facilit* or service*))) or (KW=(forensic* within 3 (unit* or hospital*))) or (KW=(secur* within 2 (hospital* or institut* or unit* or training cent* or facilit*)))))

PHI

(crimin* or incarcerat* or inmate* or offend* or correctional* or jail* or penal or prison* or remand* or therapeut* or communit*)

PsycINFO

1 (convict$or crimin$or delinquen$or felon$or incarcerat$or inmate$or offend$or parole$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
2 (borstal$or gaol$or jail$or penal or penol$or penitentia$or prison$or probation$or remand$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
Appendix 1

3 boot camp$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
4 (communit$adj2 correction$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
5 (correction$adj3 (program$or facilit$or service$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
6 correctional$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
7 (forensic adj3 (unit$or hospital$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
8 detention cent$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
9 (secure adj2 (hospital$or institut$or unit$or training cent$or facilit$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
10 therapeut$communit$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
11 youth custod$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
12 young offen$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
13 exp parole/or exp probation/
14 correctional institutions/or exp prisons/or exp reformatories/or exp halfway houses/or exp maximum security facilities/
15 exp perpetrators/or exp criminals/
16 exp criminals/
17 exp incarceration/
18 or/1–17
19 ((cohort or follow up or follow?up or longitudinal or prospective or retrospective or case control or case?control) adj stud$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
20 ((singl$or doubl$or trebl$or tripl$) adj (blind$or mask$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
21 random$8.mp. [mp=title, abstract, heading word, table of contents, key concepts]
22 (random$adj (allocat$or assign$or mask$)).mp. [mp=title, abstract, heading word, table of contents, key concepts]
23 ((cross over or cross?over) adj stud$).mp. [mp=title, abstract, heading word, table of contents, key concepts]
24 placebo$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
25 repeat$measure$.mp. [mp=title, abstract, heading word, table of contents, key concepts]
26 exp cohort analysis/
27 exp clinical trials/
28 exp longitudinal studies/or exp prospective studies/or exp followup studies/or exp prospective studies/
29 exp meta analysis/
30 exp repeated measures/
31 or/19–30
32 31 and 18

Sociological Abstracts

((crimin*) or (delinquen*) or (felon*) or (incarcerat*) or (inmate*) or (offend*) or (parole*) or (DE=(“offenders” or “career criminals” or “criminally insane” or “drug offenders” or “female offenders” or “juvenile offenders” or “sex offenders’)) or (convict*)) or ((boot camp*) or (borstal*) or (communit* within 2 correction*) or (correction* within 3 (program* or facilit* or service*)) or (correctional*) or (detention cent*) or (forensic within 3 (unit* or hospital*)) or (gaol*) or (jail*) or (penal) or (penitentia*) or (penol*) or (prison*) or (probation*) or (remand*) or (secur* within 2 (hospital* or institut* or unit* or training cent* or facilit*)) or (therapeut* communit*) or (youth custod*) or (DE=“correctional system”) or (DE=“imprisonment”) or (DE=“juvenile correctional institutions” or “correctional system”)) or (DE=“parole”) or (DE=“prisons”) or (DE=“probation”) or (DE=“detention”)) and ((random$) or (singl$ or doubl$ or trebl$ or tripl$ within 2 (blind$ or mask$)) or (placebo*) or (crossover or cross over or cross*over or cross?over within 2 stud$) or (random* within 2 (assign$ or allocat*)) or (cohort*) or (longitudinal) or (repeat$ measure*) or (follow up or follow?up or follow*up) or (prospective) or (retrospective) or (case control or case?control or case*control) or (DE=“cohort analysis”) or (DE=“longitudinal studies”)) or (DE=“random samples”))
Appendix 2

Data extraction form
# Outcomes in Forensic Mental Health

## Administration Details

<table>
<thead>
<tr>
<th>Paper ID no</th>
<th>Study no</th>
<th>No of studies in paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Extractor initials**

*Throughout use: 888 = not applicable, 999 = not stated*

<table>
<thead>
<tr>
<th>Type of report</th>
<th>1 = Journal article</th>
<th>2 = Book/chapter</th>
<th>3 = Conference</th>
<th>4 = Dissertation</th>
<th>5 = Govt. report</th>
<th>6 = Other (specify)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Published or not?</td>
<td>0 = no</td>
<td>1 = yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**First author:**

**Study name:**

**Year of publication:**

*(Combine these to give a unique name to the paper)*

**Number of studies included in this paper:**

*(if more than one, complete separate extraction forms for each, and display study no’s above)*

**Paper numbers of other studies with which this paper may link:**

*(if other papers report further results of this trial, incorporate them onto this form and note here what has been done)*

**Country of origin**

1 = USA  
2 = Canada  
3 = UK & Eire  
4 = Other European  
5 = Mid E/Asia  
6 = Africa  
7 = Australia/NZ  
8 = Latin America

**Study Design**

<table>
<thead>
<tr>
<th>Type of study</th>
<th>1 = RCT</th>
<th>2 = Other comparative designs</th>
<th>3 = Cohort study</th>
</tr>
</thead>
</table>

### Study setting

*in full*

<table>
<thead>
<tr>
<th>Study setting</th>
<th>1 = Community</th>
<th>2 = Remand</th>
<th>3 = Prison</th>
<th>4 = Probation</th>
<th>5 = Secure forensic hospital</th>
<th>6 = Juvenile centre</th>
<th>7 = therapeutic community</th>
<th>8 = Other (specify)</th>
</tr>
</thead>
</table>

### Participants

#### 1. Sample Size

<table>
<thead>
<tr>
<th>Entire study N</th>
<th>Males %</th>
<th>Males N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 2. Age

<table>
<thead>
<tr>
<th>Adolescent</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = yes</td>
<td>1 = yes</td>
</tr>
<tr>
<td>0 = no</td>
<td>0 = no</td>
</tr>
</tbody>
</table>

#### 3. Sample criminal and psychiatric history targeted by intervention

<table>
<thead>
<tr>
<th>Sample characteristics</th>
<th>Specify</th>
<th>Coding</th>
<th>Page Table no Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criminal history</td>
<td></td>
<td>1 = yes</td>
<td>0 = no</td>
</tr>
<tr>
<td></td>
<td>1 = Any offence/felony/not stated</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 = Violent offence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 = Sexual offence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 = Property offence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 = Drugs offence/use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6 = Driving offence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7 = Other specify</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatric diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 = Personality disorder</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 = Schizophrenia</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 = Affective disorder</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 = Substance abuse</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 = Sexual disorder</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6 = Behaviour disorder</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7 = Neurotic problem</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8 = Organic brain disorder</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9 = Dementia</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 = Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning disability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 = IQ below 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 = Organic brain damage</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 = Autism</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 = Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Intervention

<table>
<thead>
<tr>
<th>Intervention name</th>
<th>Type of intervention (see coding below)</th>
<th>Page no</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Intervention type codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cognitive/behavioural</td>
</tr>
<tr>
<td>2</td>
<td>family therapy</td>
</tr>
<tr>
<td>3</td>
<td>one to one psychotherapy</td>
</tr>
<tr>
<td>4</td>
<td>community penalty</td>
</tr>
<tr>
<td>5</td>
<td>prison penalty</td>
</tr>
<tr>
<td>6</td>
<td>strict daily regime</td>
</tr>
<tr>
<td>7</td>
<td>physical training</td>
</tr>
<tr>
<td>8</td>
<td>acupuncture</td>
</tr>
<tr>
<td>9</td>
<td>diet</td>
</tr>
<tr>
<td>10</td>
<td>vocational</td>
</tr>
<tr>
<td>11</td>
<td>educational</td>
</tr>
<tr>
<td>12</td>
<td>relaxation/meditation</td>
</tr>
<tr>
<td>13</td>
<td>yoga</td>
</tr>
<tr>
<td>14</td>
<td>therapeutic community</td>
</tr>
<tr>
<td>15</td>
<td>mental health court</td>
</tr>
<tr>
<td>16</td>
<td>other (specify)</td>
</tr>
</tbody>
</table>

### Outcome

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Longest follow-up period for outcome measurement (months)</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(over 6 months for non-RCT)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 = no</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 = yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specify if yes</td>
</tr>
<tr>
<td>1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Recidivism Outcome measures

<table>
<thead>
<tr>
<th>Type of recidivism (i.e. return to prison, conviction, arrest)</th>
<th>Data source (i.e. police records, self-report)</th>
<th>Follow-up periods for outcome measurement (months) (over 6 months for non-RCT)</th>
<th>Modification 0 = no 1 = yes Specify if yes</th>
<th>Page no</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Any further comments on study

Describe
Appendix 3
Included studies reference list


Dutton DG, Bodnarchuk M, Kropp R, Hart SD, Ogloff JRP. Wife assault treatment and criminal recidivism:


Lamb HR, Weinberger LE, Reston-Parham C. Court intervention to address the mental health needs of mentally ill offenders. *Psychiatr Serv* 1996;47:273–81.


Little M, Kogan J, Bullock R, Van Der Laan P. ISSP: An experiment in multi-systemic responses to persistent...


St.Lawrence J, Crosby RA, Belcher L, Yazdani N, Brasfield TL. Sexual risk reduction and anger


Health Technology Assessment reports published to date

**Volume 1, 1997**

No. 1  
Home parenteral nutrition: a systematic review.  
By Richards DM, Deeks JJ, Sheldon TA, Shaffer JL.

No. 2  
Diagnosis, management and screening of early localised prostate cancer.  
A review by Selley S, Donovan J, Faulkner A, Coast J, Gillatt D.

No. 3  
The diagnosis, management, treatment and costs of prostate cancer in England and Wales.  
A review by Chamberlain J, Melia J, Moss S, Brown J.

No. 4  
Screening for fragile X syndrome.  
A review by Murray J, Cuckle H, Taylor G, Hewison J.

No. 5  
Routine preoperative testing: a systematic review of the evidence.  
By Munro J, Booth A, Nicholl J.

No. 6  
Systematic review of the role of neonatal hearing screening in the detection of congenital hearing impairment.  
By Davis A, Bamford J, Wilson I, Ramkalawan T, Forsshaw M, Wright S.

**Volume 2, 1998**

No. 1  
Antenatal screening for Down's syndrome.  
By Wald NJ, Kennard A, Hackshaw A, McGuire A.

No. 2  
Screening for ovarian cancer: a systematic review.  
By Bell R, Petticrew M, Luengo S, Sheldon TA.

No. 3  
Consensus development methods, and their use in clinical guideline development.  
By Bell R, Petticrew M, Luengo S, Sheldon TA.

No. 4  

No. 5  
Effectiveness and efficiency of methods of dialysis therapy for end-stage renal disease: systematic reviews.  
By MacLeod A, Grant A, Donaldson C, Khan I, Campbell M, Daly C, et al.

No. 6  
Effectiveness of hip prostheses in primary total hip replacement: a critical review of evidence and an economic model.  

No. 7  
Antimicrobial prophylaxis in colorectal surgery: a systematic review of randomised controlled trials.  
By Song F, Glenny AM.

No. 8  
Bone marrow and peripheral blood stem cell transplantation for malignancy.  
A review by Johnson PWM, Simnett SJ, Sweetenham JW, Morgan GJ, Stewart LA.

No. 9  
Screening for speech and language delay: a systematic review of the literature.  
By Law J, Boyle J, Harris F, Harkness A, Nye C.

No. 10  
By Sculpher MJ, Petticrew M, Kelland JL, Elliott RA, Holdright DR, Buxton MJ.

No. 11  
Detection, adherence and control of hypertension for the prevention of stroke: a systematic review.  
By Ebrahim S.

No. 12  
Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review.  
By Britton A, Moore RA.

No. 13  
Choosing between randomised and nonrandomised studies: a systematic review.  
By Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C.

No. 14  
Evaluating patient-based outcome measures for use in clinical trials.  
A review by Fitzpatrick R, Davey C, Buxton MJ, Jones DR.
No. 15  
Ethical issues in the design and conduct of randomised controlled trials.  
A review by Edwards SJL, Lilford RJ, Braunholtz DA, Jackson JC, Hewison J, Thornton J.

No. 16  
Qualitative research methods in health technology assessment: a review of the literature.  
By Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P.

No. 17  
The costs and benefits of paramedic skills in pre-hospital trauma care.  
By Nicholl J, Hughes S, Dixon S, Turner J, Yates D.

No. 18  
Systematic review of endoscopic ultrasound in gastro-oesophageal cancer.  

No. 19  
Systematic reviews of trials and other studies.  
By Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F.

No. 20  
Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses.  

Volume 3, 1999

No. 1  
Informed decision making: an annotated bibliography and systematic review.  

No. 2  
Handling uncertainty when performing economic evaluation of healthcare interventions.  
A review by Briggs AH, Gray AM.

No. 3  
The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review.  

No. 4  

No. 5  
Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review.  
By Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ.

