National Guideline Alliance

Version 2.0

Mental health of adults in contact with the criminal justice system

Identification and management of mental health problems and integration of care for adults in contact with the criminal justice system

NICE Guideline

Methods, evidence and recommendations

4 October 2016

Draft for consultation

Commissioned by the National Institute for Health and Care Excellence

Guideline Committee Group members and National Guideline Alliance (NGA) review team

Nick Kosky (Chair)

Medical Director & Consultant Psychiatrist, Dorset Healthcare University Foundation Trust, Forston Clinic

Richard Byng (Deputy Chair)

Professor of Primary Care Research, Primary Care Group, Plymouth University Peninsula Schools of Medicine and Dentistry; Deputy Director, Peninsula CLAHRC for the south west; General Practitioner, Mount Gould Primary Care Centre and GP with Special Interest in Mental Health (Honorary, Plymouth Community Healthcare and The Zone, Plymouth)

Steve Pilling (Facilitator)

Clinical Advisor, NGA

Vikki Baker

Joint Service Director, Consultant Clinical Psychologist, Resettle

Annie Bartlett

Reader and Honorary Consultant in Forensic Psychiatry SGUL and CNWL FT; Clinical Director Offender Care CNWL FT

Diana Binding

Assistant Chief Executive – Gwent; Wales Community Rehabilitation Service (CRC)

Richard Bradshaw

Chair, Physical Health in Prisons, NICE Guideline

Marie Brown (September 2015 – February 2016)

Research Assistant, (former) NCCMH

Steffan Davies

Consultant Forensic Psychiatrist, Offender Health Services, Northamptonshire Healthcare NHS Foundation Trust

Stephen Habgood

Retired senior governor; Chairman of PAPYRUS

Teresa Hall (until March 2015)

Research Assistant, (former) NCCMH

Kay Isaacs

Manager Criminal Justice Mental Health Team Abertawe Bro Morgannwg University Health Board

Julie Kane (January 2015 onwards)

Project Manager, NGA

May Oo Khin (November 2015 onwards)

Research Fellow, NGA

Sunil Lad

Principal Counselling Psychologist Offender Health, Northamptonshire Healthcare Foundation Trust

Naomi Lumsdaine

Women Prisoners' Caseworker, Prisoners' Advice Service; Qualified barrister

Kerry Manson

Clinical Psychologist

Odette Megnin-Viggars (until July 2015)

Senior Systematic Reviewer, NGA

Maryla Moulin (until January 2015)

Project Manager, (former) NCCMH

Tony O'Connell

Detective Constable, Dorset Police

Laura O'Shea (March 2015 onwards)

Research Assistant, NGA

Emma Seymour (March 2016 onwards)

Research Assistant, NGA

Leroy Simpson

Service User

Eric Slade

Health Economist, NGA

Nicole Stanbury

Service User

Sarah Stockton

Senior Information Scientist, NGA

Julia Tabreham

Chief Executive, Carers Federation Ltd (Retired)

Jenny Talbot

Director, Care not Custody, Prison Reform Trust

Clare Taylor

Senior Editor, (former) NCMMH

Mark Warren

Forensic Liaison Practitioner, CWM TAF University Health Board, Wales

Geoffrey White

Prison Officer

Joanne White

Primary Care Mental Health Team Leader, HMP LEEDS

Acknowledgments

The Guideline Committee and the National Guideline Alliance would like to thank Professor Huw Williams, Associate Professor of Clinical Neuropsychology and Co-Director of the Centre for Clinical Neuropsychology Research (CCNR) at the University of Exeter, provided expert testimony on the relationship between traumatic brain injury (TBI) and mental health problems in young offenders.

1 Preface

This guideline has been developed to advise on identification and management of mental health problems and integration of care for adults in contact with the criminal justice system. The guideline recommendations have been developed by a multidisciplinary team of healthcare professionals, criminal justice system professionals and people with mental health problems who have been in contact with the criminal justice system, their carers and guideline methodologists after careful consideration of the best available evidence. It is intended that the guideline will be useful to clinicians and service commissioners in the identification and management of mental health problems and integration of care for adults in contact with the criminal justice system (see Appendix A for more details on the scope of the guideline).

Although the evidence base is rapidly expanding, there are a number of major gaps. The guideline makes a number of research recommendations specifically to address gaps in the evidence base. In the meantime, it is hoped that the guideline will assist clinicians, and people with mental health problems in contact with the criminal justice system and their carers, by identifying the merits of particular treatment approaches where the evidence from research and clinical experience exists.

1.1 National clinical guidelines

1.1.1 What are clinical guidelines?

Clinical guidelines are 'systematically developed statements that assist clinicians and service users in making decisions about appropriate treatment for specific conditions' (Mann, 1996). They are derived from the best available research evidence, using predetermined and systematic methods to identify and evaluate the evidence relating to the specific condition in question. Where evidence is lacking, the guidelines include statements and recommendations based upon the consensus statements developed by the Guideline Development Group (GC).

Clinical guidelines are intended to improve the process and outcomes of healthcare in a number of different ways. They can:

- provide up-to-date evidence-based recommendations for the management of conditions and disorders by healthcare professionals
- be used as the basis to set standards to assess the practice of healthcare professionals
- form the basis for education and training of healthcare professionals
- assist service users and their carers in making informed decisions about their treatment and care
- improve communication between healthcare professionals, service users and their carers
- help identify priority areas for further research.

1.1.2 Uses and limitations of clinical guidelines

Guidelines are not a substitute for professional knowledge and clinical judgement. They can be limited in their usefulness and applicability by a number of different factors: the availability of high-quality research evidence, the quality of the methodology used in the development of the guideline, the generalisability of research findings and the uniqueness of individuals.

Although the quality of research in this field is variable, the methodology used here reflects current international understanding on the appropriate practice for guideline development (Appraisal of Guidelines for Research and Evaluation Instrument [AGREE]; www.agreetrust.org; AGREE Collaboration, 2003), ensuring the collection and selection of the best research evidence available and the systematic generation of treatment recommendations applicable to the majority of people with mental health problems in contact with the criminal justice system. However, there will always be some people and situations where clinical guideline recommendations are not readily applicable. This guideline does not, therefore, override the individual responsibility of healthcare professionals to make appropriate decisions in the circumstances of the individual, in consultation with the person with mental health problems in contact with the criminal justice system or their carer.

In addition to the clinical evidence, cost-effectiveness information, where available, is taken into account in the generation of statements and recommendations in clinical guidelines. While national guidelines are concerned with clinical and cost effectiveness, issues of affordability and implementation costs are to be determined by the National Health Service (NHS).

In using guidelines, it is important to remember that the absence of empirical evidence for the effectiveness of a particular intervention is not the same as evidence for ineffectiveness. In addition, and of particular relevance in mental health, evidence-based treatments are often delivered within the context of an overall treatment programme including a range of activities, the purpose of which may be to help engage the person and provide an appropriate context for the delivery of specific interventions. It is important to maintain and enhance the service context in which these interventions are delivered, otherwise the specific benefits of effective interventions will be lost. Indeed, the importance of organising care in order to support and encourage a good therapeutic relationship is at times as important as the specific treatments offered.

1.13 Why develop national guidelines?

The National Institute for Health and Care Excellence (NICE) was established as a Special Health Authority for England and Wales in 1999, with a remit to provide a single source of authoritative and reliable guidance for service users, professionals and the public. NICE guidance aims to improve standards of care, diminish unacceptable variations in the provision and quality of care across the NHS, and ensure that the health service is person-centred. All guidance is developed in a transparent and collaborative manner, using the best available evidence and involving all relevant stakeholders.

NICE generates guidance in a number of different ways, three of which are relevant here. First, national guidance is produced by the Technology Appraisal Committee to give robust advice about a particular treatment, intervention, procedure or other health technology. Second, NICE commissions public health intervention guidance focused on types of activity (interventions) that help to reduce people's risk of developing a disease or condition, or help to promote or maintain a healthy lifestyle. Third, NICE commissions the production of national clinical guidelines focused upon the overall treatment and management of a specific condition. To enable this latter development, NICE has established the National Guideline Alliance in conjunction with a range of professional organisations involved in healthcare.

1.14 From national clinical guidelines to local protocols

Once a national guideline has been published and disseminated, local healthcare groups will be expected to produce a plan and identify resources for implementation, along with appropriate timetables. Subsequently, a multidisciplinary group involving commissioners of healthcare, primary care and specialist mental health professionals, service users and carers should undertake the translation of the implementation plan into local protocols, taking into account both the recommendations set out in this guideline and the priorities in the National

Service Framework for Mental Health (Department of Health, 1999) and related documentation. The nature and pace of the local plan will reflect local healthcare needs and the nature of existing services; full implementation may take a considerable time, especially where substantial training needs are identified.

1.1.5 Auditing the implementation of clinical guidelines

This guideline identifies key areas of clinical practice and service delivery for local and national audit. Although the generation of audit standards is an important and necessary step in the implementation of this guidance, a more broadly-based implementation strategy will be developed. Nevertheless, it should be noted that the Care Quality Commission in England, and the Healthcare Inspectorate Wales, will monitor the extent to which commissioners and providers of health and social care and Health Authorities have implemented these guidelines.

1.2 The national mental health of adults in contact with the criminal justice system guideline

1.2.1 Who has developed this guideline?

This guideline has been commissioned by NICE and developed within the National Guideline Alliance (NGA). The NGA is a collaboration of the professional organisations involved in the field of mental health, national service user and carer organisations, a number of academic institutions and NICE. The NGA is funded by NICE and is led by a partnership between the Royal College of Psychiatrists and the British Psychological Society's Centre for Outcomes Research and Effectiveness, based at University College London.

The GC was convened by the NGA and supported by funding from NICE. The GC included people with mental health problems who have been in contact with the criminal justice system and carers, and professionals from [amend as appropriate] psychiatry, clinical psychology, general practice, nursing, psychiatric pharmacy, and the private and voluntary sectors.

Staff from the NGA provided leadership and support throughout the process of guideline development, undertaking systematic searches, information retrieval, appraisal and systematic review of the evidence. Members of the GC received training in the process of guideline development from NGA staff, and the service users and carers received training and support from the NICE Patient and Public Involvement Programme. The NICE Guidelines Technical Adviser provided advice and assistance regarding aspects of the guideline development process.

All GC members made formal declarations of interest at the outset, which were updated at every GC meeting. The GC met a total of [insert number of meeting] times throughout the process of guideline development. The GC was supported by the NGA technical team, with additional expert advice from special advisers where needed. The group oversaw the production and synthesis of research evidence before presentation. All statements and recommendations in this guideline have been generated and agreed by the whole GC.

1.2.2 For whom is this guideline intended?

This guideline will be relevant for adults with mental health problems who are in contact with the criminal justice system and covers the care provided by primary, community, secondary, tertiary and other healthcare professionals who have direct contact with, and make decisions concerning the care of, adults with mental health problems who are in contact with the criminal justice system.

1.2.3 Specific aims of this guideline

The guideline makes recommendations for the identification and management of mental health problems and integration of care for adults in contact with the criminal justice system. It aims to:

- improve access and engagement with treatment and services for people with mental health problems who are in contact with the criminal justice system
- evaluate the role of specific psychological, psychosocial and pharmacological interventions in the treatment of mental health problems within the criminal justice system
- evaluate the role of specific service-level interventions for people with mental health disorders in contact with the criminal justice system
- integrate the above to provide best-practice advice on the care of individuals throughout the course of their treatment
- promote the implementation of best clinical practice through the development of recommendations tailored to the requirements of the NHS in England and Wales.

1.2.4 The structure of this guideline

The guideline is divided into chapters, each covering a set of related topics. The first three chapters provide a general introduction to guidelines, an introduction to the topic of mental health problems of adults in contact with the criminal justice system and to the methods used to develop them. Chapter 4 to Chapter 7 provide the evidence that underpins the recommendations about the treatment and management of mental health problems of adults in contact with the criminal justice system

Each evidence chapter begins with a general introduction to the topic that sets the recommendations in context. Depending on the nature of the evidence, narrative reviews or meta-analyses were conducted, and the structure of the chapters varies accordingly. Where appropriate, details about current practice, the evidence base and any research limitations are provided. Where meta-analyses were conducted, information is given about both the interventions included and the studies considered for review. Clinical summaries are then used to summarise the evidence presented. Finally, recommendations related to each topic are presented at the end of each chapter. Where meta-analyses were conducted, the data are presented using forest plots in Appendix O. (see Table 1: Appendices for details).

Table 1: Appendices

Scope for the development of the clinical guideline	Appendix A
Declarations of interests by Guideline Development Group members	Appendix B
Special advisors to the Guideline Development Group	Appendix C
Stakeholders and experts who submitted comments in response to the consultation draft of the guideline	Appendix D
Researchers contacted to request information about unpublished or soon-to-be published studies	Appendix E
Analytic framework, review protocols and questions	Appendix F
Research recommendations	Appendix G
Clinical Evidence - Search strategies	Appendix H
HE Evidence - Search strategies	Appendix I
Clinical Evidence – Study characteristics, outcomes, methodology checklist for experience of care	Appendix J
Clinical Evidence – Study characteristics, outcomes, methodology checklist for recognition and assessment	Appendix K
Clinical Evidence – Study characteristics, outcomes, methodology checklist for interventions, service delivery, staff training	Appendix L
Clinical Evidence – Flow charts	Appendix M
Clinical Evidence – GRADE evidence profiles for all studies	Appendix N
Clinical Evidence – Forest plots for all studies	Appendix O
HE Evidence – HE review flow chart	Appendix P
HE Evidence - Completed HE checklists	Appendix Q
HE Evidence - Evidence tables	Appendix R
HE Evidence - Economic profiles	Appendix S
Clinical Evidence – NGT blank questionnaires	Appendix U
Clinical Evidence – NGT statements to recommendations	Appendix V
Clinical Evidence – Expert testimony	Appendix W

In the event that amendments or minor updates need to be made to the guideline, please check the NGA website (https://www.rcog.org.uk/en/about-us/nga/), where these will be listed and a corrected PDF file available to download.

2 Introduction

2.1 Mental Health and the Criminal Justice System

In 2014 over 1.7 million people in the United Kingdom were in contact with the criminal justice system (Ministry of Justice., 2009) many such contacts will be very limited and lead to no action on the part of criminal justice services. These people will have a range of mental health problems broadly similar to those of the general population in the United Kingdom with a prevalence across all disorders of about 20%. However, for those who have more extensive contact with the criminal justice system the picture is very different. For example, an estimated 39% of people detained in police custody have some form of mental disorder, and over 25% of residents in approved premises (previously known as bail hostels) have been found to have a psychiatric diagnosis. (Ministry of Justice., 2015b). An estimated 39% of adults serving community sentences (there are currently around 120,000 people with community sentences (Ministry of Justice., 2015b) have a mental disorder, and it has been estimated that over 90% of prisoners have at least one of the following psychiatric disorders; psychosis, anxiety or depression, personality disorder and alcohol or drug misuse. Some disorders such as personality disorders have a very high prevalence in the prison population (currently around 85,000 (MoJ, 2015)) approaching 60%, compared to 5% in the general population and the rate of psychotic disorders in those serving community sentences is 11% compared to 1% in the general population. There are other troubling discrepancies between those in the general population and those in the prison system, for example 76% of female remand prisoners compared to 40% of male remand prisoners have a common mental disorder (MoJ, 2015)). In addition to considerable differences in formal psychiatric disorders, self-harm is also very common among people in contact with the criminal justice system. Of people detained in police custody, 10% reported current suicidal thoughts and 18% a suicide attempt. (Kent & Gunasekaran, 2010).

An estimated 12% of people serving community sentences are at high risk of suicide. Among prisoners, 46% of men and 21% of women said they had attempted suicide at some point in their lives (Department of Health., 2014) . This is considerably higher than in the general UK population, with 6% of people saying they have previously attempted suicide. Among adults with mental health problems serving community sentences, an estimated 72% also screened positive for either an alcohol or a drug problem. Drug and alcohol is also high with an estimated 12% of adults serving community sentences having substantial or severe levels of drug misuse, and estimates of drug misuse and dependence on reception into prison range from 10–48% for male prisoners and 30–60% for female prisoners. 56% of people serving community sentences are hazardous drinkers and 60% of male prisoners) and 38% of female prisoners report hazardous drinking in the year before going to prison.

In addition to the common mental health and severe mental illness, there are other characteristics of the population in contact with the criminal justice system that present particular challenges. 7% of the prison population have a learning disability compared to 2% of the general population and 50% of the prison population suffer from degree of traumatic brain injury compared to approximately 0.56% (Headway, 2015) of the general population. These include the 13% of the prison population that are sexual offenders compared to 0.004% in the general population (Ministry of Justice., 2013a).

Black and minority ethnic (BME) groups are over-represented in the prison population. It is estimated that BME groups constitute 26% of the prison population compared with 9% of the overall population in England and Wales (Goodman & Ruggiero, 2008) . For BME groups, in particular young black men, contact with the criminal justice system may be an important route into mental health services, with BME groups found to be 40% more likely than white British groups to access mental health services through a criminal justice system gateway.

Other groups such as those older than 50 years and groups with comorbid disorders such as severe mental illness and drug or alcohol misuse, who are typically excluded from mainstream mental health services (Drake et al, 2000) are also a cause for concern(Drake & Mueser, 2000).

2.2 Current practice

The scope of this NICE guideline covers the mental health of adults in contact with the Criminal Justice System, apart from those whose sole contact is as witness or victim. It covers first contact with police service, whether or not an arrest is made through the courts and prison system and on release from prison to continuing community support (including contact with probation services. This involves a number of complex and interweaving pathways beginning with the 1.7 million people who may have some form of contact with the criminal justice system to the unknown number of people with a mental health problem who appear before courts in the UK each year, the 85,000 who are currently in prison and the 250,000 who are in the care of probation or community rehabilitation companies.(Ministry of Justice., 2013b). Given the complexities of the difficulties experienced by people with mental health problems in the criminal justice system it is troubling to learn that services for them are not well developed. Although a significant number of people coming into contact with criminal justice services may have a mental health problem and have had recent contact with services a surprising number are currently not in contact with services. For example, in a recent evaluation of the pilots of the Street Triage programme, Reveruzzi et al (2016) reported that although an average 60.6% of service users who came into contact with Street Triage were already known to mental health services, the average number of service users currently engaged with services was relatively low at 19.2%(Reveruzzi et al., 2016). In addition, recognition of mental health problems in prison settings is poor with many common mental, disorders going unrecognised and even where problems are recognised treatment is difficult to access or simply not available. There is evidence that in significant part these problems of access to treatment for people in contact with the criminal justice system are in part due a reluctance on the part of some health care professionals to offer services to people from the criminal justice system (Thornicroft et al., 2007).

For most people in contact with the criminal justice system health care comes from the primary care and secondary care health services which are accessed by all member so the general population. In the prison services the situation is different, across the whole prison estate there is access to a primary healthcare service akin to that of general practice in the community, supported to a greater or lesser extent by mental health services. Here the dominant model has been the mental health in-reach team (Steele et al, 2007) this is moving to a hybrid model of primary care and in-reach based services. Another important difference between prison and non-prison based services is the role played by prison staff, who in addition to maintaining safety and good order in the prison, are involved in providing important role in the recognition of an immediate management of mental health problems as they arise or are identified. Other prison service staff offender management staff, substance misuse teams, educationalists and forensic psychologists (particularly in relation to sexual offenders) also have a significant role in supporting people with a mental health problem. Of these staff groups only those working in primary care and specialist mental health teams are employed by the NHS. This, along with the complex nature of the mental health and physical problems experienced by prisoners, leads to a complex relationship between the prisoner and the National Health Service which can lead to significant problems with the delivery and coordination of care particularly when a person leaves prison. A particular problem arises with the management of patients in an acute psychotic episode which needs in-patient care where problems with access to beds leads to long delays the tensions that exist between those whose main concern is reduction of offending behaviour and the maintenance of safety and security and those whose main concern is the provision of healthcare.

Unfortunately, despite the fact that people in contact have the same rights of access to health care as the general population there is clear evidence that this is not the case. For some more seriously ill people particularly those who have been released from prison a number of mandated forms of treatment are available with the intention of ensuring people get access to effective treatment such as are contained in the Community Orders introduced as a sentencing option the Criminal Justice Act 2003. As an alternative to a custodial sentence, the Courts may impose mental health treatment orders or drug rehabilitation orders. Supervision of the delivery of this rests with an individual's probation worker and should an individual subject to community sentence or on licence following a prison sentence, breach requirements of those arrangements, they can either be returned to court or to prison. This also supports the use of community order focused on mental health treatment and drug and alcohol rehabilitation. For people who have committed more serious offences, after conviction, and determined by nature of offence and degree of risk, people in contact with the Criminal Justice System may be subject to multi-agency risk assessment conference (MARAC) or multi-agency public protection arrangements (MAPPA) processes, aimed at promoting effective inter-agency working.

Although the emphasis so far has been on problems of access to mental health services by people in contact the criminal justice system, loss of contact with mental health services, particularly for the more severely ill can lead to criminal justice services having a role in crisis response services. This can be seen for example in the development of Street Triage services which aim to identify people with mental health problems and either undertake or signpost appropriate care, as soon as possible after contact with the Criminal Justice System. A related function is that of liaison and diversion teams based in police cells and visiting courts who provide advice to the Criminal Justice System about care, management and processing of people in contact with it and act as a facilitating gateway into mental health and addiction services. There is no agreed model for street triage or liaison and diversion services and not all police services and courts have access to that support.

Outside of the prison services where there are established screening tools, case recognition and identification systems are limited and not all people who may benefit from an assessment by a forensic medical examiner, or a liaison and diversion practitioner, in police custody or a specialist team in a court diversion scheme will be identified and offered a further assessment. In police custody the fact many people may be intoxicated and lack so specialist police training may further hinder effective recognition of mental health problems. In prison setting lack of similar training for prison officers can again be an impediment to improved recognition. The consequences of this may be untreated disorder and inappropriate referral and use of both criminal justice and health care services. A particular concern are those people with neurodevelopmental disorders, learning disabilities and acquired cognitive impairment which often will do go undetected with significant consequences for the person who may be denied effective treatment (for example, methylphenidate for ADHD) and require additional and unnecessary input which prompt recognition and effective assessment and treatment could have avoided.

The Relationship between Offending and Mental Health Problems

The issue if the causal relationship between offending behaviour and mental illness has been the focus on much discussion. There is some evidence which suggest that certain disorders, particularly those managed in forensic settings are associated with different and higher rates of offending. For example, Coid et al (2015) in a review of patients discharged from medium secure units showed risks of all types of offending were increased for personality disorder, violence and acquisitive offences for delusional disorder, sexual offending for mania and hypomania and violence and acquisitive offending for organic brain syndrome(Coid et al., 2015). However, in a study including non-forensic populations Fazel and Yu (2011) identified an increased risk of re-offending with psychotic disorders when compared to the general

population but not when compared to other psychiatric disorders(Fazel & Yu, 2011). Yet other studies such as that by Stevens et al (2012) have suggested that offending behaviour may pre-date presentation to mental health services and that factors other than a mental disorder may be important in determining offending behaviour(Stevens et al., 2012). Factors such as homelessness may be associated with increased offending (Roy et al., 2014), the same study also reported that homeless severely mentally ill people were themselves more likely to be victims of crime, a finding supported by a study by Teplin et al (2005)(Teplin et al., 2005 Aug). Finally, it should be remembered that the data indicates that although some disorder may contribute an increased likelihood of offending effective treatment can reduce the likelihood of further offending (Pickard & Fazel, 2013).

The precise mechanisms which underpin the relationship between crime and mental illness are complex and varied an in many cases not well understood. It appears that pre-existing social factors, for example homelessness may be important and in other areas such as substance misuse, acquisitive crime may be driven by the need to buy illicit drugs and in some illnesses such as delusional disorder there may be a direct link to the mental disorder. For other disorders, the link may be less explicit, for example in neurodevelopmental disorders such as ADHD where impulsive adolescent males act recklessly and without consideration of consequences. Less obvious are the links between mood disorders, irritability, and secondary substance misuse. Links between mental health problems and offending behaviour relate to either the direct consequence of the disorder upon behaviour (disinhibition related to perhaps frontal lobe damage), underpinning social antecedents in common that predict both mental health problems and are associated with an increased risk of offending (adverse life experience), poor adaptive functioning, particular personality variables, and finally the consequence of offending and contact with the Criminal Justice System upon mental health. This last relationship is least well studied and may be predicated upon the social consequences of conviction (for example, job loss, relationship failure, social stigmatisation) rather than the traumatising nature of contact with the Criminal Justice per se although arrest, especially wrongful arrest, and imprisonment have been cited as traumatising experiences (Scott, 2010). There are many ethical and philosophical considerations that can be made about the relationship between offending and mental health problems.

The relationship between mental health problems and the criminal justice system and the understanding of the mechanisms underpinning the relationship has important consequences for the treatment and management of people with mental health problems in the criminal justice system and the interface between mental health services and the Criminal Justice System.

The first, most profound, and one that has the most implications is that of capacity. For adults, there is a presumption of capacity unless demonstrated otherwise. From the perspective of healthcare, an adult with capacity is one who can be a more or less equal partner in their treatment; they understand what course of action is being proposed; what the consequences of agreeing or not agreeing to it are; they can make an adequate assessment of this and they are able to communicate their decision. This principle is enshrined in clinical practice and in recent years has been underpinned, reinforced and standardised by the Mental Capacity Act. In the Criminal Justice System, issues around capacity are variously determined. The time when this is given much rigorous consideration is fitness to plead. Issues around fitness to plead are only raised in a minority of criminal court appearances and usually by the defence, sometimes by the court itself and sometimes by the prosecution. Fitness to plead is determined by a medical assessment of a person's ability to instruct council, understand the nature of the charges levelled against them, follow evidence, challenge jurors who they believe may be biased against them and understand the difference between a plea of guilty and not guilty. There is an interesting difference here between the approach to assessing fitness to plea which relies on external evidence and assessing capacity which is expected to be performed by every health professional should doubts about it arise. It is arguable that this is appropriate given the potential consequences for court

appearance, although a counter argument that the court is most expert in explaining the processes of the court and checking understanding, as opposed to this being done by external medical experts.

Elsewhere in the Criminal Justice System, individual workers are alert to potential problems around capacity and how it can effect engagement with the Criminal Justice System but processes are not as well defined or described. When taken into custody, the custody sergeant will consider whether someone is fit for detention and fit for interview but how this decision is reached is variable and may rely on a single healthcare practitioner stating that the person is fit to be detained or interviewed. The Police and Criminal Evidence Act requires the use of an appropriate adult is present After conviction, there is less routine consideration of whether someone has sufficient capacity to engage effectively with the Criminal Justice System and addressing these issues is very much dependent on individual practitioners. There are many instances in clinical practice of individuals with learning disability or other severe neurodevelopmental disorders who have been through the court system and imprisoned without any consideration of their ability to participate effectively in court proceedings, fitness to plead or capacity to engage effectively with the Criminal Justice process being considered explicitly.

The next issue concerns the detention of people with “serious mental illness” in prison and whether a prison can ever be a proper place to manage a person who continues to be significantly disabled by a severe mental illness, particularly if their symptoms and poorly controlled that the require prison od intensive care which is not possible to provide in a prison setting. Similar arguments can be made about dementia which is increasing as the prison population ages and presenting increasing management problems in the prison estate (Moll, 2013). The final issue is whether sexual offences against children are seen as a paraphilia, which is a mental disorder, currently the first approach is to see the problem as a criminal offence but then to offer treatment after conviction.