No. 6  
Assessing the costs of healthcare technologies in clinical trials.  
A review by Johnston K, Buxton MJ, Jones DR, Fitzpatrick R.

No. 7  
Cooperatives and their primary care emergency centres: organisation and impact.  
By Hallam L, Henthorne K.

No. 8  
Screening for cystic fibrosis.  
A review by Murray J, Cuckle H, Taylor G, Littlewood J, Hewison J.

No. 9  
A review of the use of health status measures in economic evaluation.  
By Brazier J, Deverill M, Green C, Harper R, Booth A.

No. 10  
A review by Billingham LJ, Abrams KR, Jones DR.

No. 11  
Antenatal and neonatal haemoglobinopathy screening in the UK: review and economic analysis.  
By Zeuner D, Ades AE, Karmen J, Brown J, Dezateux C, Aulionsen EN.

No. 12  
Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses.  

No. 13  
‘Early warning systems’ for identifying new healthcare technologies.  
By Robert G, Stevens A, Gabbay J.

No. 14  
A systematic review of the role of human papillomavirus testing within a cervical screening programme.  

No. 15  
Near patient testing in diabetes clinics: appraising the costs and outcomes.  
By Grieve R, Beech R, Vincent J, Mazurkiewicz J.

No. 16  
Positron emission tomography: establishing priorities for health technology assessment.  
A review by Robert G, Milne R.

No. 17 (Pt 1)  
The debridement of chronic wounds: a systematic review.  
By Bradley M, Cullum N, Sheldon T.

No. 17 (Pt 2)  
Systematic reviews of wound care management: (2) Dressings and topical agents used in the healing of chronic wounds.  
By Bradley M, Cullum N, Nelson EA, Petticrew M, Sheldon T, Torgerson D.

No. 18  
A systematic literature review of spiral and electron beam computed tomography: with particular reference to clinical applications in hepatic lesions, pulmonary embolus and coronary artery disease.  

No. 19  
What role for statins? A review and economic model.  

No. 20  
Factors that limit the quality, number and progress of randomised controlled trials.  
A review by Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kautka S, et al.

No. 21  
Antimicrobial prophylaxis in total hip replacement: a systematic review.  
By Glenny AM, Song F.

No. 22  
Health promoting schools and health promotion in schools: two systematic reviews.  
By Lister-Sharp D, Chapman S, Stewart-Brown S, Sowden A.

No. 23  
Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee.  
Volume 4, 2000

No. 1
The estimation of marginal time preference in a UK-wide sample (TEMPUS) project.
A review by Cairns JA, van der Pol MM.

No. 2
Geriatric rehabilitation following fractures in older people: a systematic review.

No. 3
Screening for sickle cell disease and thalassaemia: a systematic review with supplementary research.
By Davies SC, Cronin E, Gill M, Greenberg P, Hickman M, Normand C.

No. 4
Community provision of hearing aids and related audiology services.
A review by Reeves DJ, Alborz A, Hickson FS, Bamford JM.

No. 5
False-negative results in screening programmes: systematic review of impact and implications.
By Petticrew MP, Sowden AJ, Lister-Sharp D, Wright K.

No. 6
Costs and benefits of community postnatal support workers: a randomised controlled trial.
By Morrell CJ, Spilby H, Stewart P, Walters S, Morgan A.

No. 7
Implantable contraceptives (subdermal implants and hormonally impregnated intrauterine systems) versus other forms of reversible contraceptives: two systematic reviews to assess relative effectiveness, acceptability, tolerability and cost-effectiveness.

No. 8
An introduction to statistical methods for health technology assessment.
A review by White SJ, Ashby D, Brown PJ.

No. 9
Disease-modifying drugs for multiple sclerosis: a rapid and systematic review.
By Clegg A, Bryant J, Milne R.

No. 10
Publication and related biases.
A review by Song F, Eastwood AJ, Gilbody S, Dudley L, Sutton AJ.

No. 11
Cost and outcome implications of the organisation of vascular services.
By Michaels J, Brazier J, Pallfreyman S, Shackley P, Slack R.

No. 12
Monitoring blood glucose control in diabetes mellitus: a systematic review.
By Coster S, Gullford MC, Seed PT, Powrie JK, Swammanathan R.

No. 13
The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature.

No. 14
The determinants of screening uptake and interventions for increasing uptake: a systematic review.

No. 15
The effectiveness and cost-effectiveness of prophylactic removal of wisdom teeth.
A rapid review by Song F, O'Meara S, Wilson P, Goldier S, Kleijnen J.

No. 16

No. 17
A rapid and systematic review of the effectiveness and cost-effectiveness of the taxanes used in the treatment of advanced breast and ovarian cancer.
By Lister-Sharp D, McDonagh MS, Khan KS, Kleijnen J.

No. 18
Liquid-based cytology in cervical screening: a rapid and systematic review.
By Payne N, Chilcott J, McGoogan E.

No. 19
Randomised controlled trial of non-directive counselling, cognitive–behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care.

No. 20
Routine referral for radiography of patients presenting with low back pain: is patients' outcome influenced by GPs' referral for plain radiography?
By Kerry S, Hilton S, Patel S, Dunstal D, Rink E, Lord J.

No. 21
Systematic reviews of wound care management: (3) antimicrobial agents for chronic wounds; (4) diabetic foot ulceration.
By O'Meara S, Callum N, Majid M, Sheldon T.

No. 22
Using routine data to complement and enhance the results of randomised controlled trials.
By Lewsey JD, Leyland AH, Murray GD, Boddy FA.

No. 23
Coronary artery stents in the treatment of ischaemic heart disease: a rapid and systematic review.
By Meads C, Cummins C, Jolly K, Stevens A, Burls A, Hyde C.

No. 24
Outcome measures for adult critical care: a systematic review.
By Hayes JA, Black NA, Jenkinson C, Young JD, Rowan KM, Daly K, et al.

No. 25
A systematic review to evaluate the effectiveness of interventions to promote the initiation of breastfeeding.
By Fairbanks L, O'Meara S, Renfrew MJ, Woolridge M, Sowden AJ, Lister-Sharp D.

No. 26
Implantable cardioverter defibrillators: arrhythmias. A rapid and systematic review.
By Parkes J, Bryant J, Milne R.

No. 27
Treatments for fatigue in multiple sclerosis: a rapid and systematic review.
By Brañas P, Jordan R, Fry-Smith A, Burls A, Hyde C.

No. 28
Early asthma prophylaxis, natural history, skeletal development and economy (EASE): a pilot randomised controlled trial.

No. 29
Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis.
By Marks D, Wonderling D, Thorogood M, Lambeth H, Humphries SE, Neil HAW.

No. 30
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of glycoprotein IIb/IIIa antagonists in the medical management of unstable angina.
By McDonagh MS, Bachmann LM, Goldier S, Kleijnen J, ter Riet G.
No. 31
A randomised controlled trial of prehospital intravenous fluid replacement therapy in serious trauma.
By Turner J, Nicholl J, Webber L, Cox H, Dixon S, Yates D.

No. 32
Intrathecal pumps for giving opioids in chronic pain: a systematic review.
By Williams JE, Loug G, Towler G.

No. 33
Combination therapy (interferon alfa and ribavirin) in the treatment of chronic hepatitis C: a rapid and systematic review.
By Shepherd J, Waugh N, Hewitson P.

No. 34
A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies.
By MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AMS.

No. 35
Intravascular ultrasound-guided interventions in coronary artery disease: a systematic literature review, with decision-analytic modelling, of outcomes and cost-effectiveness.
By Berry E, Kelly S, Hutton J, Lindsay H, Blaxill JM, Evans JA, et al.

No. 36
A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression.
By Simpson S, Corney R, Fitzgerald P, Beecham J.

No. 37
Systematic review of treatments for atopic eczema.
By Hoare C, Li Wan Po A, Williams H.

No. 38
Bayesian methods in health technology assessment: a review.
By Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR.

No. 39
The management of dyspepsia: a systematic review.

No. 40
A systematic review of treatments for severe psoriasis.
By Griffiths CEM, Clark CM, Chalmers R, Li Wan Po A, Williams HC.

Volume 5, 2001
No. 1
Clinical and cost-effectiveness of donepezil, rivastigmine and galantamine for Alzheimer’s disease: a rapid and systematic review.

No. 2
The clinical effectiveness and cost-effectiveness of rihuzole for motor neurone disease: a rapid and systematic review.

No. 3
Equity and the economic evaluation of healthcare.
By Sassi F, Archard L, Le Grand J.

No. 4
Quality-of-life measures in chronic diseases of childhood.
By Eiser C, Morse R.

No. 5
Eliciting public preferences for healthcare: a systematic review of techniques.

No. 6
General health status measures for people with cognitive impairment: learning disability and acquired brain injury.
By Riemsma RP, Forbes CA, Glanville JM, Eastwood AJ, Kleijnen J.

No. 7
An assessment of screening strategies for fragile X syndrome in the UK.
By Pembray ME, Barnicoat AJ, Carmichael B, Bobrow M, Turner G.

No. 8
Issues in methodological research: perspectives from researchers and commissioners.

No. 9
Systematic reviews of wound care management: (5) beds; (6) compression; (7) laser therapy, therapeutic ultrasound, electrotherapy and electromagnetic therapy.
By Cullum N, Nelson EA, Fleming K, Sheldon T.

No. 10
Effects of educational and psychosocial interventions for adolescents with diabetes mellitus: a systematic review.
By Hampson SE, Skinner TC, Hart J, Storey L, Gage H, Foxcroft D, et al.

No. 11
Effectiveness of autologous chondrocyte transplantation for hyaline cartilage defects in knees: a rapid and systematic review.
By Johanputra P, Parry D, Fry-Smith A, Burd A.

No. 12
Statistical assessment of the learning curves of health technologies.
By Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT.

No. 13
The effectiveness and cost-effectiveness of temozolomide for the treatment of recurrent malignant glioma: a rapid and systematic review.
By Dines J, Cave C, Huang S, Major K, Milne R.

No. 14
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of debriding agents in treating surgical wounds healing by secondary intention.
By Lewis R, Whiting P, ter Riet G, O’Meara S, Glanzvile J.

No. 15
Home treatment for mental health problems: a systematic review.

No. 16
How to develop cost-conscious guidelines.
By Eccles M, Mason J.

No. 17
The role of specialist nurses in multiple sclerosis: a rapid and systematic review.
By De Broe S, Christopher F, Waugh N.

No. 18
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of orlistat in the management of obesity.
By O’Meara S, Riemsma R, Shirran L, Mather L, ter Riet G.

No. 19
The clinical effectiveness and cost-effectiveness of pioglitazone for type 2 diabetes mellitus: a rapid and systematic review.
By Chilcott J, Wight J, Lloyd Jones M, Tappenden P.

No. 20
Extended scope of nursing practice: a multicentre randomised controlled trial of appropriately trained nurses and preregistration house officers in preoperative assessment in elective general surgery.
No. 21 Systematic reviews of the effectiveness of day care for people with severe mental disorders: (1) Acute day hospital versus admission; (2) Vocational rehabilitation; (3) Day hospital versus outpatient care.

No. 22 The measurement and monitoring of surgical adverse events.
By Bruce J, Russell EM, Mollison J, Krukowski ZH.

No. 23 Action research: a systematic review and guidance for assessment.
By Waterman H, Tillen D, Dickson R, de Koning K.

No. 24 A rapid and systematic review of the clinical effectiveness and cost-effectiveness of gemcitabine for the treatment of pancreatic cancer.

No. 25 A rapid and systematic review of the evidence for the clinical effectiveness and cost-effectiveness of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer.
By Lloyd Jones M, Hummell S, Bansback N, Orr B, Seymour M.

No. 26 Comparison of the effectiveness of inhaler devices in asthma and chronic obstructive airways disease: a systematic review of the literature.

No. 27 The cost-effectiveness of magnetic resonance imaging for investigation of the knee joint.

No. 28 A rapid and systematic review of the clinical effectiveness and cost-effectiveness of topotecan for ovarian cancer.
By Forbes C, Shirran L, Bagnall A-M, Duffy S, ter Riet G.

No. 29 Superseded by a report published in a later volume.

No. 30 The role of radiography in primary care patients with low back pain of at least 6 weeks duration: a randomised (unblinded) controlled trial.
By Kendrick D, Fielding K, Bentley E, Miller P, Kerslake R, Pringle M.

No. 31 Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients.

No. 32 A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer.
By Clegg A, Scott DA, Sidhu M, Hewitson P, Waugh N.

No. 33 Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives.
By Brooks ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G.

No. 34 Depot antipsychotic medication in the treatment of patients with schizophrenia: (1) Meta-review; (2) Patient and nurse attitudes.
By David AS, Adams C.

No. 35 A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression.

No. 36 Cost analysis of child health surveillance.
By Sanderson D, Wright D, Acton C, Duree D.

Volume 6, 2002

No. 1 A study of the methods used to select review criteria for clinical audit.
By Hearshaw H, Harker R, Cheater F, Baker R, Grimshaw G.