2.3 The Relationship between the Criminal Justice System and Mental Health Services

The interplay between two large publically funded systems both operating in a highly regulated and risk adverse environment is inevitably complex. There is enormous local variation (for example, Kosky and Hoyle, 2013) and for which only an overview can be provided here. People in contact with the Criminal Justice System who have or are suspected to have a mental health problem have access to the whole range of normal healthcare services unless they are held in prison. However, there is wide variation in the availability of specialist services, particularly those providing psychological treatments. Nevertheless, the basic building blocks of good mental health care – GP led services; community mental health teams, substance misuse services are routinely available. There are cultural and peculiar reasons why individuals may not engage with this offer, but the services themselves do exist. For those who are detained in prison, whether or remand or serving a sentence, it is a different story, one characterised by delay and under-resourcing (Forrester et al, 2013). Since 2003, the National Health Service has been responsible for the provision of care in prisons. Prior to this, responsibility lay with healthcare professionals directly employed by the Ministry of Justice. Reasons for transferring to care provided by the health service included a desire to establish equity of service provision, improved quality of care and to improve liaison and coordination with local mental health but it is not clear whether these benefits have actually been realised (Forrester et al, 2013).

2.4 Transitions between the Criminal Justice System and Mental Health Services

A central concern of those receiving and providing mental health care in the criminal justice system is the need to be able to successfully navigate the large number of transitions that can take place for someone with mental health problems in contact with the Criminal Justice System.

These transitions fall into several categories and grouping them loosely together they are: -

1. Transitions in geographical location. This particularly applies to people who are imprisoned, often at some distance from their normal place of residence, and may well be subject to several moves during their period of detention for a variety of reasons before being moved to a prison for resettlement, ideally near the place where they will be living. There then follows a further shift of location from prison to the community, perhaps after a period of some weeks, months or even years with a potential absence of established or healthy social networks to return to.
2. Transitions in healthcare provider. In an ideal situation there would be seamless transfer from the care of the General Practitioner, perhaps with the support of a community mental health team, to a custody liaison and diversion team, then should the person be imprisoned, to the prison mental health in-reach team and prison primary care services, with appropriate onward referral to services of other prisons should there be a move of prison and then release into the community with a coordinated handover of care to community services. Sadly, this is rarely the case although there are some transition points that are managed better than others.
3. Transitions in status. These are the subtlest, and often the hardest to quantify, but can have a profound effect on a person's opportunity to develop agency and demonstrate control of their life. The Criminal Justice System becomes involved when, essentially, the person's willingness or ability or choices to manage their life in a pro-social way fall short of societal norms. However, and perhaps for understandable reasons, contact with the Criminal Justice System as an offender is stigmatising and can lead to difficulty in navigating life's hurdles even after the "debt to society" has been repaid.

Problems of transition in these areas can occur for many reasons. People in contact with the Criminal Justice System are often suspicious of those they perceive to be authority figures, have a history of difficulty in establishing meaningful relationships with care providers, may have communication difficulties, may have profoundly complicated personal and medical histories; all of these conspire to make giving a reliable and complete history to medical professionals, especially upon a repeated basis, very difficult. In addition, there is often ignorance about the complexity of the Criminal Justice System and how to relate to it on the part of health professionals, an insufficiently considered approach to the management of confidentiality and the need to convey information to other agencies, a reluctance – especially for those health professionals in the community for whom contact with the Criminal Justice System is not a frequent occurrence – to deal with people with a history of offending, and a lack of appreciation of the complexity and multiple medical and social morbidities that people in contact with the Criminal Justice System demonstrate. This last factor is particularly so for disorders that an individual does not necessarily complain about directly, particularly neurodevelopmental disorders, cognitive impairment from a variety of causes, and continuing substance misuse. The most profound reasons for failure to manage transitions successfully, however, is problems with information flow. In part this is due to the aforementioned human factors but primarily because of the lack of a coherent information system among healthcare providers which is often compounded by partial or no access to the wealth of information held on Criminal Justice System databases and the legal, ethical and practical problems of getting those two systems and the people who operate them to communicate effectively with one another. There are particular problems around medicines

reconciliation at all points in a person's journey through the Criminal Justice System, and given the high level of psychoactive substances prescribed or used in this population, this is an area of particular concern.

Delivering effective treatment options in prison may also be limited by the restrictive nature of the prison environment and the fact that the Mental Health Act does not apply to the prison population (with the exception of sections 47 and 48 for the transfer of prisoners to and from hospital). Prisoners who would be sectioned if they were in the community would be transferred to NHS inpatient facilities. However, there are often long delays in transfers going ahead.

Rehabilitation and resettlement into the community is also complicated by the lifetime of social exclusion experienced by many prisoners. For example, 50% of sentenced prisoners are not registered with a GP before entering prison. There are also considerable difficulties in finding a GP willing to accept prisoners after release.

2.5 Economic Costs

Current healthcare provision, including mental healthcare, for people in contact with the criminal justice system is the responsibility of the NHS, with the exception of people under police custody and court custody. The care of people with mental health problems in contact with the criminal justice system impose a substantial burden on healthcare resources.

In England and Wales, the prison population was approximately 85,000 during the last months of 2015 (Ministry of Justice. & HM Prison Service., 2016) and there were 118,100 community resolutions given out in the 12 months ending June 2015 (Ministry of Justice., 2015a). All of the people in these groups have a very high risk of mental ill health. For example, 10% of men and 30% of women have had a previous psychiatric admission before they entered prison; 18% of prisoners were assessed as suffering from anxiety and depression (Ministry of Justice., 2015b) (MoJ, 2012); and 62% of male and 57% of female sentenced prisoners have a personality disorder (Prison Reform Trust., 2013).

It is estimated that £1.6 billion is spent annually on arresting, convicting, imprisoning and supervising people with identified mental health problems, rather than treating or supporting them (Revolving Doors Agency., 2007). In general, people with mental illness have a higher probability of having encounter with the criminal justice system. In the US Ascher-Svanum and colleagues (2010) assessed the prevalence of encounters with the criminal justice system and the estimated cost attributable to these encounters in the one-year treatment of persons with schizophrenia (Ascher-Svanum et al., 2010). Criminal justice system involvement was assessed using the service user survey. It was estimated that 278 (46%) of 609 participants reported at least one criminal justice system encounter. The mean annual per-service user cost of involvement was $1,429 per person, translating to 6% of total annual direct healthcare costs for those with involvement (11% when excluding crime victims) (in likely 2009 US dollars).

In another US study Petrila and colleagues (2010) examined the expenditures related to the criminal justice, health, mental health, and social welfare services over a 4-year period for arrestees with a serious mental illness (schizophrenia, schizoaffective disorder, delusional disorders, and other psychotic disorders and also bipolar I disorder, and major depressive or other bipolar and mood disorders) in a Florida county (Petrila et al., 2010). According to the analysis, the aggregate expenditures for the cohort were $95 million over the 4-year period, with a median per person expenditure of $15,134; in likely 2009 US dollars. Overall, as much as 39% of expenditures were associated with mental health services. Besides, individuals with mental illness remain incarcerated longer than inmates without mental illness charged with the same offences (McPherson, 2008), and upon release, re-arrest is common (Cox et al., 2001; Hartwell, 2003; Lamb et al., 2004). Robertson and colleagues (2015) examined the costs in people with mental health problems who have criminal justice involvement and those

that do not(Robertson et al., 2015). The authors reviewed administrative records from public behavioural health and criminal justice agencies of 25,133 adults with schizophrenia or bipolar disorder. It was found that costs were nearly 27% higher for those with justice involvement compared with those who had no justice involvement ($31,166 versus $24,602); in likely 2014 US dollars. Thus, people with serious mental illness who are in contact with the criminal justice sector cause considerable financial burden on public sector services.

Where mental illness is not recognised and is not treated properly there is a potential for repeat transitions between hospital admission, discharge, and readmission. People with mental illness in prison are frequently caught in a downward spiral of non-recovery. The costs of this are substantial and include transporting and reprocessing individuals who require varying levels of mental health treatment; personal costs to individuals and their families; added staff workload; and stressed and frustrated prison staff.

There seems to be a strong case for diverting offenders away from sentences in prison towards effective treatment in the community. There is an increased risk that vulnerable people's conditions are not being identified or treated, exacerbating mental health problems and frequently leading many to reoffend, self-harm or even commit suicide (Bradley Review 2009)(Bradley., 2009). The effective diversion requires some up-front investment in dedicated liaison and diversion teams working in police stations and courts. In the UK, a financial report commissioned for the Bradley review (2009) estimated that to implement an effective triage and assessment service, would cost between £3m and £9m nationally across all police forces; but there will be wider implications still on the potential impact on reducing recidivism. There is increasing evidence that well-designed interventions can reduce re-offending by 30% or more. The economic and social cost of crime committed by recently released prisoners serving short sentences amounts to £7-10 billion a year. Much of this cost falls directly on the victims of crime, but 20-30% is borne by the public sector, mainly the criminal justice system and the NHS. And the total lifetime cost of crime committed by an average offender following release from prison is of the order of £250,000 (Centre for Mental Health. et al., 2010). In another exploratory analysis conducted by the Centre for the Mental Health (2009) it was estimated that the combined costs of diversion and liaison schemes in the UK is around £10 million a year(Centre for Mental Health., 2009). The authors argued that there is good evidence that offenders with mental health problems are more likely to be held on remand than other offenders and each additional case held on remand imposes, on average, additional costs of £3,000 on the criminal justice system.

Another issue is the need for services to support community re-entry following incarceration and continuity of care initiated in prisons. Most of the evidence in this area is from the US and Australia. In the US Lin and colleagues (2015) developed an economic model to estimate the cost burden of psychiatric relapse and recidivism among service users with schizophrenia recently released from incarceration from a US state government perspective(Lin et al., 2015). Among 34,500 persons released from incarceration in the state of Florida annually, 5,307 were estimated to have schizophrenia. The cumulative 3-year costs to the state government were $21,146,000 and $25,616,000 for criminal justice and psychiatric hospitalisation costs, respectively ($3,984 per service user criminal justice costs; $4,827 per service user hospitalisation costs); in likely 2014 US dollars.

In another study, Alan and colleagues (2011) examined the resource use in ex-prisoners within the first 12 months of release from prison in Western Australia (Alan et al., 2011). It was found that one in five adults released from prisons between 2000 and 2002 were hospitalised in the 12 months that followed, which translated into 12,074 inpatient bed days and associated costs of $10.4 million. Mental health disorders such as schizophrenia and depression and injuries involving the head or face and/or fractures accounted for as much as 58.9% of all bed days. Ostermann and colleagues (2013) estimated the costs of crimes committed by reintegrated former inmates with mental illness and compared these costs to those without mental illness (Ostermann & Matejkowski, 2013). It was found that that the recidivism costs of those with mental illness over the course of 3 years of follow-up are

nearly 3 times as large as for former inmates without mental illness. This indicates the importance of treatment during the prison stay and the need for services to support community re-entry in reduction of health service costs.

Similarly, substance abuse is associated with great economic costs in this population. McKenzie and colleagues (2005) reported costs associated with opiate replacement therapy at time of release from incarceration in the US. The authors reported the annual cost of methadone replacement therapy to be approximately $4,420 per person. In another, study Werb and colleagues (2007) reported costs associated with drug treatment courts in Canada(Werb et al., 2007). The authors reported the cost per person to be $21,265 for Vancouver drug court programme participants and $13,117 for matched controls. They further went on to report the total costs of the Vancouver drug court programme during the period of 2001 and 2005 to be £4.1 million. With 42 participants who either graduated or completed the programme, the cost per graduates or completer was as high as $96,639. Bechelli and colleagues (2014) reported that in Washington State the average per client cost of substance abuse treatment for the period 1998–2007 was $6,504 (Bechelli et al., 2014).

All of the above indicates that the management of people with mental health problems who are in contact with the criminal justice system cause a substantial financial burden on the NHS, criminal justice sector and the wider public sector. Individuals with mental health problems who are in prisons are less likely to adjust to the prison life; they are vulnerable to repeat hospitalisations; and have a higher risk of future crime associated with the untreated mental illness. There is a need for UK-based evidence to better understand the interface between the mental health services and the criminal justice systems and the related economic costs; and economic evaluations to identify cost-effective treatment strategies and service configurations for this population.

3 Methods used to develop this guideline

3.1 Overview

The development of this guideline followed The Guidelines Manual (NICE, 2012). A team of health and social care professionals, lay representatives and technical experts known as the Guideline Committee (GC), with support from the NGA staff, undertook the development of a person-centred, evidence-based guideline. There are 7 basic steps in the process of developing a guideline:

1. Define the scope, which lays out exactly what will be included (and excluded) in the guidance.
2. Define review questions that cover all areas specified in the scope.
3. Develop a review protocol for each systematic review, specifying the search strategy and method of evidence synthesis for each review question.
4. Synthesise data retrieved, guided by the review protocols.
5. Produce evidence profiles and summaries using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system.
6. Consider the implications of the research findings for clinical practice and reach consensus decisions on areas where evidence is not found.
7. Answer review questions with evidence-based recommendations for clinical practice.

The clinical practice recommendations made by the GC are therefore derived from the most up-to-date and robust evidence for the clinical and cost effectiveness of the interventions and services covered in the scope. Where evidence was not found or was inconclusive, the GC adopted both formal and informal methods to reach consensus on what should be recommended, factoring in any relevant issues. In addition, to ensure a service user and carer focus, the concerns of service users and carers regarding health and social care have been highlighted and addressed by recommendations agreed by the whole GC.

3.2 The scope

Topics are referred by NHS England and the letter of referral defines the remit, which defines the main areas to be covered. The NGA developed a scope for the guideline based on the remit (see Appendix A). The purpose of the scope is to:

- provide an overview of what the guideline will include and exclude
- identify the key aspects of care that must be included
- set the boundaries of the development work and provide a clear framework to enable work to stay within the priorities agreed by NICE and the National Guideline Alliance, and the remit from the Department of Health.
- inform the development of the review questions and search strategy
- inform professionals and the public about expected content of the guideline
- keep the guideline to a reasonable size to ensure that its development can be carried out within the allocated period.

An initial draft of the scope was sent to registered stakeholders who had agreed to attend a scoping workshop. The workshop was used to:

- obtain feedback on the selected key clinical issues
- identify which population subgroups should be specified (if any)
- seek views on the composition of the GC
- encourage applications for GC membership.

The draft scope was subject to consultation with registered stakeholders over a 4-week period. During the consultation period, the scope was posted on the NICE website (www.nice.org.uk). Comments were invited from stakeholder organisations The NGA and NICE reviewed the scope in light of comments received, and the revised scope was signed off by NICE.

3.3 The Guideline committee

During the consultation phase, members of the GC were appointed by an open recruitment process. GC membership consisted of: professionals in psychiatry, clinical psychology, nursing, social work, speech and language therapy, and general practice; academic experts in psychiatry and psychology; commissioning managers; and carers and representatives from service user and carer organisations. The guideline development process was supported by staff from the NGA, who undertook the clinical and health economic literature searches, reviewed and presented the evidence to the GC, managed the process, and contributed to drafting the guideline.

3.3.1 Guideline Development Group meetings

Twelve GC meetings were held between January 2015 and July 2016. During each day-long GC meeting, in a plenary session, review questions and clinical and economic evidence were reviewed and assessed, and recommendations formulated. At each meeting, all GC members declared any potential conflicts of interest (see Appendix B), and service user and carer concerns were routinely discussed as a standing agenda item.

3.3.2 Service users and carers

Individuals with direct experience of services gave an integral service-user focus to the GC and the guideline. The GC included carers and a representative of a national service user group. They contributed as full GC members to writing the review questions, providing advice on outcomes most relevant to service users and carers, helping to ensure that the evidence addressed their views and preferences, highlighting sensitive issues and terminology relevant to the guideline, and bringing service user research to the attention of the GC. In drafting the guideline, they met with the NGA team on several occasions to develop the chapter on experience of care and they contributed to writing the guideline's introduction and identified recommendations from the service user and carer perspective.

3.3.3 Expert advisers

Expert advisers, who had specific expertise in one or more aspects of treatment and management relevant to the guideline, assisted the GC, commenting on specific aspects of the developing guideline and making presentations to the GC. Appendix C lists those who agreed to act as expert advisers.

3.3.4 National and international experts

National and international experts in the area under review were identified through the literature search and through the experience of the GC members. These experts were contacted to identify unpublished or soon-to-be published studies, to ensure that up-to-date evidence was included in the development of the guideline. They informed the GC about completed trials at the pre-publication stage, systematic reviews in the process of being published, studies relating to the cost effectiveness of treatment and trial data if the GC could be provided with full access to the complete trial report. Appendix E lists researchers who were contacted.

3.4 Review protocols

Review questions drafted during the scoping phase were discussed by the GC at the first few meetings and amended as necessary. The review questions were used as the starting point for developing review protocols for each systematic review (described in more detail below). Where appropriate, the review questions were refined once the evidence had been searched and, where necessary, sub-questions were generated. The final list of review questions can be found in Appendix F.

For questions about interventions, the PICO (Population, Intervention, Comparison and Outcome) framework was used to structure each question (see Table 2:).

Table 2: Features of a well-formulated question on the effectiveness of an intervention – PICO

Population:	Which population of service users are we interested in? How can they be best described? Are there subgroups that need to be considered?
Intervention:	Which intervention, treatment or approach should be used?
Comparison:	What is/are the main alternative/s to compare with the intervention?
Outcome:	What is really important for the service user? Which outcomes should be considered: intermediate or short-term measures; mortality; morbidity and treatment complications; rates of relapse; late morbidity and readmission; return to work, physical and social functioning and other measures such as quality of life; general health status?

Questions relating to case identification and assessment tools and methods do not involve an intervention designed to treat a particular condition, and therefore the PICO framework was not used. Rather, the questions were designed to pick up key issues specifically relevant to clinical utility, for example their accuracy, reliability, safety and acceptability to the service user.

In some situations, the prognosis of a particular condition is of fundamental importance, over and above its general significance in relation to specific interventions. Areas where this is particularly likely to occur relate to assessment of risk, for example in terms of behaviour modification or screening and early intervention. In addition, review questions related to issues of service delivery are occasionally specified in the remit from the Department of Health/Welsh Assembly Government. In these cases, appropriate review questions were developed to be clear and concise.

Where review questions about service user experience were specified in the scope, the SPICE format was used to structure the questions (Table 3).

Table 3: Features of a well-formulated question about the experience of care (qualitative evidence) – SPICE

Setting	In what environment? In what context?
Perspective	For who?
Intervention (phenomenon of interest):	Which intervention/interest should be included?
Comparison:	What?
Evaluation:	How well? What result?
Adapted from (Booth, 2003)	

For each topic, addressed by one or more review questions, a review protocol was drafted by the technical team using a standardised template (based on PROSPERO[a]), reviewed and agreed by the GC (all protocols are included in Appendix F).

To help facilitate the literature review, a note was made of the best study design type to answer each question. There are five main types of review question of relevance to NICE guidelines. These are listed in Table 4. For each type of question, the best primary study design varies, where 'best' is interpreted as 'least likely to give misleading answers to the question'. For questions about the effectiveness of interventions, where randomised controlled trials (RCTs) were not available, the review of other types of evidence was pursued only if there was reason to believe that it would help the GC to formulate a recommendation.

However, in all cases, a well-conducted systematic review (of the appropriate type of study) is likely to always yield a better answer than a single study.

Table 4: Best study design to answer each type of question

Type of question	Best primary study design
Effectiveness or other impact of an intervention	Randomised controlled trial (RCT); other studies that may be considered in the absence of RCTs are the following: internally/externally controlled before and after trial, interrupted time-series, cohort study
Diagnostic accuracy (for example diagnostic test or prediction rule)	Cross sectional study, RCT for test and treat questions
Prognostic factors	Prospective cohort
Rates (incidence and prevalence of disease, service user experience, rare side effects)	Prospective cohort, registry, cross-sectional study
Experience of care	Qualitative research (for example, grounded theory, ethnographic research)

3.5 Clinical review methods

The aim of the clinical literature review was to systematically identify and synthesise relevant evidence from the literature in order to answer the specific review questions developed by the GC. Thus, clinical practice recommendations were evidence-based, where possible, and, if evidence was not available, informal consensus methods were used to try and reach general agreement between GC members (see Section 3.8.3) and the need for future research was specified.

3.6 The search process

3.6.1 Scoping searches

A broad preliminary search of the literature was undertaken in July 2014 to obtain an overview of the issues likely to be covered by the scope, and to help define key areas. The searches were restricted to clinical guidelines, Health Technology Assessment (HTA) reports, key systematic reviews and RCTs. A list of databases and websites searched can be found in Appendix H.

[a] http://www.crd.york.ac.uk/prospero/

3.6.2 Systematic literature searches

After the scope was finalised, a systematic search strategy was developed to locate as much relevant evidence as possible. The balance between sensitivity (the power to identify all studies on a particular topic) and specificity (the ability to exclude irrelevant studies from the results) was carefully considered, and a decision made to utilise a broad approach to searching to maximise retrieval of evidence to all parts of the guideline. Searches were restricted to certain study designs if specified in the review protocol, and conducted in the following databases:

- Cochrane Database of Abstracts of Reviews of Effects (DARE)
- Cochrane Database of Systematic Reviews (CDSR)
- CENTRAL
- Embase
- HTA database (technology assessments)
- MEDLINE/MEDLINE In-Process
- Psychological Information Database (PsycINFO)

The search strategies were initially developed for MEDLINE before being translated for use in other databases/interfaces. Strategies were built up through a number of trial searches and discussions of the results of the searches with the review team and GC to ensure that all possible relevant search terms were covered. In order to assure comprehensive coverage, search terms for mental health and the criminal justice system were kept purposely broad to help counter dissimilarities in database indexing practices and thesaurus terms, and imprecise reporting of study populations by authors in the titles and abstracts of records. The search terms for each search are set out in full in Appendix H.

3.6.3 Reference Management

Citations from each search were downloaded into reference management software and duplicates removed. Records were then screened against the eligibility criteria of the reviews before being appraised for methodological quality (see below). The unfiltered search results were saved and retained for future potential re-analysis to help keep the process both replicable and transparent.

3.6.4 Search filters

To aid retrieval of relevant and sound studies, filters were used to limit a number of searches to systematic reviews, RCTs and qualitative studies. The search filters for systematic reviews and RCTs are adaptations of validated filters designed by the Health Information Research Unit (HIRU) at McMaster University. The qualitative research filter was developed in-house. Each filter comprises index terms relating to the study type(s) and associated text words for the methodological description of the design(s). The filters have been recorded and can be found listed in the search strategies in Appendix H.

3.6.5 Date and language restrictions

Systematic database searches were initially conducted in February 2015 up to the most recent searchable date. Search updates were generated on a 6-monthly basis, with the final re-runs carried out in June 2016 ahead of the guideline consultation. After this point, studies were only included if they were judged by the GC to be exceptional (for example, if the evidence was likely to change a recommendation).

Although no language restrictions were applied at the searching stage, foreign language papers were not requested or reviewed, unless they were of particular importance to a review question.

Date restrictions were not applied, except for searches of systematic reviews which were limited to research published from 2000. The search for systematic reviews was restricted to the last 15 years as older reviews were thought to be less useful.

3.6.6 Other search methods

Other search methods involved: (a) scanning the reference lists of all eligible publications (systematic reviews and stakeholder evidence) for more published reports and citations of unpublished research; (b) tracking key papers in the Science Citation Index (prospectively) over time for further useful references; (c) conducting searches in ClinicalTrials.gov for unpublished trial reports; (d) contacting included study authors for unpublished or incomplete datasets. Searches conducted for existing NICE guidelines were updated where necessary. Other relevant guidelines were assessed for quality using the AGREE instrument (AGREE Collaboration., 2003). The evidence base underlying high-quality existing guidelines was utilised and updated as appropriate.

Full details of the search strategies and filters used for the systematic review of clinical evidence are provided in Appendix H.

3.6.7 Study selection and assessment of methodological quality

All primary-level studies included after the first scan of citations were acquired in full and re-evaluated for eligibility at the time they were being entered into the study information database (standardised template created in Microsoft Excel). More specific eligibility criteria were developed for each review question and are described in the relevant clinical evidence chapters. Eligible systematic reviews were critically appraised for methodological quality (risk of bias) using a checklist (see *The Guidelines Manual* (NICE, 2012) for template). Primary intervention studies were appraised using a checklist based on the Cochrane Risk of Bias tool, but with additional items for non-randomised studies (e.g. non-random allocation method and confounders) and for indirectness and imprecision (see Appendices I, J and K).

However, some checklists recommended in the 2014 manual update (NICE., 2014) were also used (for example, for qualitative studies [The Critical Appraisal Skills Programme, CASP, (2013) checklist], for effectiveness of intervention/service delivery studies [appropriate NICE quality assessment checklist]. The eligibility of each study was confirmed by at least 1 member of the GC.

The Quality Assessment of Diagnostic Accuracy Studies – Revised (QUADAS-II) (Whiting, 2011) was used for diagnostic studies and was adapted for use with risk assessment studies as follows:

- Index test question signalling question: ‘If a threshold was used, was it pre-specified?’ This was amended to: ‘Is information available to facilitate clinical judgment?’ (that is, how scores should be translated to risk level)
- Flow and timing signalling question: ‘Was there an appropriate interval between index test(s) and reference standard?’ This was interpreted as: ‘Was there sufficient time for events of interest to occur?’

The CASP clinical prediction rule checklist suggested in the in the 2014 manual update (NICE., 2014) covers similar risk of bias domains as QUADAS II, but the CASP tool does not explicitly cover whether there is sufficient follow up time for events to occur in the study. For this reason QUADAS-II was used and modified to capture this specific aspect.

The eligibility of studies was confirmed by the GC. A flow diagram of the search process for selection of studies for inclusion in the literature review conducted for this guideline is provided in Appendix O.

For some review questions, it was necessary to prioritise the evidence with respect to the UK context (that is, external validity). To make this process explicit, the GC took into account the following factors when assessing the evidence:

- participant factors (for example, gender, age and ethnicity)
- provider factors (for example, model fidelity, the conditions under which the intervention was performed and the availability of experienced staff to undertake the procedure)
- cultural factors (for example, differences in standard care and differences in the welfare system).

It was the responsibility of the GC to decide which prioritisation factors were relevant to each review question in light of the UK context.

3.6.18 Double-Sifting

Titles and abstracts of identified studies were screened by two reviewers against inclusion criteria specified in the protocols, until a good inter-rater reliability was observed (percentage agreement ≥90% or Kappa statistics, K>0.60). Any disagreements between raters were resolved through discussion. Initially 10% of references were double-screened. If inter-rater agreement was good, then the remaining references were screened by one reviewer.

3.6.19 Unpublished evidence

Stakeholders were invited to submit any relevant unpublished data using the call for evidence process set out in the NICE manual (NICE, 2012). Additionally, authors and principal investigators were approached for unpublished evidence. The GC used a number of criteria when deciding whether or not to accept unpublished data. First, the evidence must have been accompanied by a trial report containing sufficient detail to properly assess risk of bias. Second, the evidence must have been submitted with the understanding that data from the study and a summary of the study's characteristics would be published in the full guideline. Therefore, in most circumstances the GC did not accept evidence submitted 'in confidence'. However, the GC recognised that unpublished evidence submitted by investigators might later be retracted by those investigators if the inclusion of such data would jeopardise publication of their research.