No. 2 Fludarabine as second-line therapy for B cell chronic lymphocytic leukaemia: a technology assessment.

No. 3 Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin’s lymphoma: a systematic review and economic evaluation.

No. 4 A systematic review of discharge arrangements for older people.

No. 5 The clinical effectiveness and cost-effectiveness of inhaler devices used in the routine management of chronic asthma in older children: a systematic review and economic evaluation.
By Peters J, Stevenson M, Beverley C, Lim J, Smith S.

No. 6 The clinical effectiveness and cost-effectiveness of sibutramine in the management of obesity: a technology assessment.
By O’Meara S, Riemsma R, Shirran L, Mather L, ter Riet G.

No. 7 The cost-effectiveness of magnetic resonance angiography for carotid artery stenosis and peripheral vascular disease: a systematic review.

No. 8 Promoting physical activity in South Asian Muslim women through ‘exercise on prescription’.
By Carroll B, Ali N, Azam N.

No. 9 Zanamivir for the treatment of influenza in adults: a systematic review and economic evaluation.

No. 10 A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models.
By Richards RG, Sampson FC, Beard SM, Tappenden P.

No. 11 Screening for gestational diabetes: a systematic review and economic evaluation.
By Scott DA, Loveman E, McIntyre L, Waugh N.

No. 12 The clinical effectiveness and cost-effectiveness of surgery for people with morbid obesity: a systematic review and economic evaluation.

No. 13 The clinical effectiveness of trastuzumab for breast cancer: a systematic review.

No. 14 The clinical effectiveness and cost-effectiveness of vinorelbine for breast cancer: a systematic review and economic evaluation.
No. 15 A systematic review of the effectiveness and cost-effectiveness of metal-on-metal hip resurfacing arthroplasty for treatment of hip disease.
By Vale L, Wyness L, McCormack K, McKenzie I, Brazzelli M, Stearns SC.

No. 16 The clinical effectiveness and cost-effectiveness of bupropion and nicotine replacement therapy for smoking cessation: a systematic review and economic evaluation.
By Woolacott NF, Jones L, Forbes CA, Mather LC, Sowden AJ, Song FJ, et al.

No. 17 A systematic review of effectiveness and economic evaluation of new drug treatments for juvenile idiopathic arthritis: etanercept.
By Cummins C, Connock M, Fry-Smith A, Burls A.

No. 18 Clinical effectiveness and cost-effectiveness of growth hormone in children: a systematic review and economic evaluation.

By Bryant J, Loveman E, Chase D, Mihaylova B, Cave C, Gerard K, et al.

No. 20 Clinical medication review by a pharmacist of patients on repeat prescriptions in general practice: a randomised controlled trial.
By Zermansky AG, Petty DR, Raynor DK, Lowe CJ, Freementle N, Vail A.

No. 21 The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation.
By Jobanputra P, Barton P, Bryan S, Burls A.

No. 22 A systematic review and economic evaluation of computerised cognitive behaviour therapy for depression and anxiety.
By Kaltenhauser E, Shackley P, Stevens K, Beverley C, Parry G, Chilcott J.

No. 23 A systematic review and economic evaluation of pegylated liposomal doxorubicin hydrochloride for ovarian cancer.
By Forbes C, Wilby J, Richardson G, Sculptor M, Mather L, Riemsma R.

No. 24 A systematic review of the effectiveness of interventions based on a stages-of-change approach to promote individual behaviour change.

No. 25 A systematic review update of the clinical effectiveness and cost-effectiveness of glycoprotein IIb/IIIa antagonists.

No. 26 A systematic review of the effectiveness, cost-effectiveness and barriers to implementation of thrombolytic and neuroprotective therapy for acute ischaemic stroke in the NHS.

No. 27 A randomised controlled crossover trial of nurse practitioner versus doctor-led outpatient care in a bronchiectasis clinic.

No. 28 Clinical effectiveness and cost – consequences of selective serotonin reuptake inhibitors in the treatment of sex offenders.
By Adi Y, Ashcroft D, Browne K, Beech A, Fry-Smith A, Hyde C.

No. 29 Treatment of established osteoporosis: a systematic review and cost–utility analysis.
By Kanis JA, Brazier JE, Stevenson M, Calvert NW, Lloyd Jones M.

No. 30 Which anaesthetic agents are cost-effective in day surgery? Literature review, national survey of practice and randomised controlled trial.

No. 31 Screening for hepatitis C among injecting drug users and in genitourinary medicine clinics: systematic reviews of effectiveness, modelling study and national survey of current practice.

No. 32 The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature.

No. 33 The effectiveness and cost-effectiveness of imatinib in chronic myeloid leukaemia: a systematic review.
By Garside R, Round A, Dalziel K, Stein K, Royle R.

No. 34 A comparative study of hypertonic saline, daily and alternate-day rhDNase in children with cystic fibrosis.

No. 35 A systematic review of the costs and effectiveness of different models of paediatric home care.

Volume 7, 2003

No. 1 How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study.
By Egger M, Juni P, Bartlett C, Holenstein F, Sterne J.

No. 2 Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of home versus hospital or satellite unit haemodialysis for people with end-stage renal failure.

No. 3 Systematic review and economic evaluation of the effectiveness of infliximab for the treatment of Crohn’s disease.
By Clark W, Raftery J, Barton P, Song F, Fry-Smith A, Burls A.

No. 4 A review of the clinical effectiveness and cost-effectiveness of routine anti-D prophylaxis for pregnant women who are rhesus negative.

No. 5 Systematic review and evaluation of the use of tumour markers in paediatric oncology: Ewing’s sarcoma and neuroblastoma.

No. 6 The cost-effectiveness of screening for Helicobacter pylori to reduce mortality and morbidity from gastric cancer and peptic ulcer disease: a discrete-event simulation model.
No. 7
The clinical effectiveness and cost-effectiveness of routine dental checks: a systematic review and economic evaluation.

No. 8
A multicentre randomised controlled trial assessing the costs and benefits of using structured information and analysis of women’s preferences in the management of menorrhagia.

No. 9
Clinical effectiveness and cost-utility of photodynamic therapy for wet age-related macular degeneration: a systematic review and economic evaluation.
By Meads C, Salas C, Roberts T, Moore D, Fry-Smith A, Hyde C.

No. 10
Evaluation of molecular tests for prenatal diagnosis of chromosome abnormalities.

No. 11
First and second trimester antenatal screening for Down’s syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURESS).
By Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM.

No. 12
The effectiveness and cost-effectiveness of ultrasound locating devices for central venous access: a systematic review and economic evaluation.
By Calvert N, Hind D, McWilliams RG, Thomas SM, Beverley C, Davidson A.

No. 13
A systematic review of atypical antipsychotics in schizophrenia.
By Bagnall A-M, Jones L, Robertson G, Dwyer N, Riemsma R.

No. 14
Prostate Testing for Cancer and Treatment (ProtecT) feasibility study.
By Donovan J, Hamdy F, Neal D, Peters T, Oliver S, Brindle L, et al.

No. 15
Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation.

No. 16
Screening for fragile X syndrome: a literature review and modelling.
By Song FJ, Barton P, Sleightholme V, Yao GL, Fry-Smith A.

No. 17
Systematic review of endoscopic sinus surgery for nasal polyps.
By Dalziel K, Stein K, Round A, Garside R, Royle P.

No. 18
Towards efficient guidelines: how to monitor guideline use in primary care.
By Hutchinson A, McIntosh A, Cox S, Gilbert C.

No. 19
Effectiveness and cost-effectiveness of acute hospital-based spinal cord injuries services: systematic review.
By Bagnall A-M, Jones L, Richardson G, Duffy S, Riemsma R.

No. 20
Prioritisation of health technology assessment. The PATHS model: methods and case studies.
By Townsend J, Buxton M, Harper G.

No. 21

No. 22
By Loveman E, Cave C, Green C, Royle P, Dunn N, Waugh N.

No. 23
The role of modelling in prioritising and planning clinical trials.
By Chikov J, Brennan A, Booth A, Karron J, Tappenden P.

No. 24
Cost–benefit evaluation of routine influenza immunisation in people 65–74 years of age.
By Allsop S, Gosney M, Haycox A, Regan M.

No. 25
The clinical and cost-effectiveness of pulsatile machine perfusion versus cold storage of kidneys for transplantation retrieved from heart-beating and non-heart-beating donors.
By Wight J, Chikov J, Holmes M, Brewer N.

No. 26
Can randomised trials rely on existing electronic data? A feasibility study to explore the value of routine data in health technology assessment.
By Williams JG, Cheung WY, Cohen DR, Hutchings HA, Longo MF, Russell IT.

No. 27
Evaluating non-randomised intervention studies.

No. 28
A randomised controlled trial to assess the impact of a package comprising a patient-orientated, evidence-based self-help guidebook and patient-centred consultations on disease management and satisfaction in inflammatory bowel disease.

No. 29
The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review.
By Dinnes J, Loveman E, McIntyre L, Waugh N.

No. 30
The value of digital imaging in diabetic retinopathy.

No. 31
Lowering blood pressure to prevent myocardial infarction and stroke: a new preventive strategy.
By Law M, Wald N, Morris J.

No. 32
Clinical and cost-effectiveness of capicitabine and tegafur with uracil for the treatment of metastatic colorectal disease.
By Ward S, Kaltenthaler E, Cowan J, Brewer N.

No. 33
Evaluating non-randomised intervention studies.

No. 34
By Hummel S, Paisley S, Morgan A, Currie E, Brewer N.
No. 35 Systematic review and economic decision modelling for the prevention and treatment of influenza A and B.

No. 36 A randomised controlled trial to evaluate the clinical and cost-effectiveness of Hickman line insertions in adult cancer patients by nurses.
  By Boland A, Haycox A, Bagust A, Fitzsimmons L.

No. 37 Redesigning postnatal care: a randomised controlled trial of protocol-based midwifery-led care focused on individual women's physical and psychological health needs.

No. 38 Estimating implied rates of discount in healthcare decision-making.
  By West RR, McNabb R, Thompson AGH, Sheldon TA, Grimley Evans J.

  By Cooper BS, Stone SP, Kibbler CC, Cookson BD, Roberts JA, Medley GF, et al.

No. 40 Treatments for spasticity and pain in multiple sclerosis: a systematic review.
  By Beard S, Humm A, Wight J.

No. 41 The inclusion of reports of randomised trials published in languages other than English in systematic reviews.
  By Moher D, Pham B, Lawson ML, Klassen TP.

No. 42 The impact of screening on future health-promoting behaviours and health beliefs: a systematic review.

Volume 8, 2004

No. 1 What is the best imaging strategy for acute stroke?
  By Wardlaw JM, Keir SL, Seymour J, Lewis S, Sandercoc PAG, Dennis MS, et al.

No. 2 Systematic review and modelling of the investigation of acute and chronic chest pain presenting in primary care.
  By Mant J, McManus RJ, Oakes RAL, Delaney BC, Barton PM, Deeks JJ, et al.

No. 3 The effectiveness and cost-effectiveness of microwave and thermal balloon endometrial ablation for heavy menstrual bleeding: a systematic review and economic modelling.

No. 4 A systematic review of the role of bisphosphonates in metastatic disease.

No. 5 Systematic review of the clinical effectiveness and cost-effectiveness of capetabibine (Xeloda®) for locally advanced and/or metastatic breast cancer.
  By Jones L, Hawkins N, Westwood M, Wright K, Richardson G, Riemsma R.

No. 6 Effectiveness and efficiency of guideline dissemination and implementation strategies.

No. 7 Clinical effectiveness and costs of the Sugarbaker procedure for the treatment of pseudomyxoma peritonei.
  By Bryant J, Clegg AJ, Siddhu MK, Brodin H, Royle P, Davidson P.

No. 8 Psychological treatment for insomnia in the regulation of long-term hypnotic drug use.
  By Morgan K, Dixon S, Mathers N, Thompson J, Tonymen M.

No. 9 Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome.
  By Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ.

No. 10 A systematic review and economic evaluation of magnetic resonance cholangiopancreatography compared with diagnostic endoscopic retrograde cholangiopancreatography.

No. 11 The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis.

  By Pandor A, Eastham J, Beverley C, Chilcott J, Paisley S.

  By Czoski-Murray C, Warren E, Chilcott J, Beverley C, Pyllaki MA, Cowan J.

No. 14 Routine examination of the newborn: the EMREN study. Evaluation of an extension of the midwife role including a randomised controlled trial of appropriately trained midwives and paediatric senior house officers.

No. 15 Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach.

No. 16 A multi-centre randomised controlled trial of minimally invasive direct coronary bypass grafting versus percutaneous transluminal coronary angioplasty with stenting for proximal stenosis of the left anterior descending coronary artery.