3.6.20 Experience of care

Reviews were sought of qualitative studies that used relevant first-hand experiences of service users and their families, partners or carers. A particular outcome was not specified by the GC. Instead, the review was concerned with narrative data that highlighted the experience of care.

3.7 Data extraction

3.7.1 Quantitative analysis

Study characteristics, aspects of methodological quality, and outcome data were extracted from all eligible studies, using Review Manager Version 5.3.5 (Cochrane Collaboration, 2014) and an Excel-based form (see Appendix L).

In most circumstances, for a given outcome (continuous and dichotomous), where more than 50% of the number randomised to any group were missing or incomplete, the study results were excluded from the analysis (except for the outcome 'leaving the study early', in which case, the denominator was the number randomised). Where there were limited data for a particular review, the 50% rule was not applied. In these circumstances the evidence was downgraded (see section 3.7.4).

In some circumstances it was not possible to extract any efficacy data for the interventions and outcomes of interest and in such cases the study was excluded from the analysis.

Where possible, outcome data from an intention-to-treat analysis (ITT) (that is, a 'once-randomised-always-analyse' basis) were used. Where ITT had not been used or there were missing data, the effect size for dichotomous outcomes were recalculated using worse-case scenarios (for instance, if the outcome of missing participants was positive, it was assumed that they did not have the positive result). Where conclusions varied between scenarios, the evidence was downgraded (see section 3.7.4).

Where some of the studies failed to report standard deviations (for a continuous outcome), and where an estimate of the variance could not be computed from other reported data or obtained from the study author, the following approach was taken.[b] When the number of studies with missing standard deviations was less than one-third and when the total number of studies was at least 10, the pooled standard deviation was imputed (calculated from all the other studies in the same meta-analysis that used the same version of the outcome measure). In this case, the appropriateness of the imputation was made by comparing the standardised mean differences (SMDs) of those trials that had reported standard deviations against the hypothetical SMDs of the same trials based on the imputed standard deviations. If they converged, the meta-analytical results were considered to be reliable.

When the conditions above could not be met, standard deviations were taken from another related systematic review (if available). In this case, the results were considered to be less reliable and the evidence downgraded.

For continuous outcomes, final scores in each group were the preferred outcome for extraction. However, if final or change scores (from baseline) were not reported for each group in a study (for example, the study reported an F-value, p-value or t-value), the SMD was estimated, if possible, using statistical calculator.

The meta-analysis of survival data, such as time to any mood episode, was based on log hazard ratios and standard errors. Since individual participant data were not available in included studies, hazard ratios and standard errors calculated from a Cox proportional hazard model were extracted. Where necessary, standard errors were calculated from confidence intervals (CIs) or *p* value according to standard formulae (see the Cochrane Reviewers' Handbook 5.1.0 (Higgins & Green, 2011)). Data were summarised using the generic inverse variance method using Review Manager.

Data from studies included in existing systematic reviews were extracted independently by 1 reviewer and cross-checked with the existing dataset. Where possible, two independent reviewers extracted data from new studies. Where double data extraction was not possible, data extracted by one reviewer was checked by the second reviewer. Disagreements were resolved through discussion. Where consensus could not be reached, a third reviewer or GC members resolved the disagreement. Masked assessment (that is, blind to the journal from which the article comes, the authors, the institution and the magnitude of the effect) was not used since it is unclear that doing so reduces bias (Berlin, 2001; Jadad et al., 1996).

3.7.12 Qualitative analysis

After transcripts/reviews or primary studies of service user experience were identified, each was read and re-read and sections of the text were collected under different headings using an Excel-based form. Initially the text from the transcripts/reviews was organised using a matrix of service user experience (see Table 5)

[b] Based on the approach suggested by Furukawa and colleagues (2006).

The matrix was formed by creating a table with the eight dimensions of patient-centred care developed by the Picker Institute Europe[c], down the vertical axis, and the key points on a pathway of care (as specified by the GC), across the horizontal axis. With regard to terminology, the GC preferred the term 'person-centred' rather than 'patient-centred', therefore the former is used in the matrix. The Picker Institute's dimensions of patient-centred care were chosen because they are well established, comprehensive, and based on research. In addition, a variation of these dimensions has been adopted by the US Institute of Medicine (Institute of Medicine., 2001).

Table 5: Matrix of service user experience

		Key points on the pathway of care		**Themes that apply to all points on the pathway**
Experience of the mental health problem				
The relationship between individual service users and professionals	Involvement in decisions and respect for preferences			
	Clear, comprehensible information and support for self-care			
	Emotional support, empathy and respect			
The way that services and systems work	Fast access to reliable health advice			
	Effective treatment delivered by trusted professionals			
	Attention to physical and environmental needs			
	Involvement of, and support for, family and carers			
	Continuity of care and smooth transitions			

Under the broad headings in the matrix, specific emergent themes were identified and coded by two researchers working independently. Then, a sample of each other's work (10%) for reliability. Discrepancies or difficulties with the interpretation of study results were resolved through discussion between reviewers or with members of the GC. Overlapping themes and themes with the highest frequency count across all testimonies were extracted and

[c] http://www.pickereurope.org/patientcentred

regrouped using the matrix. The findings from the qualitative analysis can be found in Appendix J.

3.7.3 Evidence synthesis

The method used to synthesise evidence depended on the review question and availability and type of evidence (see Appendix F for full details). Briefly, for questions about the psychometric properties of instruments, reliability, validity and clinical utility were synthesised narratively based on accepted criteria. For questions about test accuracy, bivariate test accuracy meta-analysis was conducted where appropriate. For questions about the effectiveness of interventions, standard meta-analysis was used where appropriate, otherwise narrative methods were used with clinical advice from the GC. In the absence of high-quality research, formal and informal consensus processes were used (see 3.8.3).

3.7.4 Grading the quality of evidence

For questions about the effectiveness of interventions and the organisation and delivery of care, the GRADE approach[d] was used to grade the quality of evidence from group comparisons for each outcome (Guyatt et al., 2011). Evidence from systematic reviews of Small Case and Small-N (SCSn) designs was graded as 'low' or 'very low' quality without using the formal GRADE approach because specific methodology has not been developed to grade this type of evidence (see section 3.7.2 for limitations, which account for the low or very low-quality grade). For questions about the experience of care and the organisation and delivery of care, methodology checklists (see section 3.6.7) were used to assess the risk of bias, and this information was taken into account when interpreting the evidence. The technical team produced modified GRADE evidence profiles (see below) using GRADEpro guideline development tool (GRADEpro) software (Version 3.6), following advice set out in the GRADE handbook (Schünemann et al., 2009). All staff doing GRADE ratings were trained, and calibration exercises were used to improve reliability (Mustafa et al., 2013).

For questions about diagnostic accuracy, while the QUADAS framework does not provide an overall quality index for each study, this was deemed important to assist interpretation of the data tools to augment assessment of mental health problems. We adopted the terminology used within GRADE (high, moderate, low or very low quality evidence). For each of the first 3 domains (patient selection, index text, reference standard) we used the 'risk of bias' and 'concerns about applicability' ratings (low, unclear and high risk for each) to create a 3x3 table (see Table 6). For domain 4 (flow and timing), which has only a 'risk of bias' rating, the same method was used, but 'risk of bias' was entered on both axes. We then used the 4 total domain ratings to generate an overall quality index. For the overall quality rating we took the mode classification and upgraded or downgraded from that point; that is, if a study had 2 ratings of 'high', one of 'moderate' and one of 'very low', then the final quality rating would be 'moderate'. Although there is overlap between the concepts of indirectness in GRADE and applicability in QUADAS we did not explicitly downgrade for indirectness or imprecision.

Table 6: Process for determining overall quality ratings for QUADAS-II domains 1-3 (patient selection', index test and reference standard)

	Concerns about applicability			
Risk of bias		Low risk	Unclear risk	High risk
	Low risk	High	Moderate	Moderate
	Unclear risk	Moderate	Low	Low
	High risk	Moderate	Low	Very low
Note. QUADAS = Quality Assessment of Diagnostic Accuracy Studies.				

[d] For further information about GRADE, see www.gradeworkinggroup.org

3.7.15 Evidence profiles

For questions about the effectiveness of interventions and the organisation and delivery of care a GRADE evidence profile was used to summarise both the quality of the evidence and the results of the evidence synthesis for each 'critical' and 'important' outcome (see Appendix N for completed evidence profiles). The GRADE approach is based on a sequential assessment of the quality of evidence, followed by judgment about the balance between desirable and undesirable effects, and subsequent decision about the strength of a recommendation (Table 7).

Within the GRADE approach to grading the quality of evidence, the following is used as a starting point:

- RCTs without important limitations provide high-quality evidence
- observational studies without special strengths or important limitations provide low-quality evidence.

For each outcome, quality may be reduced depending on 5 factors: limitations, inconsistency, indirectness, imprecision and publication bias. For the purposes of the guideline, each factor was evaluated using criteria provided in Table 8.

For observational studies without any reasons for down-grading, the quality may be up-graded if there is a large effect, all plausible confounding would reduce the demonstrated effect (or increase the effect if no effect was observed), or there is evidence of a dose-response gradient (details would be provided under the 'other' column).

Each evidence profile includes a summary of findings: number of participants included in each group, an estimate of the magnitude of the effect, and the overall quality of the evidence for each outcome. Under the GRADE approach, the overall quality for each outcome is categorised into 1 of 4 groups (high, moderate, low, very low).

Table 7: Example of a GRADE evidence profile

Quality assessment							No. of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Intervention**	**Control group**	**Relative (95% CI)**	**Absolute**		
Outcome 1 (measured with: any valid method; better indicated by lower values)												
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious[1]	None	47	43	-	SMD 0.20 lower (0.61 lower to 0.21 higher)	⊕⊕⊕⊖ MODERATE	CRITICAL
Outcome 2 (measured with: any valid rating scale; better indicated by lower values)												
4	Randomised trials	Serious[2]	No serious inconsistency	No serious indirectness	Serious[1]	None	109	112	-	SMD 0.42 lower (0.69 to 0.16 lower)	⊕⊕⊖⊖ LOW	CRITICAL
Outcome 3 (measured with: any valid rating scale; better indicated by lower values)												
26	Randomised trials	No serious risk of bias	Serious[3]	No serious indirectness	No serious imprecision	None	521/5597 (9.3%)	798/3339 (23.9%)	RR 0.43 (0.36 to 0.51)	136 fewer per 1000 (from 117 fewer to 153 fewer)	⊕⊕⊕⊖ MODERATE	CRITICAL
Outcome 4 (measured with: any valid rating scale; better indicated by lower values)												
5	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	503	485	-	SMD 0.34 lower (0.67 to 0.01 lower)	⊕⊕⊕⊕ HIGH	CRITICAL

Note.
[1] OIS (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[2] Risk of bias across domains was generally high or unclear.
[3] There is evidence of moderate heterogeneity of study effect sizes.
CI = confidence interval; OIS = optimal information size; RR = risk ratio; SMD = standardised mean difference.

Table 8: Factors that decrease the quality of evidence

Factor	Description	Criteria
Limitations	Methodological quality/ risk of bias.	Serious risks across most studies (that reported a particular outcome). The evaluation of risk of bias was made for each study using NICE methodology checklists (see Section 3.6).
Inconsistency	Unexplained heterogeneity of results.	Moderate or greater heterogeneity (using the methods suggested by GRADE[1])
Indirectness	How closely the outcome measures, interventions and participants match those of interest.	If the comparison was indirect, or if the available evidence was substantially different from the population, intervention, comparator, or an outcome specified in the protocol for the question being addressed by the GC.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect.	If either of the following 2 situations were met: • the OIS (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) was not achieved • the 95% confidence interval around the pooled or best estimate of effect included both (a) no effect and (b) appreciable benefit or appreciable harm (using default minimally important differences, as suggested by GRADE)
Publication bias	Systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.	Evidence of selective publication. This may be detected during the search for evidence, or through statistical analysis of the available evidence.

Note.
[1] An I^2 of 50% was used as the cut-off to downgrade for inconsistency. If heterogeneity was found, subgroup analysis was performed using the pre-specified subgroups in the protocol (see Appendix F); if subgroup analysis did not explain the heterogeneity, a random-effects model was used and the outcome was downgraded.
GC = Guideline Committee; GRADE = Grading of Recommendations Assessment, Development and Evaluation; NICE = National Institute for Health and Care Excellence; OIS = optimal information size.

3.8 Presenting evidence to the Guideline Committee

Study characteristics tables and, where appropriate, forest plots generated with Review Manager Version 5.3 and GRADE summary of findings tables (see below) were presented to the GC.

Where meta-analysis was not appropriate and/ or possible, the reported results from each primary-level study were reported in the study characteristics table and presented to the GC. The range of effect estimates were included in the GRADE profile, and where appropriate, described narratively.

The GC were also provided with evidence statements reflecting the key findings, the quantity, quality and consistency of the evidence. Evidence statements were prioritised for the critical outcomes, especially when the evidence contained multiple correlated outcomes.

3.8.1 Summary of findings tables

Summary of findings tables generated from GRADEpro were used to summarise the evidence for each outcome and the quality of that evidence.

For continuous outcomes the mean difference (MD) was generally used as the effect estimate. In some cases, for example when studies used different scales to measure the same outcome, the standardized mean difference (SMD) was used. Minimally important difference (MID) boundaries were determined as +/- 0.5 times the SD of the control arm. When SMD was used the MID was +/- 0.5 on the SMD scale. The MID boundaries considered for relative risk (RR) were 0.8 to 1.25. If the 95% confidence interval of the effect estimate spanned both the upper or lower MID threshold and no effect (0 for MD or SMD; 1 for RR) then the effect estimate was considered very imprecise. If the MD, SMD or RR was below the lower MID threshold or above the upper MID threshold, the effect was considered as potentially clinically important. Where the GC felt that effects were of sufficient magnitude to be clinically important, this is described within the Linking Evidence to Recommendations (LETR) tables.

For questions about diagnostic tests the GC considered a test as potentially clinically useful if both sensitivity and specificity were 75% or greater.

Table 9 is an example of a GRADE summary of findings table. The summary of findings tables provide anticipated comparative risks, which are especially useful when the baseline risk varies for different groups within the population.

For continuous outcomes the mean difference (MD) was generally used as the effect estimate. In some cases, for example when studies used different scales to measure the same outcome, the standardized mean difference (SMD) was used. Minimally important difference (MID) boundaries were determined as +/- 0.5 times the SD of the control arm. When SMD was used the MID was +/- 0.5 on the SMD scale. The MID boundaries considered for relative risk (RR) were 0.8 to 1.25. If the 95% confidence interval of the effect estimate spanned both the upper or lower MID threshold and no effect (0 for MD or SMD; 1 for RR) then the effect estimate was considered very imprecise. If the MD, SMD or RR was below the lower MID threshold or above the upper MID threshold, the effect was considered as potentially clinically important. Where the GC felt that effects were of sufficient magnitude to be clinically important, this is described within the Linking Evidence to Recommendations (LETR) tables.

For questions about diagnostic tests the GC considered a test as potentially clinically useful if both sensitivity and specificity were 75% or greater.

Table 9: Example of a GRADE summary of findings table

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with PLB	Risk difference with intervention (95% CI)
Global impression: no improvement - short term	102 (1 study)	⨁⨁⊖⊖ LOW[1,2]	RR 0.89 (0.69 to 1.16)	725 per 1000	80 fewer per 1000 (from 225 fewer to 116 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with PLB	Risk difference with intervention (95% CI)
Behaviour: average change score (ABS) - medium term	101 (1 study)	⨁⨁⨀⨀ LOW[1,2]	-	-	The mean behaviour: 1. average change score (ABS =) - medium term, in the intervention groups was 0.60 standard deviations lower (1 to 0.21 lower)
Adverse effects: extrapyramidal symptoms - medium term	243 (2 studies)	⨁⨁⨀⨀ LOW[1,2]	RR 0.34 (0.05 to 2.1)	33 per 1000	21 fewer per 1000 (from 31 fewer to 36 more)

1 Generally unclear risk of bias and funded by manufacturer.

2 OIS (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

Table 10 is an aid to the interpretation of the psychometric scales used as outcomes in some of the summary of findings tables.

Table 10: Range and direction of psychometric scales used in summary of findings tables

Scale	Range	Direction
Abel and Becker Cognition Scale [ABCS]	26 to 130	Higher better
Addiction Severity Index (ASI-6): alcohol composite score	0 to 9	Lower better
Addiction Severity Index (ASI-6): drug composite score	0 to 9	Lower better
Adult Nowicki-Strickland Locus of Control Scale (ANS)	0 to 40	Lower better
Alabama Parenting Questionnaire (APQ): corporal punishment	3 to 15	Lower better
Alabama Parenting Questionnaire (APQ): inconsistent discipline	6 to 30	Lower better
Alabama Parenting Questionnaire (APQ): involvement	10 to 50	Higher better
Alabama Parenting Questionnaire (APQ): poor monitoring/supervision	10 to 50	Lower better
Alabama Parenting Questionnaire (APQ): positive parenting	6 to 30	Higher better
Beck Depression Inventory (BDI)	0 to 63	Lower better
Beck Hopelessness Scale (BHS)	0 to 20	Lower better
Bipolar Disorder Symptom Scale (BDSS)	7 to 70	Lower better
Brief Psychiatric Rating Scale (BPRS)	18 to 126	Lower better
Brief Symptom Inventory (BSI) total score	0 to 212	Lower better
Centre for Epidemiological Studies Depression (CESD)	0 to 60	Lower better
Clinical Anxiety Scale	0 to 100	Lower better
Conners Adult ADHD rating scale - Observer: Screening Version (CAARS-OSV)	0 to 90	Lower better
Eyberg Child Behavior Inventory (ECBI) : intensity scale	36 to 252	Lower better
Eyberg Child Behavior Inventory (ECBI) : problem scale	0 to 36	Lower better
Formal Elements of Arts Therapy Scale rating guide (FEATS) – prominence of color	1 to 5	Higher better
Formal Elements of Arts Therapy Scale rating guide (FEATS) –color fit	1 to 5	Higher

Scale	Range	Direction
		better
Generalized Expectancy for Success Scale	30 to 150	Higher better
Hamilton rating scale for depression (HRSD) score	0 to 52	Lower better
Hamilton rating scale for depression (HRSD) score	0 to 52	Lower better
Health of the Nation Outcome Scales (HoNOS) total score	0 to 28	Lower better
Heartland Forgiveness Scale (HFS)	18 to 126	Higher better
Hospital Anxiety and Depression Scale (HADS) - Anxiety	0 to 21	Lower better
Hospital Anxiety and Depression Scale (HADS) - Depression	0 to 21	Lower better
Inventory of Interpersonal Problems (IIP-32)	0 to 128	Lower better
Mothers object relations scale (MORS) invasiveness	0 to 35	Lower better
Mothers object relations scale (MORS) warmth	0 to 35	Higher better
Parent development interview (PDI): reflexive functioning	-1 to 9	Higher better
Perceived stress scale (PSS)	0 to 40	Lower better
Positive and Negative Affect Schedule (PANAS): negative affect score	10 to 50	Lower better
Positive and Negative Affect Schedule (PANAS): positive affect score	10 to 50	Higher better
PTSD Symptom Scale (PSS)	0 to 51	Lower better
Rosenberg Self-Esteem Scale	0 to 30	Higher better
Short Inventory of Problems (SIP) follow-up	0 to 45	Lower better
Social Avoidance and Distress Scale (SADS)	0 to 28	Lower better
State-Trait Anxiety Inventory (STAI) - State	20 to 80	Lower better
State-Trait Anxiety Inventory (STAI) - Trait	20 to 80	Lower better
Symptom checklist 90 (SCL-90): global severity index	0 to 4	Lower better
Symptom checklist 90 (SCL-90): positive symptom distress index	0 to 4	Lower better
Symptom checklist 90 (SCL-90): positive symptom total	0 to 90	Lower better
Symptom Checklist-8D (SCL-8D)	0 to 8	Lower better
Texas Social Behavior Inventory (TSBI)	0 to 128	Higher better

3.8.2 Extrapolation

When answering review questions, if there was no direct evidence from a primary dataset,[e] based on the initial search for evidence, it was be appropriate to extrapolate from another data set. In this situation, the following principles were used to determine when to extrapolate:

- a primary dataset was absent, of low quality or was judged to be not relevant to the review question under consideration, and
- a review question was deemed by the GC to be important, such that in the absence of direct evidence, other data sources were considered, and

[e] A primary data set is defined as a data set which contains evidence on the population and intervention under review

- non-primary data source(s) were, in the view of the GC, available which may have informed the review question.

When the decision to extrapolate was made, the following principles were used to inform the choice of the non-primary dataset:

- the populations (usually in relation to the specified diagnosis or problem which characterises the population) under consideration shared some common characteristic but differed in other ways, such as age, gender or in the nature of the disorder (for example, a common behavioural problem; acute versus chronic presentations of the same disorder), and
- the interventions under consideration in the view of the GC have one or more of the following characteristics:
  - shared a common mode of action (for example, the pharmacodynamics of drug; a common psychological model of change - operant conditioning)
  - feasibility to deliver in both populations (for example, in terms of the required skills or the demands of the health care system)
  - shared common side effects/harms in both populations, and
- the context or comparator involved in the evaluation of the different datasets shared some common elements which support extrapolation, and
- the outcomes involved in the evaluation of the different datasets shared some common elements which support extrapolation (for example, improved mood or a reduction in behaviour that challenges).

When the choice of the non-primary dataset was made, the following principles were used to guide the application of extrapolation:

- the GC had to first consider the need for extrapolation through a review of the relevant primary dataset and be guided in these decisions by the principles for the use of extrapolation
- in all areas of extrapolation datasets were assessed against the principles for determining the choice of datasets. In general, the criteria in the 4 principles set out above for determining the choice should be met
- in deciding on the use of extrapolation, the GC had to determine if the extrapolation can be held to be reasonable, including ensuring that:
  - the reasoning behind the decision could be justified by the clinical need for a recommendation to be made
  - the absence of other more direct evidence, and by the relevance of the potential dataset to the review question could be established
  - the reasoning and the method adopted is clearly set out in the relevant section of the guideline.

3.8.3 Method used to answer a review question in the absence of appropriately designed, high-quality research

In the absence of appropriately designed, high-quality research (including indirect evidence where it would be appropriate to use extrapolation), both formal and informal consensus processes were adopted.

3.8.3.1 Formal method of consensus

The modified nominal group technique (Bernstein et al., 1992) was chosen due to its suitability within the guideline development process. The method is concerned with deriving a group decision from a set of expert individuals and has been identified as the method most commonly used for the development of consensus in health care (Murphy et al., 1998). The

nominal group technique requires participants to indicate their agreement with a set of statements about the intervention(s) of concern. These statements were developed by the NGA technical team drawing on the available sources of evidence on the methods of delivery and outcomes of the interventions. These sources of evidence could be supplemented by advice from external experts in the intervention(s). Agreement with the statements were rated on a 9-point Likert scale, where 1 represented least agreement and 9 represented most agreement. In the first round participants indicated the extent of their agreement with the statements and also provided written comment on their reason for any disagreement and how the statement could be modified.

In round 1, members were presented with an overview of the modified nominal group technique, a short summary of the available evidence, a consensus questionnaire containing the statements and instructions on the use of the questionnaire. Members were asked to rate their agreement with the statements taking into account the available evidence and their expertise. For the purpose of determining agreement, ratings were grouped into 3 categories to calculate the percentage agreement: 1–3 (inappropriate strategy), 4–6 (uncertain), or 7–9 (appropriate strategy or adaptation).

Where possible, in the afternoon of the GC meeting or at the subsequent GC meeting, anonymised distributions of responses to each statement were given to all members, together with members' additional comments and a ranking of statements based on consensus percentage agreement. Those statements with 80% or greater agreement were used to inform the drafting of recommendations, where appropriate taking into account the initial comments from and subsequent discussions with the GC.

For statements where there were 60 – 80% agreement a judgement was made based on the nature of the comments from the GC. If it appeared from the comments that the general principle included within the statement was agreed but that the comments could be addressed with some minor amendments incorporating the comments, the statements were used to inform the development of recommendations. Other statements that fell within this range were re-drafted based on the comments from the first rating and re-rated as in round 1 (round 2). If agreement at 80% or above on the re-rated was achieved, the statements were used to inform recommendations. Those that did not were discarded.

Any distribution of ratings with less than 60% agreement in round 1 was generally regarded as no consensus and discarded, unless obvious and addressable issues were identified from the comments.

3.8.3.2 Informal method of consensus

The informal consensus process involved a group discussion of what is known about the issues. The views of GC were synthesised narratively by a member of the review team, and circulated after the meeting. Feedback was used to revise the text, which was then included in the appropriate evidence review chapter.

3.9 Health economics methods

The aim of the health economics was to contribute to the guideline's development by providing evidence on the cost effectiveness of interventions and services covered in this guideline. This was achieved by a systematic literature review of existing economic evidence in all areas covered in the guideline.

Economic modelling was planned to be undertaken in areas with likely major resource implications, where the current extent of uncertainty over cost effectiveness was significant and economic analysis was expected to reduce this uncertainty, in accordance with The Guidelines Manual(NICE., 2014).Prioritisation of areas for economic modelling was a joint decision between the Health Economist and the GC. The rationale for prioritising review

questions for economic modelling was set out in an economic plan agreed between NICE, the GC, the Health Economist and the other members of the technical team. The following economic questions were selected as key issues to be addressed by economic modelling:

- Interventions to promote mental health and wellbeing, and modifications needed to psychological, social, pharmacological or physical interventions recommended in other NICE guidance
- Interventions for adults with a personality disorder
- Interventions for adults with a paraphilic disorder
- Recognition and assessment tools

In addition, literature on the health-related quality of life (HRQoL) of people covered by this guideline was systematically searched to identify studies reporting appropriate utility scores that could be utilised in a cost-utility analysis.

The identified clinical evidence on the areas prioritised for economic modelling was very sparse and allowed only a simple exploratory cost analysis assessing the impact of therapeutic community treatment for substance misuse treatment in imprisoned adults. The methods and results of this analysis are reported in Chapter 7.

In areas where modelling was not possible, the GC took into consideration resource implications and anticipated the cost effectiveness of interventions and services for people with mental health problems who are in contact with the criminal justice system when making recommendations.

The methods adopted in the systematic literature review of economic evidence are described in the remainder of this section.

3.9.1 Search strategy for economic evidence

3.9.1.1 Scoping searches

A broad preliminary search of the literature was undertaken in July 2014 to obtain an overview of the issues likely to be covered by the scope, and help define key areas. Searches were restricted to economic studies and HTA reports, and conducted in the following databases:

- Embase
- MEDLINE/MEDLINE In-Process
- HTA database (technology assessments)
- NHS Economic Evaluation Database (NHS EED).

Any relevant economic evidence arising from the clinical scoping searches was also made available to the health economist during the same period.

3.9.1.2 Systematic literature searches

After the scope was finalised, a systematic search strategy was developed to locate all the relevant evidence. The balance between sensitivity (the power to identify all studies on a particular topic) and specificity (the ability to exclude irrelevant studies from the results) was carefully considered, and a decision made to utilise a broad approach to searching to maximise retrieval of evidence to all parts of the guideline. Searches were restricted to economic studies and health technology assessment reports, and conducted in the following databases:

- Embase
- HTA database (technology assessments)

- MEDLINE/MEDLINE In-Process
- NHS EED
- PsycINFO.