No. 17 Does early magnetic resonance imaging influence management or improve outcome in patients referred to secondary care with low back pain? A pragmatic randomised controlled trial.
  By Gilbert FJ, Grant AM, Gillan MGC, Vale L, Scott NW, Campbell MK, et al.

No. 18 The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis.
  By Clark W, Jobanputra P, Barton P, Burls A.
A rapid and systematic review and economic evaluation of the clinical and cost-effectiveness of newer drugs for treatment of mania associated with bipolar affective disorder.


No. 19

Liquor-based cytology in cervical screening: an updated rapid and systematic review and economic analysis.


No. 20

Systematic review of the long-term effects and economic consequences of treatments for obesity and implications for health promotion.


No. 21

Autoantibody testing in children with newly diagnosed type 1 diabetes mellitus.

By Dretzke J, Cummins C, Sandercock J, Fry-Smith A, Barrett T, Burs A.

No. 22

Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients.

By Dretzke J, Sandercock J, Bayliss S, Burs A.

No. 23

Newer hypnotic drugs for the short-term management of insomnia: a systematic review and economic evaluation.


No. 24

Development and validation of methods for assessing the quality of diagnostic accuracy studies.

By Whiting P, Rutjes AWS, Dinnes J, Reitsma JB, Bossuyt PMM, Kleijnen J.

No. 25

EVALUATE hysterectomy trial: a multicentre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy.


No. 26


By Tappenden P, Chilcott JB, Eggington S, Oakley J, McCabe C.

No. 27


By Dalsziel R, Round A, Stein K, Garside R, Price A.

No. 28

VenUS E: a randomised controlled trial of two types of bandage for treating venous leg ulcers.

By Iglesias C, Nelson EA, Cullum NA, Torgerson DJ, on behalf of the VenUS Team.

No. 29

Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of myocardial perfusion scintigraphy for the diagnosis and management of angina and myocardial infarction.


No. 30

A pilot study on the use of decision theory and value of information analysis as part of the NHS Technology Assessment programme.

By Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S.

No. 31

The Social Support and Family Health Study: a randomised controlled trial and economic evaluation of two alternative forms of postnatal support for mothers living in disadvantaged inner-city areas.


No. 32

Psychosocial aspects of genetic screening of pregnant women and newborns: a systematic review.

By Green JM, Hewson J, Bekker HL, Bryant, Cuckle HS.

No. 33

Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status.


No. 34

Coronary artery stents: a rapid systematic review and economic evaluation.


No. 35

Review of guidelines for good practice in decision-analytic modelling in health technology assessment.


No. 36

Rituximab (MabThera) for aggressive non-Hodgkin’s lymphoma: systematic review and economic evaluation.

By Knight C, Hind D, Brewer N, Abbott V.

No. 37

Clinical effectiveness and cost-effectiveness of clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events: a systematic review and economic evaluation.

By Jones L, Griffin S, Palmer S, Main C, Orton V, Sculpher M, et al.

No. 38

Pegylated interferon α-2a and -2b in combination with ribavirin in the treatment of chronic hepatitis C: a systematic review and economic evaluation.

By Shepher J, Brodin H, Cave C, Waugh N, Price A, Gabbay J.

No. 39

Clopidogrel used in combination with aspirin compared with aspirin alone in the treatment of non-ST-segment-elevation acute coronary syndromes: a systematic review and economic evaluation.

By Main C, Palmer S, Griffin S, Jones L, Orton V, Sculpher M, et al.

No. 40

Provision, uptake and cost of cardiac rehabilitation programmes: improving services to under-represented groups.

By Beswick AD, Rees K, Griebsch I, Taylor FC, Burke M, West RR, et al.

No. 41

Involving South Asian patients in clinical trials.

By Hussain-Gambles M, Leese B, Akin K, Brown J, Mason S, Tovey P.

No. 42

Clinical and cost-effectiveness of continuous subcutaneous insulin infusion for diabetes.

By Colquitt JL, Green C, Sidhu MK, Hartwell D, Waugh N.

No. 43

Identification and assessment of ongoing trials in health technology assessment reviews.


No. 44

Systematic review and economic evaluation of a long-acting insulin analogue, insulin glargine.

By Warren E, Weatherley-Jones E, Chilcott J, Beverley C.

No. 45
No. 46
Supplementation of a home-based exercise programme with a class-based programme for people with osteoarthritis of the knees: a randomised controlled trial and health economic analysis.

No. 47
Clinical and cost-effectiveness of once-daily versus more frequent use of same potency topical corticosteroids for atopic eczema: a systematic review and economic evaluation.
By Green C, Colquitt JL, Kirby J, Davidson P, Payne E.

No. 48
Acupuncture of chronic headache disorders in primary care: randomised control trial and economic analysis.

No. 49
Generalisability in economic evaluation studies in healthcare: a review and case studies.

No. 50
Virtual outreach: a randomised control trial and economic evaluation of joint teleconferenced medical consultations.

Volume 9, 2005

No. 1
Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne.

No. 2
Do the findings of case series studies vary significantly according to methodological characteristics?
By Dalziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L.

No. 3
Improving the referral process for familial breast cancer genetic counselling: findings of three randomised controlled trials of two interventions.

No. 4
Randomised evaluation of alternative electroceutical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia.
By Fowler C, McAllister W, Plail R, Karim O, Yang Q.

No. 5
A pragmatic randomised controlled trial of the cost-effectiveness of palliative therapies for patients with inoperable oesophageal cancer.
By Shenefine J, McNamee P, Steen N, Bond J, Griffin SM.

No. 6
Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography.

No. 7
Issues in data monitoring and interim analysis of trials.
By Grant AM, Altman DG, Babiker AB, Campbell MK, Clemens FJ, Darbyshire JH, et al.

No. 8
Lay public’s understanding of equipoise and randomisation in randomised controlled trials.

No. 9
Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies.
By Greenhalgh J, Knight C, Hind D, Beverley C, Walters S.

No. 10
Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology.

No. 11
Clinical effectiveness and cost-effectiveness of drotrecogin alfa (activated) (Xigris) for the treatment of severe sepsis in adults: a systematic review and economic evaluation.

No. 12
A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy.
By Dinnes J, Deeks J, Kirby J, Roderick P.

No. 13
Cervical screening programmes: can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK.
By Willis BH, Barton P, Pearmain P, Bryan S, Hyde C.

No. 14
Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation.

No. 15
Clinical effectiveness, tolerability and cost-effectiveness of newer drugs for epilepsy in adults: a systematic review and economic evaluation.

No. 16
A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine.

No. 17
Clinical effectiveness and cost-effectiveness of immediate angioplasty for acute myocardial infarction: systematic review and economic evaluation.

No. 18
A randomised controlled comparison of alternative strategies in stroke care.
By Kafra L, Evans A, Perez I, Knupp M, Swift C, Donaldson N.

No. 19
The investigation and analysis of critical incidents and adverse events in healthcare.
By Wodoszynowycz M, Rogers S, Taylor-Adams S, Vincent C.

No. 20
Potential use of routine databases in health technology assessment.
By Rafferty J, Roderick P, Stevens A.

No. 21

No. 22
A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis.
By Stevenson M, Lloyd Jones M, De Nigris E, Brewer N, Davis S, Oakley J.
No. 23
A systematic review to examine the impact of psycho-educational interventions on health outcomes and costs in adults and children with difficult asthma.

No. 24
An evaluation of the costs, effectiveness and quality of renal replacement therapy provision in renal satellite units in England and Wales.

No. 25
Imatinib for the treatment of patients with unresectable and/or metastatic gastrointestinal stromal tumours: systematic review and economic evaluation.

No. 26
Indirect comparisons of competing interventions.

No. 27
Cost-effectiveness of alternative strategies for the initial medical management of non-ST elevation acute coronary syndrome: systematic review and decision-analytical modelling.

No. 28
Outcomes of electrically stimulated gracilis neosphincter surgery.
By Tillin T, Chambers M, Feldman R.

No. 29
The effectiveness and cost-effectiveness of pimecrolimus and tacrolimus for atopic eczema: a systematic review and economic evaluation.
By Bartlett C, Doyal L, Ebrahim S, Davey P, Arden N, Coglan L, Rogers P.

No. 30
The British Rheumatoid Outcome Study Group (BROSG) randomised controlled trial to compare the effectiveness and cost-effectiveness of aggressive versus symptomatic therapy in established rheumatoid arthritis.
By Symmons D, Tricker K, Roberts C, Davies L, Dawes P, Scott DL.

No. 31
The clinical and cost-effectiveness of implantable cardioverter defibrillators: a systematic review.
By Bryant J, Brodin H, Loveman E, Payne E, Clegg A.

No. 32
A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study.

No. 33
Displaced intracapsular hip fractures in fit, older people: a randomised comparison of reduction and fixation, bipolar hemiarthroplasty and total hip arthroplasty.
By Grant JF, Grant A, Masson M, Scott NW, Forbes JE.

No. 34
Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland.

No. 35
Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis.

No. 36
The effectiveness and cost-effectiveness of dual-chamber pacemakers compared with single-chamber pacemakers for bradyarrhythmia due to atrioventricular block or sick sinus syndrome: systematic review and economic evaluation.
By Castelnuovo E, Stein K, Pitt M, Garside R, Payne E.

No. 37
Systematic review on urine albumin testing for early detection of diabetic complications.

No. 38
Systematic review of the cost-effectiveness of water-based therapy for lower limb oedema.
By Cochrane T, Davey RC, Matthews Edwards SM.

No. 39
Clinical and cost-effectiveness of autologous chondrocyte implantation for cartilage defects in knee joints: systematic review and economic evaluation.
Health Technology Assessment reports published to date

No. 48  Systematic review of effectiveness of different treatments for childhood retinoblastoma.

No. 49  Towards evidence-based guidelines for the prevention of venous thromboembolism: systematic reviews of mechanical methods, oral anticoagulation, dextran and regional anaesthesia as thromboprophylaxis.

No. 50  The effectiveness and cost-effectiveness of parent training/education programmes for the treatment of conduct disorder, including oppositional defiant disorder, in children.

Volume 10, 2006

No. 1  The clinical and cost-effectiveness of-donepezil, rivastigmine, galantamine and memantine for Alzheimer’s disease.

No. 2  FOOD: a multicentre randomised trial evaluating feeding policies in patients admitted to hospital with a recent stroke.
   By Dennis M, Lewis S, Cranwisk G, Forbes J.

No. 3  The clinical effectiveness and cost-effectiveness of computed tomography screening for lung cancer: systemic reviews.

No. 4  A systematic review of the effectiveness and cost-effectiveness of neuroimaging assessments used to visualise the seizure focus in people with refractory epilepsy being considered for surgery.

No. 5  Comparison of conference abstracts and presentations with full-text articles in the health technology assessments of rapidly evolving technologies.
   By Dundar Y, Dodd S, Dickson R, Walley T, Haycox A, Williamson PR.

No. 6  Systematic review and evaluation of methods of assessing urinary incontinence.

No. 7  The clinical effectiveness and cost-effectiveness of newer drugs for children with epilepsy. A systematic review.

No. 8  Surveillance of Barrett’s oesophagus: exploring the uncertainty through systematic review, expert workshop and economic modelling.
   By Garside R, Pitt M, Somerville M, Stein K, Price A, Gilbert N.

No. 9  Topotecan, pegylated liposomal doxorubicin hydrochloride and paclitaxel for second-line or subsequent treatment of advanced ovarian cancer: a systematic review and economic evaluation.

No. 10  Evaluation of molecular techniques in prediction and diagnosis of cytomegalovirus disease in immunocompromised patients.
   By Szczepura A, Westmoreland D, Vinogradova Y, Fox J, Clark M.

No. 11  Screening for thrombophilia in high-risk situations: systematic review and cost-effectiveness analysis. The Thrombosis: Risk and Economic Assessment of Thrombophilia Screening (TREATS) study.

No. 12  A series of systematic reviews to inform a decision analysis for sampling and treating infected diabetic foot ulcers.

No. 13  Randomised clinical trial, observational study and assessment of cost-effectiveness of the treatment of varicose veins (REACTIV trial).

No. 14  The cost-effectiveness of screening for oral cancer in primary care.
   By Speight PM, Palmer S, Moles DR, Downer MC, Smith DH, Henriksson M, et al.


No. 17  Randomised controlled trials of conventional antipsychotic versus new atypical drugs, and new atypical drugs versus clozapine, in people with schizophrenia responding poorly to, or intolerant of, current drug treatment.
   By Lewis SW, Davies L, Jones PB, Barnes TRE, Murray RM, Kerwin R, et al.

No. 18  Diagnostic tests and algorithms used in the investigation of haematuria: systematic reviews and economic evaluation.

No. 19  Cognitive behavioural therapy in addition to antispasmodic therapy for irritable bowel syndrome in primary care: randomised controlled trial.

No. 20  A systematic review of the clinical effectiveness and cost-effectiveness of enzyme replacement therapies for Fabry’s disease and mucopolysaccharidosis type 1.

No. 21  Health benefits of antiviral therapy for mild chronic hepatitis C: randomised controlled trial and economic evaluation.
   By Wright M, Grieve R, Roberts J, Main J, Thomas HC, on behalf of the UK Mild Hepatitis C Trial Investigators.

No. 22  Pressure relieving support surfaces: a randomised evaluation.
No. 23
A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents.

No. 24
The clinical effectiveness and cost-effectiveness of enzyme replacement therapy for Gaucher’s disease: a systematic review.

No. 25
Effectiveness and cost-effectiveness of salicylic acid and cryotherapy for cutaneous warts. An economic decision model.

No. 26
A systematic literature review of the effectiveness of non-pharmacological interventions to prevent wandering in dementia and evaluation of the ethical implications and acceptability of their use.

No. 27
A review of the evidence on the effects and costs of implantable cardioverter defibrillator therapy in different patient groups, and modelling of cost-effectiveness and cost-utility for these groups in a UK context.

No. 28
Adefovir dipivoxil and pegylated interferon alfa-2a for the treatment of chronic hepatitis B: a systematic review and economic evaluation.
By Shepherd J, Jones J, Takeda A, Davidson P, Price A.

No. 29
By Harvey S, Stevens K, Harrison D, Young D, Brampton W, McCabe C, et al.

No. 30
Accurate, practical and cost-effective assessment of carotid stenosis in the UK.
By Wardlaw JM, Chappell FM, Stevenson M, De Nigris E, Thomas S, Gillard J, et al.

No. 31
Etanercept and infliximab for the treatment of psoriatic arthritis: a systematic review and economic evaluation.

No. 32
The cost-effectiveness of testing for hepatitis C in former injecting drug users.

No. 33
Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation.

No. 34
Cost-effectiveness of using prognostic information to select women with breast cancer for adjuvant systemic therapy.

No. 35
Psychological therapies including dialectical behaviour therapy for borderline personality disorder: a systematic review and preliminary economic evaluation.

No. 36
Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model.

No. 37
Cognitive behavioural therapy in chronic fatigue syndrome: a randomised controlled trial of an outpatient group programme.
By O’ Dowd H, Gladwell P, Rogers CA, Hollinghurst S, Gregory A.

No. 38

No. 39
The effectiveness and cost-effectiveness of computed tomography screening for coronary artery disease: systematic review.
By Waugh N, Black C, Walker S, McIntyre L, Cummins E, Hills G.

No. 40
What are the clinical outcome and cost-effectiveness of endoscopy undertaken by nurses when compared with doctors? A Multi-Institution Nurse Endoscopy Trial (MINuET).

No. 41
The clinical and cost-effectiveness of oxaliplatin and capetitabine for the adjuvant treatment of colon cancer: systematic review and economic evaluation.
By Farsol A, Eggington S, Paisley S, Tappenden P, Sutcliffe P.

No. 42
A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness.

No. 43
Telemedicine in dermatology: a randomised controlled trial.
By Bowns IR, Collins K, Walters SJ, McDonagh AJG.

No. 44

No. 45
Clinical effectiveness and cost-effectiveness of laparoscopic surgery for colorectal cancer: systematic reviews and economic evaluation.

No. 46
Etanercept and efalizumab for the treatment of psoriasis: a systematic review.

No. 47
Systematic reviews of clinical decision tools for acute abdominal pain.

No. 48
Evaluation of the ventricular assist device programme in the UK.
No. 49

No. 50
Amniocentesis results: investigation of anxiety. The ARIA trial.

Volume 11, 2007

No. 1
Pemetrexed disodium for the treatment of malignant pleural mesothelioma: a systematic review and economic evaluation.

No. 2
A systematic review and economic model of the clinical effectiveness and cost-effectiveness of docetaxel in combination with prednisone or prednisolone for the treatment of hormone-refractory metastatic prostate cancer.

No. 3
A systematic review of rapid diagnostic tests for the detection of tuberculosis infection.

No. 4
The clinical effectiveness and cost-effectiveness of stromium ranelate for the prevention of osteoporotic fragility fractures in postmenopausal women.
By Stevenson M, Davis S, Lloyd-Jones M, Beverley C.

No. 5
A systematic review of quantitative and qualitative research on the role and effectiveness of written information available to patients about individual medicines.

No. 6
oral naltrexone as a treatment for relapse prevention in formerly opioid-dependent drug users: a systematic review and economic evaluation.

No. 7
Glucocorticoid-induced osteoporosis: a systematic review and cost-utility analysis.
By Kanis JA, Stevenson M, McCloskey EV, Davis S, Lloyd-Jones M.

No. 8
Epidemiological, social, diagnostic and economic evaluation of population screening for genital chlamydial infection.

No. 9
Methadone and buprenorphine for the management of opioid dependence: a systematic review and economic evaluation.

No. 10
Exercise Evaluation Randomised Trial (EXERT1): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only.

No. 11
Interferon alfa (pegylated and non-pegylated) and ribavirin for the treatment of mild chronic hepatitis C: a systematic review and economic evaluation.
By Shepherd J, Jones J, Hartwell D, Davidson P, Price A, Waugh N.

No. 12
Systematic review and economic evaluation of bevacizumab and cetuximab for the treatment of metastatic colorectal cancer.
By Tappenden P, Jones R, Paisley S, Carroll C.

No. 13
A systematic review and economic evaluation of epoetin alfa, epoetin beta and darbepoetin alfa in anaemia associated with cancer, especially that attributable to cancer treatment.

No. 14
A systematic review and economic evaluation of statins for the prevention of coronary events.

No. 15
A systematic review of the effectiveness and cost-effectiveness of different models of community-based respite care for frail older people and their carers.

No. 16
Additional therapy for young children with spastic cerebral palsy: a randomised controlled trial.
By Weindling AM, Cunningham CC, Glenn SM, Edwards RT, Reeves DJ.

No. 17
Screening for type 2 diabetes: literature review and economic modelling.

No. 18
The effectiveness and cost-effectiveness of cinacalcet for secondary hyperparathyroidism in end-stage renal disease patients on dialysis: a systematic review and economic evaluation.

No. 19
The clinical effectiveness and cost-effectiveness of gemcitabine for metastatic breast cancer: a systematic review and economic evaluation.
By Takeda AL, Jones J, Loveman E, Tan SC, Clegg AJ.

No. 20
A systematic review of duplex ultrasound, magnetic resonance angiography and computed tomography angiography for the diagnosis and assessment of symptomatic, lower limb peripheral arterial disease.

No. 21
The clinical effectiveness and cost-effectiveness of treatments for children with idiopathic steroid-resistant nephrotic syndrome: a systematic review.
By Colquitt JL, Kirby J, Green C, Cooper K, Trompeter RS.

No. 22
A systematic review of the routine monitoring of growth in children of primary school age to identify growth-related conditions.

No. 23
Systematic review of the effectiveness of preventing and treating Staphylococcus aureus carriage in reducing periportal catheter-related infections.
No. 24  The clinical effectiveness and cost of repetitive transcranial magnetic stimulation versus electroconvulsive therapy in severe depression: a multicentre pragmatic randomised controlled trial and economic analysis.

No. 25  A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment.
By Boyle J, McCartney E, Forbes J, O’Hare A.

No. 26  Hormonal therapies for early breast cancer: systematic review and economic evaluation.
By Hind D, Ward S, De Nigris E, Simpson E, Carroll C, Wyld L.

No. 27  Cardioprotection against the toxic effects of anthracyclines given to children with cancer: a systematic review.
By Bryant J, Picot J, Levitt G, Sullivan I, Baxter L, Clegg A.

No. 28  Adalimumab, etanercept and infliximab for the treatment of ankylosing spondylitis: a systematic review and economic evaluation.

No. 29  Prenatal screening and treatment strategies to prevent group B streptococcal and other bacterial infections in early infancy: cost-effectiveness and expected value of information analyses.

No. 30  Clinical effectiveness and cost-effectiveness of bone morphogenetic proteins in the non-healing of fractures and spinal fusion: a systematic review.

No. 31  A randomised controlled trial of postoperative radiotherapy following breast-conserving surgery in a minimum-risk older population. The PRIME trial.

No. 32  Current practice, accuracy, effectiveness and cost-effectiveness of the school entry hearing screen.

No. 33  The clinical effectiveness and cost-effectiveness of inhaled insulin in diabetes mellitus: a systematic review and economic evaluation.
By Black C, Cummins E, Royle P, Philip S, Waugh N.

No. 34  Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis.

No. 35  The Birmingham Rehabilitation Uptake Maximisation Study (BRUM). Homebased compared with hospital-based cardiac rehabilitation in a multi-ethnic population: cost-effectiveness and patient adherence.

No. 36  A systematic review of the clinical, public health and cost-effectiveness of rapid diagnostic tests for the detection and identification of bacterial intestinal pathogens in faeces and food.

No. 37  A randomised controlled trial examining the longer-term outcomes of standard versus new antiepileptic drugs. The SANAD trial.

No. 38  Clinical effectiveness and cost-effectiveness of different models of managing long-term oral anti-coagulation therapy: a systematic review and economic modelling.

No. 39  A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder.

No. 40  Taxanes for the adjuvant treatment of early breast cancer: systematic review and economic evaluation.
By Ward S, Simpson E, Davis S, Hind D, Rees A, Wilkinson A.

No. 41  The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation.

No. 42  Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models.
By Davis A, Smith P, Ferguson M, Stephens D, Gianopoulous I.

No. 43  Contamination in trials of educational interventions.

No. 44  Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers.
By Facy K, Bradbury I, Laking G, Payne E.

No. 45  The effectiveness and cost-effectiveness of carmustine implants and temozolomide for the treatment of newly diagnosed high-grade glioma: a systematic review and economic evaluation.

No. 46  Drug-eluting stents: a systematic review and economic evaluation.

No. 47  The clinical effectiveness and cost-effectiveness of cardiac resynchronisation (biventricular pacing) for heart failure: systematic review and economic model.

No. 48  Recruitment to randomised trials: strategies for trial enrolment and participation study. The STEPS study.
By Campbell MK, Snowden C, Francis D, Elbourne D, McDonald AM, Knight R, et al.

© 2010 Queen’s Printer and Controller of HMSO. All rights reserved.
No. 49  
Cost-effectiveness of functional  
cardiac testing in the diagnosis and  
management of coronary artery  
disease: a randomised controlled trial.  
The CECat trial.  
By Sharples L, Hughes V, Crean A,  

No. 50  
Evaluation of diagnostic tests when  
there is no gold standard: A review of  
methods.  
By Rutjes AWS, Reitsma JB, Coomarasamy A, Khan KS,  
Bosuut PMM.

No. 51  
Systematic reviews of the clinical  
effectiveness and cost-effectiveness of  
proton pump inhibitors in acute upper  
gastrointestinal bleeding.  
By Leontiadis GI, Sreedharan A, Dorward S, Barton P, Delaney B,  
Howden CW, et al.

No. 52  
A review and critique of modelling in  
prioritising and designing screening  
programmes.  

No. 53  
An assessment of the impact of the  
NHS Health Technology Assessment  
Programme.  
By Hamney S, Buxton M, Green C,  
Coullson D, Rafferty J.

Volume 12, 2008

No. 1  
A systematic review and economic  
model of switching from  
nonglycopeptide to glycopeptide  
antibiotic prophylaxis for surgery.  
By Cranny G, Elliott R, Weatherly H,  
Chambers D, Hawkins N, Myers L, et al.

No. 2  
‘Cut down to quit’ with nicotine  
replacement therapies in smoking  
cessation: a systematic review of  
effectiveness and economic analysis.  
By Wang D, Connock M, Barton P,  
Fry-Smith A, Aveyard P, Moore D.

No. 3  
A systematic review of the effectiveness  
of strategies for reducing fracture risk  
in children with juvenile idiopathic  
arthritis with additional data on long- 
term risk of fracture and cost of disease  
management.  
By Thornton J, Ashcroft D, O’Neill T,  

No. 4  
Does befriending by trained lay workers  
improve psychological well-being and  
quality of life for carers of people  
with dementia, and at what cost? A  
randomised controlled trial.  
By Charlesworth G, Shepstone L,  
Wilson E, Thalanany M, Mugford M,  
Poland F.

No. 5  
A multi-centre retrospective cohort  
study comparing the efficacy, safety  
and cost-effectiveness of hysterectomy  
and uterine artery embolisation for  
the treatment of symptomatic uterine  
fibroids. The HOPEFUL study.  

No. 6  
Methods of prediction and prevention  
of pre-eclampsia: systematic reviews of  
accuracy and effectiveness literature  
with economic modelling.  
By Meads CA, Cranny G, Elliott R, Weatherly H,  
Fox D, Golder S, et al.

No. 7  
The use of economic evaluations in  
NHS decision-making: a review and  
empirical investigation.  
By Williams I, McIver S, Moore D,  
Bryan S.