Any relevant economic evidence arising from the clinical searches was also made available to the health economist during the same period.

The search strategies were initially developed for MEDLINE before being translated for use in other databases/interfaces. Strategies were built up through a number of trial searches, and discussions of the results of the searches with the review team and GC to ensure that all possible relevant search terms were covered. In order to assure comprehensive coverage, search terms for the guideline topic were kept purposely broad to help counter dissimilarities in database indexing practices and thesaurus terms, and imprecise reporting of study interventions by authors in the titles and abstracts of records.

For standard mainstream bibliographic databases (Embase, MEDLINE and PsycINFO) search terms for the guideline topic combined with a search filter for health economic studies. For searches generated in topic-specific databases (HTA, NHS EED) search terms for the guideline topic were used without a filter. The sensitivity of this approach was aimed at minimising the risk of overlooking relevant publications, due to potential weaknesses resulting from more focused search strategies. The search terms are set out in full in Appendix F.

3.9.2.3 Reference Management

Citations from each search were downloaded into reference management software and duplicates removed. Records were then screened against the inclusion criteria of the reviews before being quality appraised. The unfiltered search results were saved and retained for future potential re-analysis to help keep the process both replicable and transparent.

3.9.2.4 Search filters

The search filter for health economics is an adaptation of a pre-tested strategy designed by CRD (2007). The search filter is designed to retrieve records of economic evidence (including full and partial economic evaluations) from the vast amount of literature indexed to major medical databases such as MEDLINE. The filter, which comprises a combination of controlled vocabulary and free-text retrieval methods, maximises sensitivity (or recall) to ensure that as many potentially relevant records as possible are retrieved from a search. A full description of the filter is provided in Appendix F.

3.9.2.5 Date and language restrictions

Systematic database searches were initially conducted in February 2015 up to the most recent searchable date. Search updates were generated on a 6-monthly basis, with the final re-runs carried out in June 2016. After this point, studies were included only if they were judged by the GC to be exceptional (for example, the evidence was likely to change a recommendation).

Although no language restrictions were applied at the searching stage, foreign language papers were not requested or reviewed, unless they were of particular importance to an area under review. All the searches were restricted to research published from 2000 onwards in order to obtain data relevant to current healthcare settings and costs.

3.9.16 Other search methods

Other search methods involved scanning the reference lists of all eligible publications (systematic reviews, stakeholder evidence and included studies from the economic and clinical reviews) to identify further studies for consideration.

Full details of the search strategies and filter used for the systematic review of health economic evidence are provided in Appendix I.

3.9.2 Inclusion criteria for economic studies

The following inclusion criteria were applied to select studies identified by the economic searches for further consideration:

1. Only studies from Organisation for Economic Co-operation and Development countries were included, as the aim of the review was to identify economic information transferable to the UK context.
2. Selection criteria based on types of clinical conditions and service users as well as interventions assessed were identical to the clinical literature review.
3. Studies were included provided that sufficient details regarding methods and results were available to enable the methodological quality of the study to be assessed, and provided that the study's data and results were extractable. Poster presentations of abstracts were excluded.
4. Full economic evaluations that compared 2 or more relevant options and considered both costs and consequences as well as costing analyses that compared only costs between 2 or more interventions were included in the review. Non-comparative studies were not considered in the review.
5. Economic studies were included if they used clinical effectiveness data from a clinical trial, a prospective or retrospective cohort study, a study with a before-and–after design, or from a literature review. Studies with clinical effectiveness based on author's assumptions only were excluded.

3.9.3 Applicability and quality criteria for economic studies

All economic papers eligible for inclusion were appraised for their applicability and quality using the methodology checklist for economic evaluations recommended in The Guidelines Manual (NICE, 2014). All studies that fully or partially met the applicability and quality criteria described in the methodology checklist were considered during the guideline development process. The completed methodology checklists for all economic evaluations considered in the guideline are provided in Appendix R.

3.9.4 Presentation of economic evidence

The economic evidence considered in the guideline is provided in the respective evidence chapters, following presentation of the relevant clinical evidence. The references to included studies and the respective evidence tables with the study characteristics and results are provided in Appendix S. Characteristics and results of all economic studies considered during the guideline development process are summarised in economic evidence profiles provided in Appendix T.

3.9.5 Results of the systematic search of economic literature

The titles of all studies identified by the systematic search of the literature were screened for their relevance to the topic (that is, economic issues and information on HRQoL). References that were clearly not relevant were excluded first. The abstracts of all potentially relevant studies (41 references) were then assessed against the inclusion criteria for economic

evaluations by the health economist. Full texts of the studies potentially meeting the inclusion criteria (including those for which eligibility was not clear from the abstract) were obtained. Studies that did not meet the inclusion criteria, were duplicates, were secondary publications of 1 study, or had been updated in more recent publications were subsequently excluded. All economic evaluations eligible for inclusion (27 studies in 29 publications) were then appraised for their applicability and quality using the methodology checklist for economic evaluations. Finally, those studies that fully or partially met the applicability and quality criteria set by NICE were considered at formulation of the guideline recommendations.

3.10 Using NICE evidence reviews and recommendations from existing NICE clinical guidelines

When review questions overlapped and evidence from another guideline applied to a question in the current guideline, it was desirable and practical to incorporate or adapt recommendations published in NICE guidelines. Adaptation refers to the process by which an existing recommendation is modified in order to facilitate its placement in a new guideline. Incorporation refers to the placement of a recommendation that was developed for another guideline into a new guideline, with no material changes to wording or structure. In most cases incorporation was not used, as cross-referring to the other guideline was all that was necessary.

3.10.1 Incorporation

The following criteria were used to determine when a recommendation could be incorporated:

- a review question in the current guideline was addressed in another NICE guideline
- evidence for the review question and related recommendation(s) has not changed in important ways
- evidence for the previous question is judged by the GC to support the existing recommendation(s), and be relevant to the current question
- the relevant recommendation can ‘stand alone’ and does not need other recommendations from the original guideline to be relevant or understood within the current guideline.

3.10.2 Adaptation

The following criteria were used to determine when a recommendation could be adapted:

- a review question in the current guideline is similar to a question addressed in another NICE guideline
- evidence for the review question and related recommendations has not changed in important ways
- evidence for the previous question is judged by the GC to support the existing recommendation(s), and be relevant to the current question
- the relevant recommendation can ‘stand alone’ and does not need other recommendations from the original guideline to be relevant
- contextual evidence, such as background information about how an intervention is provided in the healthcare settings that are the focus of the guideline, informs the re-drafting or re-structuring of the recommendation but does not alter its meaning or intent (if meaning or intent were altered, a new recommendation should be developed).

In deciding whether to choose incorporation or adaption of existing guideline recommendations, the GC considered whether the direct evidence obtained from the current guideline dataset was of sufficient quality to allow development of recommendations. It was

only where (a) such evidence was not available or insufficient to draw robust conclusions and (b) where methods used in other NICE guidelines were sufficiently robust that the 'incorporate and adapt' method could be used. Recommendations were only incorporated or adapted after the GC had reviewed evidence supporting previous recommendations and confirmed that they agreed with the original recommendations.

When adaptation is used, the meaning and intent of the original recommendation is preserved but the wording and structure of the recommendation may change. Preservation of the original meaning (that is, that the recommendation faithfully represents the assessment and interpretation of the evidence contained in the original guideline evidence reviews) and intent (that is, the intended action[s] specified in the original recommendation will be achieved) is an essential element of the process of adaptation.

3.103 Roles and responsibilities

The guideline review team, in consultation with the guideline Facilitator and Chair, were responsible for identifying overlapping questions and deciding if it would be appropriate to incorporate or to adapt following the principles above. For adapted recommendations, at least 2 members of the GC for the original guideline were consulted to ensure the meaning and intent of the original recommendation was preserved. The GC confirmed the process had been followed, that there was insufficient evidence to make new recommendations, and agreed all adaptations to existing recommendations.

In evidence chapters where incorporation and adaptation have been used, the original review questions are listed with the rationale for the judgement on the similarity of questions. Tables are then provided that set out the original recommendation, a brief summary of the original evidence, the new recommendation, and the reasons for adaptation. For an adapted recommendation, details of any contextual information are provided, along with information about how the GC ensured that the meaning and intent of the adapted recommendation was preserved.

3.104 Drafting of adapted recommendations

The drafting of adapted recommendations conformed to standard NICE procedures for the drafting of guideline recommendations, preserved the original meaning and intent, and aimed to minimise the degree or re-writing and re-structuring.

3.11 From evidence to recommendations

Once the clinical and health economic evidence was summarised, the GC drafted the recommendations. In making recommendations, the GC took into account the trade-off between the benefits and harms of the intervention/instrument, as well as other important factors, such as the trade-off between net health benefits and resource use, values of the GC and society, the requirements to prevent discrimination and to promote equality[f], and the GC's awareness of practical issues (Eccles et al., 1998; NICE, 2012).

Finally, to show clearly how the GC moved from the evidence to the recommendations, each chapter (or sub-section) has a section called 'recommendations and link to evidence'. Underpinning this section is the concept of the 'strength' of a recommendation (Schünemann et al., 2003). This takes into account the quality of the evidence but is conceptually different. Some recommendations are 'strong' in that the GC believes that the vast majority of healthcare professionals and service users would choose a particular intervention if they considered the evidence in the same way that the GC has. This is generally the case if the benefits clearly outweigh the harms for most people and the intervention is likely to be cost effective. However, there is often a closer balance between benefits and harms, and some

[f]See NICE's equality scheme: www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp

service users would not choose an intervention whereas others would. This may happen, for example, if some service users are particularly averse to some side effect and others are not. In these circumstances the recommendation is generally weaker, although it may be possible to make stronger recommendations about specific groups of service users. The strength of each recommendation is reflected in the wording of the recommendation, rather than by using ratings, labels or symbols. The word 'offer' was used for recommendations with strong evidence whereas 'consider' was used to make recommendations with limited evidence.

Where the GC identified areas in which there were uncertainties or where robust evidence was lacking, they developed research recommendations. Those that were identified as 'high priority' were developed further in the NICE version of the guideline, and presented in Appendix G.

3.12 Stakeholder contributions

Professionals, service users, and companies have contributed to and commented on the guideline at key stages in its development. Stakeholders for this guideline include:

- service user and carer stakeholders: national service user and carer organisations that represent the interests of people whose care will be covered by the guideline
- local service user and carer organisations: but only if there is no relevant national organisation
- professional stakeholders' national organisations: that represent the healthcare professionals who provide the services described in the guideline
- commercial stakeholders: companies that manufacture drugs or devices used in treatment of the condition covered by the guideline and whose interests may be significantly affected by the guideline
- providers and commissioners of health services in England and Wales
- statutory organisations: including the Department of Health, the Welsh Assembly
- Government, NHS Quality Improvement Scotland, the Care Quality Commission and the National Patient Safety Agency
- research organisations: that have carried out nationally recognised research in the area.

NICE clinical guidelines are produced for the NHS in England and Wales, so a 'national' organisation is defined as one that represents England and/or Wales, or has a commercial interest in England and/or Wales.

Stakeholders have been involved in the guideline's development at the following points:

- commenting on the initial scope of the guideline and attending a scoping workshop held by NICE
- commenting on the draft of the guideline.

3.13 Validation of the guideline

Registered stakeholders had an opportunity to comment on the draft guideline, which was posted on the NICE website during the consultation period. Following the consultation, all comments from stakeholders and experts were responded to, and the guideline updated as appropriate. NICE also reviewed the guideline and checked that stakeholders' comments had been addressed.

Following the consultation period, the GC finalised the recommendations and the NGA produced the final documents. These were then submitted to NICE for a quality assurance check. Any errors were corrected by the NGA, then the guideline was formally approved by NICE and issued as guidance to the NHS in England and Wales.

1

4 Staff Training

4.1 Introduction

Mental health work with adults in contact with criminal justice system is not new. It is therefore surprising that bespoke mental health training for the range of practitioners and range of clinical problems is rare and piecemeal. Both the patterns of morbidity, with high rates of co-morbidity, and the multi-agency contexts in which mental health care is delivered make the acquisition and application of mental health skills and knowledge skills important.

In recent years, notably following the transfer of commissioning responsibility for prison health care to the NHS in 2006, the scope of mental health work with this client group has increased and encompasses not only expanded prison health care but also liaison and diversion in police custody suites and criminal courts, as well as established forensic services in hospitals and community settings. Much of this work is undertaken by clinicians in partnership with criminal justice practitioners whose own agencies have traditionally offered little by way of relevant professional development in mental health. Health personnel require not only practice familiarity with criminal justice settings but also an understanding of relevant justice roles, principles and procedures.

The context in which this guideline has been developed in one of increasing recognition at national and regional level of the need for such specialist training and accreditation for clinicians. Relevant professional bodies, Health Education England, Higher Education Institutes and clinical services commissioners and providers will all have roles in taking this forward. Relatedly, health providers' partner agencies, that are prisons (in particular prison officers), the police service, National Probation Service and Community Rehabilitation Companies, need training and support to be equipped to play their part in the identification of mental health difficulties, the assessment of such problems were appropriate, and signposting to available care and in some cases its coordination and delivery. The complexity of service delivery in this field underlines the importance of effective multidisciplinary work and the need to educate staff groups in how to make inevitably complicated systems work to the advantage of those with mental health problems. Members of the judiciary, crown prosecutors, defence lawyers and court staff similarly require mental health and learning disability awareness training.

These challenges underlie the attempt in this chapter to examine the available literature on effective support, training, education and supervision in this area of work. The range of clinical problems, the variety of settings and the multiplicity of practitioners from different agencies with different levels of experience, training and interest, make it inevitable that continuing professional development should cover wide territory and require multiple delivery systems. Given the enhanced scope of services nationally and evidence of an ongoing government commitment to further expansion e.g. liaison and diversion schemes and re-organisation of prisons, there is both a need and opportunity to build on the limited progress on specialist training so far available.