No. 8  
Stapled haemorrhoidectomy  
(haemorrhoidopexy) for the treatment  
of haemorrhoids: a systematic review  
and economic evaluation.  
By Burch J, Epstein D, Baba-Akbari A,  
Weatherly H, Fox D, Golder S, et al.

No. 9  
The clinical effectiveness of diabetes  
education models for Type 2 diabetes:  
a systematic review.  
By Loveman E, Frampton GK,  
Clegg AJ.

No. 10  
Payment to healthcare professionals for  
patient recruitment to trials: systematic  
review and qualitative study.  
By Rafferty J, Bryant J, Powell J,  
Kerr C, Hawker S.

No. 11  
Cyclooxygenase-2 selective non-  
steroidal anti-inflammatory drugs  
(etodolac, meloxicam, celecoxib,  
rofecoxib, etoricoxib, valdecoxib and  
lumiracoxib) for osteoarthritis and  
rheumatoid arthritis: a systematic  
review and economic evaluation.  
By Chen Y-F, Johanputra P, Barton P,  

No. 12  
The clinical effectiveness and cost-  
effectiveness of central venous catheters  
treated with anti-infective agents in  
preventing bloodstream infections:  
a systematic review and economic  
evaluation.  

No. 13  
Stepped treatment of older adults on  
laxatives. The STOOL trial.  
By Mihaylov S, Stark C, McColl E,  

No. 14  
A randomised controlled trial of  
cognitive behaviour therapy in  
adolescents with major depression  
treated by selective serotonin reuptake  
inhibitors. The ADAPT trial.  
By Goodyer IM, Dubicka B, Wilkinson P, Kelvin R, Roberts C,  
Byford S, et al.

No. 15  
The use of irinotecan, oxaliplatin  
and raltitrexed for the treatment of  
advanced colorectal cancer: systematic  
review and economic evaluation.  
By Hind D, Tappenden P, Tumur I,  
Eggington E, Sutcliffe P, Ryan A.

No. 16  
Ranibizumab and pegaptanib for  
the treatment of age-related macular  
degeneration: a systematic review and  
economic evaluation.  
By Colquitt JL, Jones J, Tan SC,  
Takeda A, Clegg AJ, Price A.

No. 17  
Systematic review of the clinical  
effectiveness and cost-effectiveness  
of 64-slice or higher computed  
tomography angiography as an  
alternative to invasive coronary  
angiography in the investigation  
of coronary artery disease.  
By Mowatt G, Cummins E, Waugh N,  

No. 18  
Structural neuroimaging in psychosis:  
a systematic review and economic  
evaluation.  
By Albon E, Tsourapas A, Frew E,  

No. 19  
Systematic review and economic  
analysis of the comparative  
effectiveness of different inhaled  
corticosteroids and their usage with  
long-acting beta, agonists for the  
treatment of chronic asthma in adults  
and children aged 12 years and over.  
By Shepherd J, Rogers G, Anderson R, Main C, Thompson-Coon J,  
Hartwell D, et al.
No. 20
Systematic review and economic analysis of the comparative effectiveness of different inhaled corticosteroids and their usage with long-acting beta_2_ agonists for the treatment of chronic asthma in children under the age of 12 years.

No. 21
Ezetimibe for the treatment of hypercholesterolaemia: a systematic review and economic evaluation.

No. 22
Topical or oral ibuprofen for chronic knee pain in older people. The TOIB study.

No. 23
A prospective randomised comparison of minor surgery in primary and secondary care. The MiSTIC trial.

No. 24
A review and critical appraisal of measures of therapist–patient interactions in mental health settings.

No. 25
The clinical effectiveness and cost-effectiveness of screening programmes for amблиopía and strabismus in children up to the age of 4–5 years: a systematic review and economic evaluation.
By Carlton J, Karnon J, Czoski-Murray C, Smith KJ, Marr J.

No. 26
A systematic review of the clinical effectiveness and cost-effectiveness and economic modelling of minimal incision total hip replacement approaches in the management of arthritic disease of the hip.

No. 27
A preliminary model-based assessment of the cost–utility of a screening programme for early age-related macular degeneration.

No. 28
Intravenous magnesium sulphate and sotalol for prevention of atrial fibrillation after coronary artery bypass surgery: a systematic review and economic evaluation.
By Shepherd J, Jones J, Frampton GK, Tanajewski L, Turner D, Price A.

No. 29
Absorbent products for urinary/faecal incontinence: a comparative evaluation of key product categories.

No. 30
A systematic review of repetitive functional task practice with modelling of resource use, costs and effectiveness.

No. 31
The effectiveness and cost-effectiveness of minimal access surgery amongst people with gastro-oesophageal reflux disease – a UK collaborative study. The REFLUX trial.

No. 32
Time to full publication of studies of anti-cancer medicines for breast cancer and the potential for publication bias: a short systematic review.
By Takeda A, Loveman E, Harris P, Hartwell D, Welch K.

No. 33
Performance of screening tests for child physical abuse in accident and emergency departments.
By Woodman J, Pitt M, Wentz R, Taylor B, Hodes D, Gilbert RE.

No. 34
Curative catheter ablation in atrial fibrillation and typical atrial flutter: systematic review and economic evaluation.

No. 35
Systematic review and economic modelling of effectiveness and cost utility of surgical treatments for men with benign prostatic enlargement.

No. 36
Immunophrophylaxis against respiratory syncitial virus (RSV) with palivizumab in children: a systematic review and economic evaluation.
By Wang D, Cummins C, Bayliss S, Sandercock J, Burd A.

Volume 13, 2009

No. 1
Deferasirox for the treatment of iron overload associated with regular blood transfusions (transfusional haemosiderosis) in patients suffering with chronic anaemia: a systematic review and economic evaluation.

No. 2
Thrombophilia testing in people with venous thromboembolism: systematic review and cost-effectiveness analysis.
By Simpson EL, Stevenson MD, Rawlin A, Papaioannou D.

No. 3
Surgical procedures and non-surgical devices for the management of non-apnoeeic snoring: a systematic review of clinical effects and associated treatment costs.
By Main C, Liu Z, Welch K, Weiner G, Quentin Jones S, Stein K.

No. 4
Continuous positive airway pressure devices for the treatment of obstructive sleep apnoea–hypopnoea syndrome: a systematic review and economic analysis.

No. 5
Use of classical and novel biomarkers as prognostic risk factors for localised prostate cancer: a systematic review.

No. 6
The harmful health effects of recreational ecstasy: a systematic review of observational evidence.

No. 7
Systematic review of the clinical effectiveness and cost-effectiveness of oesophageal Doppler monitoring in critically ill and high-risk surgical patients.

No. 8
The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports.
By Taylor RS, Elston J.

No. 9
Controlling Hypertension and Hypotension Immediately Post Stroke (CHHIPS) – a randomised controlled trial.
No. 10 Routine antenatal anti-D prophylaxis for RhD-negative women: a systematic review and economic evaluation.
By Pilgrim H, Lloyd-Jones M, Rees A.

No. 11 Amantadine, oseltamivir and zanamivir for the prophylaxis of influenza (including a review of existing guidance no. 67): a systematic review and economic evaluation.

No. 12 Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods.
By Hobart J, Cano S.

No. 13 Treatment of severe ankle sprain: a pragmatic randomised controlled trial comparing the clinical effectiveness and cost-effectiveness of three types of mechanical ankle support with tubular bandage. The CAST trial.
By Cooke MW, Marsh JL, Clark M, Nakash R, Jarvis RM, Hutton JL, et al., on behalf of the CAST trial group.

No. 14 Non-occupational postexposure prophylaxis for HIV: a systematic review.
By Bryant J, Baxter L, Hird S.

No. 15 Blood glucose self-monitoring in type 2 diabetes: a randomised controlled trial.

No. 16 How far does screening women for domestic (partner) violence in different health-care settings meet criteria for a screening programme? Systematic reviews of nine UK National Screening Committee criteria.

No. 17 Spinal cord stimulation for chronic pain of neuropathic or ischaemic origin: systematic review and economic evaluation.
By Simpson, EL, Duenas A, Holmes MW, Papaioannou D, Chilcott J.

No. 18 The role of magnetic resonance imaging in the identification of suspected acoustic neuroma: a systematic review of clinical and cost-effectiveness and natural history.

No. 19 Dipsticks and diagnostic algorithms in urinary tract infection: development and validation, randomised trial, economic analysis, observational cohort and qualitative study.

No. 20 Systematic review of respite care in the frail elderly.

No. 21 Neuroleptics in the treatment of aggressive challenging behaviour for people with intellectual disabilities: a randomised controlled trial (NACHBID).

No. 22 Randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of selective serotonin reuptake inhibitors plus supportive care, versus supportive care alone, for mild to moderate depression with somatic symptoms in primary care: the THREAD (THREshold for AntiDepressant response) study.

No. 23 Diagnostic strategies using DNA testing for hereditary haemochromatosis in at-risk populations: a systematic review and economic evaluation.

No. 24 Enhanced external counterpulsation for the treatment of stable angina and heart failure: a systematic review and economic analysis.

No. 25 Development of a decision support tool for primary care management of patients with abnormal liver function tests without clinically apparent liver disease: a record-linkage population cohort study and decision analysis (ALFIE).

No. 26 A systematic review of presumed consent systems for deceased organ donation.
By Rithalia A, McDaid C, Suekarran S, Norman G, Myers L, Swoden A.

No. 27 Paracetamol and ibuprofen for the treatment of fever in children: the PITCH randomised controlled trial.

No. 28 A randomised controlled trial to compare minimally invasive glucose monitoring devices with conventional monitoring in the management of insulin-treated diabetes mellitus (MITRE).

No. 29 Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making.
By Andronis L, Barton P, Bryan S.

Suppl. 1 Trastuzumab for the treatment of primary breast cancer in HER2-positive women: a single technology appraisal.
By Ward S, Pilgrim H, Hind D.

Docetaxel for the adjuvant treatment of early node-positive breast cancer: a single technology appraisal.
By Chilcott J, Lloyd Jones M, Wilkinson A.

The use of pacitaxel in the management of early stage breast cancer.

Rituximab for the first-line treatment of stage III/IV follicular non-Hodgkin’s lymphoma.

Bortezomib for the treatment of multiple myeloma patients.

Fludarabine phosphate for the first-line treatment of chronic lymphocytic leukaemia.

Erlotinib for the treatment of relapsed non-small cell lung cancer.

Cetuximab plus radiotherapy for the treatment of locally advanced squamous cell carcinoma of the head and neck.

Infliximab for the treatment of adults with psoriasis.
By Loveman E, Turner D, Hartwell D, Cooper K, Clegg A.
No. 30 Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PnNDER trial.

No. 31 The effect of different treatment durations of clopidogrel in patients with non-ST-segment elevation acute coronary syndromes: a systematic review and value of information analysis.

No. 32 Systematic review and individual patient data meta-analysis of diagnosis of heart failure, with modelling of implications of different diagnostic strategies in primary care.

No. 33 A multicentre randomised controlled trial of the use of continuous positive airway pressure and non-invasive positive pressure ventilation in the early treatment of patients presenting to the emergency department with severe acute cardiogenic pulmonary oedema: the 3CPO trial.
By Gray AJ, Goodacre S, Newby DE, Masson MA, Sampson F, Dixon S, et al., on behalf of the 3CPO study investigators.

No. 34 Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation.
By Ara R, Pandor A, Stevens J, Rees A, Raffa R.

No. 35 Adefovir dipivoxil and pegylated interferon alpha for the treatment of chronic hepatitis B: an updated systematic review and economic evaluation.

No. 36 Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis.

No. 37 A double-blind randomised placebo-controlled trial of topical intranasal corticosteroids in 4- to 11-year-old children with persistent bilateral otitis media with effusion in primary care.

No. 38 The effectiveness and cost-effectiveness of methods of storing donated kidneys from deceased donors: a systematic review and economic model.
By Bond M, Pitt M, Akoh J, Moxham T, Hoyle M, Anderson R.

No. 39 Rehabilitation of older patients: day hospital compared with rehabilitation at home. A randomised controlled trial.
By Parker SG, Oliver P, Pennington M, Bond J, Jagger C, Enderby PM, et al.

No. 40 Breastfeeding promotion for infants in neonatal units: a systematic review and economic analysis.

No. 41 The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation.

No. 42 Rapid testing for group B streptococcus during labour: a test accuracy study with evaluation of acceptability and cost-effectiveness.

No. 43 Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling.

No. 44 The effectiveness and cost-effectiveness of cochlear implants for severe to profound deafness in children and adults: a systematic review and economic model.

Suppl. 2 Gemcitabine for the treatment of metastatic breast cancer.
By Jones J, Takeda A, Tan SC, Cooper K, Loveman E, Clegg A.

Varenicline in the management of smoking cessation: a single technology appraisal.
By Hind D, Tappenden P, Peters J, Kenjegaliyeva K.

Alefacept for the treatment of acute ischaemic stroke: a single technology appraisal.
By Lloyd Jones M, Holmes M, Rituximab for the treatment of rheumatoid arthritis.

Omalizumab for the treatment of severe persistent allergic asthma.

Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin’s lymphoma.
By Boland A, Bagust A, Hockenhull J, Davis H, Chi F, Dickson R.

Adalimumab for the treatment of psoriasis.
By Turner D, Picot J, Cooper K, Loveman E.

Dabigatran etexilate for the prevention of venous thromboembolism in patients undergoing elective hip and knee surgery: a single technology appraisal.
By Holmes M, C Carroll C, Papaioannou D.

Romiplostim for the treatment of chronic immune or idiopathic thrombocytopenic purpura: a single technology appraisal.

Sunitinib for the treatment of gastrointestinal stromal tumours: a critique of the submission from Pfizer.
By Bond M, Hoyle M, Moxham T, Napier M, Anderson R.

No. 45 Vitamin K to prevent fractures in older women: systematic review and economic evaluation.
By Stevenson M, Lloyd-Jones M, Papaioannou D.

No. 46 The effects of biofeedback for the treatment of essential hypertension: a systematic review.
By Greenhalgh J, Dickson R, Dundar Y.

No. 47 A randomised controlled trial of the use of acidivor and/or prednisolone for the early treatment of Bell’s palsy: the BELLS study.

Suppl. 3 Lapatinib for the treatment of HER2-overexpressing breast cancer.
By Jones J, Takeda A, Picot J, von Keyserlingk C, Clegg A.

Infliximab for the treatment of ulcerative colitis.
By Hyde C, Bryan S, Juarez-Garcia A, Andronis I, Fry-Smith A.
Rimonabant for the treatment of overweight and obese people.

Telbivudine for the treatment of chronic hepatitis B infection.
By Hartwell D, Jones J, Harris P, Cooper K.

Entecavir for the treatment of chronic hepatitis B infection.
By Shepherd J, Gospodarevskaya E, Frampton G, Cooper K.

Febuxostat for the treatment of hyperuricaemia in people with gout: a single technology appraisal.
By Stevenson M, Pandor A.

Rivaroxaban for the prevention of venous thromboembolism: a single technology appraisal.
By Stevenson M, Scope A, Holmes M, Rees A, Kalentlichaler E.

Cetuximab for the treatment of recurrent and/or metastatic squamous cell carcinoma of the head and neck.

Mifamurtide for the treatment of osteosarcoma: a single technology appraisal.
By Pandor A, Fitzgerald P, Stevenson M, Papaioannou D.

Ustekinumab for the treatment of moderate to severe psoriasis.
By Gospodarevskaya E, Picot J, Cooper K, Loveman E, Takeda A.

Endovascular stents for abdominal aortic aneurysms: a systematic review and economic model.

Clinical and cost-effectiveness of epoprostenol, iloprost, bosentan, sitaxentan and sildenafil for pulmonary arterial hypertension within their licensed indications: a systematic review and economic evaluation.

Cessation of attention deficit hyperactivity disorder drugs

No. 52
The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation.

No. 53
Randomised preference trial of medical versus surgical termination of pregnancy less than 14 weeks’ gestation (TOPS).

No. 54
Randomised controlled trial of the use of three dressing preparations in the management of chronic ulceration of the foot in diabetes.

No. 55
VenUS II: a randomised controlled trial of larval therapy in the management of leg ulcers.

No. 56
A prospective randomised controlled trial and economic modelling of antimicrobial silver dressings versus non-adherent control dressings for venous leg ulcers: the VULCAN trial.

No. 57
Communication of carrier status information following universal newborn screening for sickle cell disorders and cystic fibrosis: qualitative study of experience and practice.
By Kai J, Ulph F, Cullinan T, Qureshi N.

No. 58
Antiviral drugs for the treatment of influenza: a systematic review and economic evaluation.

No. 59
Development of a toolkit and glossary to aid in the adaptation of health technology assessment (HTA) reports for use in different contexts.
By Chase D, Rosten C, Turner S, Hicks N, Milne R.

No. 60
Colour vision testing for diabetic retinopathy: a systematic review of diagnostic accuracy and economic evaluation.

No. 61
By Bond M, Wyatt K, Lloyd J, Welch K, Taylor R.

No. 62
Are adverse effects incorporated in economic models? An initial review of current practice.
By Craig D, McDaid C, Fonseca T, Stock C, Duffy S, Woolacott N.

Volume 14, 2010

No. 1
Multicentre randomised controlled trial examining the cost-effectiveness of contrast-enhanced high field magnetic resonance imaging in women with primary breast cancer scheduled for wide local excision (COMICE).

No. 2
Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation.

No. 3
The clinical effectiveness and cost-effectiveness of testing for cytochrome P450 polymorphisms in patients with schizophrenia treated with antipsychotics: a systematic review and economic evaluation.

No. 4
Systematic review of the clinical effectiveness and cost-effectiveness of photodynamic diagnosis and urine biomarkers (FISH, ImmunoCyt, NMP22) and cytology for the detection and follow-up of bladder cancer.
No. 5
Effectiveness and cost-effectiveness of arthroscopic lavage in the treatment of osteoarthritis of the knee: a mixed methods study of the feasibility of conducting a surgical placebo-controlled trial (the KORAL study).
By Campbell MK, Skea ZC, Sutherland AG, Cuthbertson BH, Entwistle VA, McDonald AM, et al.

No. 6
A randomised 2 × 2 trial of community versus hospital pulmonary rehabilitation for chronic obstructive pulmonary disease followed by telephone or conventional follow-up.
By Waterhouse JC, Walters SJ, Oluboyede Y, Lawson RA.

No. 7

No. 8
Dissemination and publication of research findings: an updated review of related biases.

No. 9
The effectiveness and cost-effectiveness of biomarkers for the prioritisation of patients awaiting coronary revascularisation: a systematic review and decision model.

No. 10
Comparison of case note review methods for evaluating quality and safety in health care.
By Hutchinson A, Coster JE, Cooper KL, McIntosh A, Walters SJ, Bath PA, et al.

No. 11
Clinical effectiveness and cost-effectiveness of continuous subcutaneous insulin infusion for diabetes: systematic review and economic evaluation.

No. 12
Self-monitoring of blood glucose in type 2 diabetes: systematic review.
By Clar C, Barnard K, Cummins E, Royle P, Waugh N.

No. 13

No. 14
Multicentre randomised controlled trial of the clinical and cost-effectiveness of a bypass-surgery-first versus a balloon-angioplasty-first revascularisation strategy for severe limb ischaemia due to infrainguinal disease. The Bypass versus Angioplasty in Severe Ischaemia of the Leg (BASIL) trial.

No. 15
A randomised controlled multicentre trial of treatments for adolescent anorexia nervosa including assessment of cost-effectiveness and patient acceptability – the TOuCAN trial.

No. 16
Randomised controlled trials for policy interventions: a review of reviews and meta-regression.

No. 17
Paracetamol and selective and non-selective non-steroidal anti-inflammatory drugs (NSAIDs) for the reduction of morphine-related side effects after major surgery: a systematic review.
By McDaid C, Maund E, Rice S, Wright K, Jenkins B, Woolacott N.

DOI: 10.3310/hta14180
Health Technology Assessment 2010; Vol. 14: No. 18
© 2010 Queen’s Printer and Controller of HMSO. All rights reserved.
Health Technology Assessment programme

Director,
Professor Tom Walley,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Deputy Director,
Professor Jon Nicholl,
Director, Medical Care Research Unit, University of Sheffield

Prioritisation Strategy Group

Chair,
Professor Tom Walley,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Deputy Chair,
Professor Jon Nicholl,
Director, Medical Care Research Unit, University of Sheffield

Dr Bob Coates,
Consultant Advisor, NETSCC, HTA

Dr Andrew Cook,
Consultant Advisor, NETSCC, HTA

Professor Robin E Ferner,
Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham

Dr Peter Davidson,
Director of Science Support, NETSCC, HTA

Dr Nick Hicks,
Director of NHS Support, NETSCC, HTA

Professor Paul Glasziou,
Professor of Evidence-Based Medicine, University of Oxford

Dr Edmund Jessop,
Medical Adviser, National Specialist, National Commissioning Group (NCG), Department of Health, London

Professor Paul Glasziou,
Professor of Evidence-Based Medicine, University of Oxford

Dr Nick Hicks,
Director of NHS Support, NETSCC, HTA

Professor Paul Glasziou,
Professor of Evidence-Based Medicine, University of Oxford

Dr Edmund Jessop,
Medical Adviser, National Specialist, National Commissioning Group (NCG), Department of Health, London

Professor Paul Glasziou,
Professor of Evidence-Based Medicine, University of Oxford

Dr Edmund Jessop,
Medical Adviser, National Specialist, National Commissioning Group (NCG), Department of Health, London

Ms Lynn Kerridge,
Chief Executive Officer, NETSCC and NETSCC, HTA

Dr Ruairidh Milne,
Director of Strategy and Development, NETSCC

Ms Kay Pattison,
Section Head, NHS R&D Programme, Department of Health

Dr Ruairidh Milne,
Director of Strategy and Development, NETSCC

Ms Kay Pattison,
Section Head, NHS R&D Programme, Department of Health

Ms Pamela Young,
Specialist Programme Manager, NETSCC, HTA

Ms Lynn Kerridge,
Chief Executive Officer, NETSCC and NETSCC, HTA

Dr Ruairidh Milne,
Director of Strategy and Development, NETSCC

Ms Kay Pattison,
Section Head, NHS R&D Programme, Department of Health

Dr Ruairidh Milne,
Director of Strategy and Development, NETSCC

Ms Kay Pattison,
Section Head, NHS R&D Programme, Department of Health

Ms Pamela Young,
Specialist Programme Manager, NETSCC, HTA

Dr Ruairidh Milne,
Director of Strategy and Development, NETSCC

Ms Kay Pattison,
Section Head, NHS R&D Programme, Department of Health

Ms Pamela Young,
Specialist Programme Manager, NETSCC, HTA

© 2010 Queen’s Printer and Controller of HMSO. All rights reserved.
## Diagnostic Technologies & Screening Panel

### Members

| Chair, Professor Paul Glasziou, Professor of Evidence-Based Medicine, University of Oxford |
| Dr Stephanie Dancer, Consultant Microbiologist, Hairmyres Hospital, East Kilbride |
| Dr Ron Gray, Consultant Clinical Epidemiologist, Department of Public Health, University of Oxford |
| Professor Paul D Griffiths, Professor of Radiology, University of Sheffield |
| Dr Jennifer J Kurinczuk, Consultant Clinical Epidemiologist, National Perinatal Epidemiology Unit, Oxford |
| Dr Susanne M Ludgate, Medical Director, Medicines & Healthcare Products Regulatory Agency, London |

| Deputy Chair, Dr David Elliman, Consultant Paediatrician and Honorary Senior Lecturer, Great Ormond Street Hospital, London |
| Professor Glyn Ewryn, Primary Medical Care Research Group, Swansea Clinical School, University of Wales |
| Dr David Elliman, Deputy Chair, Medicine, University of Oxford |
| Professor Lindsay Wilson, Consultant Microbiologist, Hull Royal Infirmary |
| Ms Dawn Talbot, Service User Representative |

| Observers |
| Dr Tim Elliott, Team Leader, Cancer Screening, Department of Health |
| Dr Catherine Moody, Programme Manager, Neuroscience and Mental Health Board |
| Dr Ursula Wells, Principal Research Officer, Department of Health |

## Pharmaceuticals Panel

### Members

| Chair, Professor Robin Ferner, Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham |
| Dr Peter Elton, Director of Public Health, Bury Primary Care Trust |
| Dr Ben Goldacre, Research Fellow, Division of Psychological Medicine and Psychiatry, King’s College London |
| Mrs Barbara Greggains, Service User Representative |
| Dr Bill Gutteridge, Medical Adviser, London Strategic Health Authority |
| Dr Dyfrig Hughes, Reader in Pharmacoeconomics and Deputy Director, Centre for Economics and Policy in Health, IMSCaR, Bangor University |
| Professor Jonathan Ledermann, Professor of Medical Oncology and Director of the Cancer Research UK and University College London Cancer Trials Centre |
| Dr Yoon K Loke, Senior Lecturer in Clinical Pharmacology, University of East Anglia |
| Professor Femi Oyebode, Consultant Psychiatrist and Head of Department, University of Birmingham |
| Dr Andrew Prentice, Senior Lecturer and Consultant Obstetrician and Gynaecologist, The Rosie Hospital, University of Cambridge |
| Dr Martin Shelly, General Practitioner, Leeds, and Associate Director, NHS Clinical Governance Support Team, Leicester |
| Dr Gillian Shepherd, Director, Health and Clinical Excellence, Merck Serono Ltd |
| Mrs Katrina Simister, Assistant Director New Medicines, National Prescribing Centre, Liverpool |
| Mr David Symes, Service User Representative |
| Dr Lesley Wise, Unit Manager, Pharmacoepidemiology Research Unit, VRMM, Medicines, National Prescribing Centre, Liverpool |

### Observers

| Ms Kay Pattison, Section Head, NHS R&D Programme, Department of Health |
| Mr Simon Reeve, Head of Clinical and Cost-Effectiveness, Medicines, Pharmacy and Industry Group, Department of Health |
| Dr Heike Weber, Programme Manager, Medical Research Council |
| Dr Ursula Wells, Principal Research Officer, Department of Health |

Current and past membership details of all HTA programme ‘committees’ are available from the HTA website (www.hta.ac.uk)
Therapeutic Procedures Panel

Members

Chair,
Dr John C Pounsford,
Consultant Physician, North Bristol NHS Trust

Deputy Chair,
Professor Scott Weich,
Professor of Psychiatry, Division of Health in the Community, University of Warwick, Coventry

Professor Jane Barlow,
Professor of Public Health in the Early Years, Health Sciences Research Institute, Warwick Medical School, Coventry

Ms Maree Barnett,
Acting Branch Head of Vascular Programme, Department of Health

Mrs Val Carlill,
Service User Representative

Mrs Anthea De Barton-Watson,
Service User Representative

Mr Mark Emberton,
Senior Lecturer in Oncological Urology, Institute of Urology, University College Hospital, London

Professor Steve Goodacre,
Professor of Emergency Medicine, University of Sheffield

Professor Christopher Griffiths,
Professor of Primary Care, Barts and The London School of Medicine and Dentistry

Mr Paul Hilton,
Consultant Gynaecologist and Urogynaecologist, Royal Victoria Infirmary, Newcastle upon Tyne

Professor Nicholas James,
Professor of Clinical Oncology, University of Birmingham, and Consultant in Clinical Oncology, Queen Elizabeth Hospital

Dr Peter Martin,
Consultant Neurologist, Addenbrooke’s Hospital, Cambridge

Mrs Val Carlill,
Service User Representative

Mrs Anthea De Barton-Watson,
Service User Representative

Mr Mark Emberton,
Senior Lecturer in Oncological Urology, Institute of Urology, University College Hospital, London

Professor Steve Goodacre,
Professor of Emergency Medicine, University of Sheffield

Professor Christopher Griffiths,
Professor of Primary Care, Barts and The London School of Medicine and Dentistry

Mr Paul Hilton,
Consultant Gynaecologist and Urogynaecologist, Royal Victoria Infirmary, Newcastle upon Tyne

Professor Nicholas James,
Professor of Clinical Oncology, University of Birmingham, and Consultant in Clinical Oncology, Queen Elizabeth Hospital

Dr Peter Martin,
Consultant Neurologist, Addenbrooke’s Hospital, Cambridge

Observers

Dr Phillip Leech,
Principal Medical Officer for Primary Care, Department of Health

Ms Kay Pattison,
Section Head, NHS R&D Programme, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Medical Research Council

Professor Tom Walley,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Mr Jim Reece
Service User Representative

Dr Karen Roberts,
Nurse Consultant, Dunston Hill Hospital Cottages

Dr Ursula Wells,
Principal Research Officer, Department of Health

Disease Prevention Panel

Members

Chair,
Dr Edmund Jessop,
Medical Adviser, National Specialist, National Commissioning Group (NCG), London

Deputy Chair,
Dr David Pencheon,
Director, NHS Sustainable Development Unit, Cambridge

Dr Elizabeth Fellow-Smith,
Medical Director, West London Mental Health Trust, Middlesex

Dr John Jackson,
General Practitioner, Parkway Medical Centre, Newcastle upon Tyne

Professor Mike Kelly,
Director, Centre for Public Health Excellence, NICE, London

Dr Chris McCall,
General Practitioner, The Hadleigh Practice, Corfe Mullen, Dorset

Ms Jeanett Martin,
Director of Nursing, BarnDoc Limited, Lewisham Primary Care Trust

Dr Julie Mytton,
Locum Consultant in Public Health Medicine, Bristol Primary Care Trust

Miss Nicky Mullany,
Service User Representative

Professor Ian Roberts,
Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine

Professor Ken Stein,
Senior Clinical Lecturer in Public Health, University of Exeter

Dr Julia Mytton,
Locum Consultant in Public Health Medicine, Bristol Primary Care Trust

Miss Nicky Mullany,
Service User Representative

Professor Ian Roberts,
Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine

Professor Ken Stein,
Senior Clinical Lecturer in Public Health, University of Exeter

Observers

Ms Christine McGuire,
Research & Development, Department of Health

Dr Caroline Stone,
Programme Manager, Medical Research Council

Dr Kieran Sweeney,
Honorary Clinical Senior Lecturer, Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth

Professor Carol Tannahill,
Glasgow Centre for Population Health

Professor Margaret Thorogood,
Professor of Epidemiology, University of Warwick Medical School, Coventry

Dr Kate Radford,
Senior Lecturer (Research), Clinical Practice Research Unit, University of Central Lancashire, Preston

Mr Jim Reece
Service User Representative

Dr Karen Roberts,
Nurse Consultant, Dunston Hill Hospital Cottages

Dr Ursula Wells,
Principal Research Officer, Department of Health

© 2010 Queen’s Printer and Controller of HMSO. All rights reserved.
Expert Advisory Network

Members

Professor Douglas Altman, Professor of Statistics in Medicine, Centre for Statistics in Medicine, University of Oxford

Professor John Bond, Professor of Social Gerontology & Health Services Research, University of Newcastle upon Tyne

Professor Andrew Bradbury, Professor of Vascular Surgery, Solihull Hospital, Birmingham

Mr Shaun Brogan, Chief Executive, Ridgeway Primary Care Group, Aylesbury

Mrs Stella Burnside OBE, Chief Executive, Regulation and Improvement Authority, Belfast

Ms Tracy Bury, Project Manager, World Confederation for Physical Therapy, London

Professor Iain T Cameron, Professor of Obstetrics and Gynaecology and Head of the School of Medicine, University of Southampton

Dr Christine Clark, Medical Writer and Consultant Pharmacist, Rossendale

Professor Collette Clifford, Professor of Nursing and Head of Research, The Medical School, University of Birmingham

Professor Barry Cookson, Director, Laboratory of Hospital Infection, Public Health Laboratory Service, London

Dr Carl Counsell, Clinical Senior Lecturer in Neurology, University of Aberdeen

Professor Howard Cuckle, Professor of Reproductive Epidemiology, Department of Paediatrics, Obstetrics & Gynaecology, University of Leeds

Dr Katherine Barton, Information Unit, MIND – The Mental Health Charity, London

Professor Carol Dezateux, Professor of Paediatric Epidemiology, Institute of Child Health, London

Mr John Dunning, Consultant Cardiothoracic Surgeon, Papworth Hospital NHS Trust, Cambridge

Mr Jonathan Earnshaw, Consultant Vascular Surgeon, Gloucestershire Royal Hospital, Gloucester

Professor Martin Eccles, Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne

Professor Pam Enderby, Dean of Faculty of Medicine, Institute of General Practice and Primary Care, University of Sheffield

Professor Gene Feder, Professor of Primary Care Research & Development, Centre for Health Sciences, Barts and The London School of Medicine and Dentistry

Mr Leonard R Fenwick, Chief Executive, Freeman Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher, Antenatal Teacher and Tutor and President, National Childbirth Trust, Hayfield

Professor Jayne Franklyn, Professor of Medicine, University of Birmingham

Mr Tam Fry, Honorary Chairman, Child Growth Foundation, London

Professor Fiona Gilbert, Consultant Radiologist and NCRN Member, University of Aberdeen

Professor Paul Gregg, Professor of Orthopaedic Surgical Science, South Tees Hospital NHS Trust

Bec Hanley, Co-director, TwoCan Associates, West Sussex

Dr Maryann L Hardy, Senior Lecturer, University of Bradford

Mrs Sharon Hart, Healthcare Management Consultant, Reading

Professor Robert E Hawkins, CRC Professor and Director of Medical Oncology, Christie CRC Research Centre, Christie Hospital NHS Trust, Manchester

Professor Richard Hobbs, Head of Department of Primary Care & General Practice, University of Birmingham

Professor Alan Horwich, Dean and Section Chairma, The Institute of Cancer Research, London

Professor Allen Hutchinson, Director of Public Health and Deputy Dean of SchARR, University of Sheffield

Professor Peter Jones, Professor of Psychiatry, University of Cambridge, Cambridge

Professor Stan Kaye, Cancer Research UK Professor of Medical Oncology, Royal Marsden Hospital and Institute of Cancer Research, Surrey

Dr Duncan Kesley, General Practitioner (Dr Burch & Pins), The Health Centre, Thame

Dr Donna Lamping, Research Degrees Programme Director and Reader in Psychology, Health Services Research Unit, London School of Hygiene and Tropical Medicine, London

Mr George Levy, Chief Executive, Motor Neurone Disease Association, Northampton

Professor James Lindsey, Professor of Psychiatry for the Elderly, University of Leicester

Professor Julian Little, Professor of Human Genome Epidemiology, University of Ottawa

Professor Alistaire McGuire, Professor of Health Economics, London School of Economics

Professor Rajan Madhok, Medical Director and Director of Public Health, Directorate of Clinical Strategy & Public Health, North & East Yorkshire & Northern Lincolnshire Health Authority, York

Professor Alexander Markham, Director, Molecular Medicine Unit, St James’s University Hospital, Leeds

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Susan Schonfield, Consultant in Public Health, Hillingdon Primary Care Trust, Middlesex

Dr Eamonn Sheridan, Consultant in Clinical Genetics, St James’s University Hospital, Leeds

Dr Margaret Somerville, Director of Public Health, Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown, Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Professor Alla Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Mrs Joan Webster, Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle, Clinical Co-director, National Co-ordinating Centre for Women’s and Children’s Health, Lympstone

Professor Miranda Mugford, Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson, Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick, National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveeler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price, Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg, Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sanderson, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Susan Schonfield, Consultant in Public Health, Hillingdon Primary Care Trust, Middlesex

Dr Eamonn Sheridan, Consultant in Clinical Genetics, St James’s University Hospital, Leeds

Dr Margaret Somerville, Director of Public Health, Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown, Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Professor Alla Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Mrs Joan Webster, Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle, Clinical Co-director, National Co-ordinating Centre for Women’s and Children’s Health, Lympstone

Professor Miranda Mugford, Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson, Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick, National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveeler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price, Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg, Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Susan Schonfield, Consultant in Public Health, Hillingdon Primary Care Trust, Middlesex

Dr Eamonn Sheridan, Consultant in Clinical Genetics, St James’s University Hospital, Leeds

Dr Margaret Somerville, Director of Public Health, Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown, Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Professor Alla Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Mrs Joan Webster, Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle, Clinical Co-director, National Co-ordinating Centre for Women’s and Children’s Health, Lympstone

Professor Miranda Mugford, Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson, Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick, National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveeler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price, Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg, Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Susan Schonfield, Consultant in Public Health, Hillingdon Primary Care Trust, Middlesex

Dr Eamonn Sheridan, Consultant in Clinical Genetics, St James’s University Hospital, Leeds

Dr Margaret Somerville, Director of Public Health, Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown, Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Professor Alla Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Mrs Joan Webster, Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle, Clinical Co-director, National Co-ordinating Centre for Women’s and Children’s Health, Lympstone

Professor Miranda Mugford, Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson, Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick, National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveeler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price, Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg, Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Susan Schonfield, Consultant in Public Health, Hillingdon Primary Care Trust, Middlesex

Dr Eamonn Sheridan, Consultant in Clinical Genetics, St James’s University Hospital, Leeds

Dr Margaret Somerville, Director of Public Health, Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown, Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Professor Alla Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Mrs Joan Webster, Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle, Clinical Co-director, National Co-ordinating Centre for Women’s and Children’s Health, Lympstone

Professor Miranda Mugford, Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson, Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick, National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveeler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price, Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg, Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Susan Schonfield, Consultant in Public Health, Hillingdon Primary Care Trust, Middlesex

Dr Eamonn Sheridan, Consultant in Clinical Genetics, St James’s University Hospital, Leeds

Dr Margaret Somerville, Director of Public Health, Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown, Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Professor Alla Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Mrs Joan Webster, Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle, Clinical Co-director, National Co-ordinating Centre for Women’s and Children’s Health, Lympstone

Professor Miranda Mugford, Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson, Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick, National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveeler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price, Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg, Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Susan Schonfield, Consultant in Public Health, Hillingdon Primary Care Trust, Middlesex

Dr Eamonn Sheridan, Consultant in Clinical Genetics, St James’s University Hospital, Leeds

Dr Margaret Somerville, Director of Public Health, Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown, Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Professor Alla Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Mrs Joan Webster, Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle, Clinical Co-director, National Co-ordinating Centre for Women’s and Children’s Health, Lympstone
Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (www.hta.ac.uk) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.