4.2 Review question: What are the most effective support, training and education, and supervision programmes for health, social care or criminal justice practitioners to improve awareness, recognition, assessment, intervention and management of mental health problems in adults in contact with the criminal justice system?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 11. A complete list of review questions and full review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 11: Clinical review protocol summary for the review of staff training and supervision

Component	Description
Population	Adults with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Intervention(s)	Any staff support, training or supervision programme
Comparison	• Training or education as usual • Training or education Waitlist control • Placebo (including attention control) • Any alternative staff training or education programme
Outcomes	• Critical – Offending and re-offending outcomes; Mental health outcomes; Identification of mental health problems • Important – Number of staff assessed as being competent
Study design	Systematic reviews, RCTs

4.2.1 Group consensus for the most effective support, training and education, and supervision programmes for health, social care or criminal justice practitioners

In a search of the literature only 1 RCT was identified for this question. The GC reviewed this trial (see section 4.2.2 below) and also considered the economic evidence in section 4.2.3. The GC considered this to be a very limited evidence base on which to make any recommendations for training. For this question the GC thought that recommendations should be based on robust (RCT) evidence given the potential cost and feasibility of implementing support, education and training changes across the entire criminal justice system. They also deemed it inappropriate to go down the evidence hierarchy as in this case they were concerned about the implications of making recommendations based upon poorer quality evidence about specific approaches. They also choose not to examine indirect evidence as they decided that it would not be possible to extrapolate from other populations to the criminal justice system, which is provided in a diverse range of set of settings and is significantly different from areas covered in NICE mental health guidelines. The GC therefore decided to develop a set of principles to inform practice in this area (rather than supporting specific forms of intervention) and recommendations for staff training in those working within the criminal justice system using a modified form of the Nominal Group Technique. The method used for the technique is described in full within the methods section in Chapter 3.

Key issues related to staff training were identified by the NGA technical team from published evidence identified in literature searches and from discussions within the GC meetings.

These issues were used to generate nominal statements covering a range of areas that had been identified as important by the GC. These included mental health awareness and knowledge and the knowledge and skills needed to deliver interventions. These statements were grouped together in the form of a questionnaire and distributed to the GC to be rated. An example of a statement that was rated highly by the committee is 'Staff should receive training about commonly occurring mental health problems (e.g., substance misuse, neurodevelopmental disorders, acquired cognitive impairment, personality disorder) in the criminal justice system and the impact these may have'.

The questionnaire was completed by 15 of the 19 GC members. Some members were unable to attend the relevant committee meeting, however they had the opportunity to discuss the statements from the nominal group process and contributed to the subsequent recommendations. Percentage consensus values were calculated, and comments collated, for each statement. The rankings and comments were then presented to the GC members, and used to inform a structured discussion within the GC meeting. Agreement within the GC was high enough that a second round of ratings was not deemed necessary. This discussion led to the development of recommendations in this area. A brief summary of the outcome of this process is given in Table 12 below. The full list of statements and ratings can be found in Appendix V and blank copies of the questionnaires used can be found in Appendix U.

Table 12: Summary of the nominal group technique process followed for the development of recommendations for staff training of those who work within the criminal justice system

Round 1		**Round 2**		**No. of recommendations generated**
Level of agreement	Statements N (total=23)	Level of agreement	Statements N (total=0)	6 recommendations
High	22	High	n/a	
Moderate	1	Moderate	n/a	
Low	0	Low	n/a	

High agreement was 80% or greater agreement, moderate was 60 to 80% agreement and low was less than 60% agreement.

4.2.2 Clinical Evidence

One RCT (N = 847) met the eligibility criteria for this review: Friedmann 2015 (Friedmann et al., 2015). This US study was judged by the GC to have relatively limited application to the UK health care system due to the different criteria for initiating opiate substitution treatment, (OST) the populations for which it is prescribed and the relatively limited knowledge of OST in many parts of the US health care and criminal justice systems.

An overview of the included study can be found in Table 13. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in

Table 14. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 13 Study information for Friedman et al (2015) included in the analysis of organisational linkage intervention (OLI) in addition to training of medication-assisted therapy (MAT)

	OLI plus training of MAT vs MAT alone
Total no. of studies (N[1])	1 (847)
Study ID	Friedmann 2015

	OLI plus training of MAT vs MAT alone
Study design	RCT
Country	USA
Underlying Mental Health Disorder	The training was to educate staff about treatment substance (alcohol and/or drug) misuse disorders.
Diagnosis	Clinical
Age (mean) years	46
Gender (% female)	63
Ethnicity (% white)	61.5
Intervention	OLI plus training of MAT
Comparator	Training of MAT alone
Criminal Justice setting	Community correction agency
Format (number of participants per group)	Group (10/group)
Dose/Intensity	Not reported
Treatment length (weeks)	52
Follow-up length (weeks)	Not reported
Notes. RCT = randomised controlled trial [1]Number randomised.	

Table 14 Summary of findings table for OLI plus training of MAT vs training alone for substance misuse disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with training alone	Risk difference with Organisational Linkage Intervention (OLI) plus training (RQ 5.1) (95% CI)
Familiarity with methadone[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.26 (SD 1.01)	MD 0.14 higher (0.03 lower to 0.31 higher)
Referral knowledge for methadone[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.24 (SD 1.23)	MD 0.04 higher (0.11 lower to 0.19 higher)
Intent to refer clients for MAT with methadone[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.05 (SD 1.24)	MD 0.38 higher (0.19 to 0.57 higher)
Overall perception and knowledge of methadone[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.01 (SD 0.04)	MD 0.2 higher (0.13 to 0.27 higher)
Familiarity with buprenorphine[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.39 (SD 1.52)	MD 0.01 higher (0.19 lower to 0.21 higher)
Referral knowledge for buprenorphine[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.34 (SD 1.33)	MD 0.07 higher (0.12 lower to 0.26 higher)
Intent to refer clients to MAT with buprenorphine[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.15 (SD 1.35)	MD 0.15 higher (0.02 lower to 0.32 higher)
Overall perception and knowledge of buprenorphine[2]	847 (1 study)	⊕⊕⊕⊖ MODERATE[1]	-	Mean 0.03 (SD 0.66)	MD 0.13 higher (0.05 to 0.21 higher)
1. Friedmann 2015 - unclear randomisation and concealment; comparable management of experimental and control group; appropriate outcome report *2. Change from baseline to post intervention; range -4 to 4; higher is better*					

4.213 Economic evidence

The systematic search of the literature identified 1 study that assessed the costs and consequences of a police training programme in Canada (Krameddine et al., 2013). Details on the methods used for the systematic review of the economic literature are described in Chapter 3; full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix S. Completed methodology checklists of the studies are provided in Appendix R. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix T.

Krameddine and colleagues (2013) assessed the costs and consequences of police officer training to improve interaction with people who might have mental health problems versus no training in Canada. The intervention involved a 1 day scripted role-play training, which required police officers interacting with highly trained actors during six realistic scenarios with the aim of improving empathy, communication skills, and the ability of officers to de-escalate potentially difficult situations when dealing with people who have mental health problems. The economic analysis was based on a before-after observational study (N=663). The time horizon of the economic analysis was 7 months, and its perspective was a service provider. Cost elements comprised only the programme provision (staff time, actors' training and attendance). Clinical effectiveness and resource use data were obtained from the before-after study. The source of unit cost data was not reported. The measures of outcome utilised in the economic analysis were: measures of police officer attitude (total Community Attitudes toward Mental Illness [CAMI] scale, total Social distance scale [SDS]; measurement of police officer knowledge (mental illness recognition scale, mental illness knowledge); police officer behavioural measures (supervising officer survey using 5-point Likert scale), number of mental health calls identified, time spent on mental health calls, and use of force. Additionally, all indirect behavioural measures compares the same 6-month period (July–December) for the years 2009–2011.This is because the training took place throughout the second quarter of 2011 and thus it wasn't possible to accurately compare the first 6-months of these 3 years.

According to the analysis, to train a cohort of a total of 663 officers on 19 separate training days, and with several days of advance training of the actors, the cost was slightly less than $80,000 CAN dollars or ~$120 per officer (in likely 2012 prices). In terms of effectiveness no significant changes were observed on CAMI and SDS scales, or mental illness knowledge scores. The mean scores on the mental illness recognition scale were 1.9 (SD 2.8) and 1.3 (SD 2.9) at baseline and follow-up, respectively; an improvement of 0.6, $p = 0.011$. The mean scores relating to the ability to communicate with the public (as rated by the supervising officer) were 3.49 (SD 0.86) and 3.73 (SD 0.77) at baseline and follow-up, respectively; an improvement of 0.24, $p = 0.001$. The mean scores relating to the ability to verbally de-escalate situation (as rated by the supervising officer) were 3.39 (SD 0.87) and 3.65 (SD 0.79) at baseline and follow-up, respectively; an improvement of 0.26, $p < 0.001$. Similarly, the mean scores relating to the level of empathy with the public were 3.51 (SD 0.73) and 3.73 (SD 0.73) at baseline and follow-up, respectively; an improvement of 0.22, $p = 0.003$.

The mean number of mental health calls during the 6 month period (July to December) was 162 for the year 2009, 182 for the year 2010, and 257 for the year 257. An increase of 20 calls between 2010 and 2009 was statistically significant, $p = 0.031$. Similarly, there was a statistically significant increase in calls between 2011 and 2010 (an increase of 75 calls), $p < 0.001$. This indicates that police officers were better equipped to identify a call as being due to mental health issues. An increase in the number of calls being identified as being due to mental health issues was a positive outcome of the training programme.

The mean time per mental health call during the 6 month period (July to December) was 221 min (SD 142) for the year 2009, 251 min (SD 164) for the year 2010, and 205 min (SD 146)

for the year 2011. An increase between 2010 and 2009 was 30 min and it was statistically significant, $p \leq 0.001$. Between 2011 and 2010 there was a decrease in the mean duration of mental health call by 46 min and it was also statistically significant, $p < 0.001$. This indicates that police officers became more confident in their interactions with mentally ill individuals, and that they became more efficient in the use of their time when dealing with mentally ill individuals. A reduction in the time taken per mental health call was a positive outcome of the training programme.

The percentage of times police force was used in any Mental Health Call during the 6 month period (July to December) was 11.5 (SD 1.9) for the year 2009, 8.0 (SD 1.2) for the year 2010, and 5.2 (SD 0.9) for the year 2011. The reduction of 3.5% between 2010 and 2009 was statistically significant, $p = 0.011$; and also the reduction of 2.6% between 2011 and 2010 was statistically significant, $p = 0.0004$.

The authors estimated that if the time spent per mental health call in 2011 was the same as the time spent in 2010, then there would have been an additional expenditure of approximately $84,000 in the 6-month period from July to December 2011. Based on the above findings, the authors concluded that 1-day training course significantly changed behaviour of police officers in meaningful ways and also led to cost-savings.

4.24 The study is only partially applicable to the NICE decision-making context, as it has been conducted in Canada and adopted a narrow local service provider perspective. The study was judged by the GC to have potentially serious limitations, including its design (before-after study), the relatively short time horizon (7 months), the inclusion of intervention costs only, the lack of consideration of outcomes (including health outcomes) for people with mental health problems and the omission of wider healthcare and social care costs, and the source of unit cost data was unclear.Clinical evidence statements

Moderate quality evidence from one RCT (n=847) showed that OLI plus training of MAT in the US health care system lead to a clinically important increase in the overall perception and knowledge of methadone and increase intent to refer clients who were on methadone to MAT compared to training only, but the difference was not seen in buprenorphine substitute therapy. The referral knowledge and familiarity with medication did not differ between the two groups. There considerable concerns about the applicability of the trial to the UK health care and criminal justice systems.

4.25 Clinical evidence statements based on formal consensus ratings

To improve mental health awareness, the GC agreed that staff should receive training (provided in a multi-disciplinary setting where possible) about:

- the prevalence of mental health problems within criminal justice settings
- the most commonly occurring mental health problems, their impact upon the service users and how to recognise these
- how to communicate effectively with those with mental health problems
- how to notice changes in behaviour that may indicate mental health problems
- the stigma associated with mental health problems and the need to avoid judgemental attitudes
- common protocols for dealing with mental health problems in this group.

Regarding the delivery of interventions, the GC agreed that:

- staff should be trained in how to make appropriate referrals, the effectiveness of interventions and different management strategies, assessment and management of self-harm and in de-escalation techniques
- staff should receive training in stress management for themselves
- they should receive regular clinical supervision to enable them to deliver interventions within the criminal justice system
- teams who carry out assessments and intervention work should receive training, including on the management of critical incidents, and supervision to ensure they are competent to do so

Regarding knowledge of mental health and criminal justice services the GC decided that:

- all staff should receive a comprehensive induction including information about the purpose of their service and others available locally with which their service users may be involved
- staff should receive training relating to relevant legislation and local policies on information sharing.

The GC expressed moderate support for staff receiving information regarding commonly used terms and acronyms.

4.2.6 Economic evidence statements

There was evidence from 1 cost-consequences analysis based on a before-after observational study (N=663). The findings showed that there were no changes in the attitudes of the police toward the mentally ill. However, there was a significant increase in the recognition of mental health issues as a reason for a call, training was associated with an improvement in efficiency when dealing with mental health issues, and a decrease in weapon or physical interactions with mentally ill individuals. Overall, 1-day training course for police officers seemed to lead to better outcomes and potential cost-savings. This is a Canadian study and it is only partially applicable to the NICE decision-making context. It is characterised by potentially serious limitations including its before-after design, it's relatively short time horizon (7 months), the inclusion of intervention costs only, the lack of consideration of outcomes (including health outcomes) for people with mental health problems and the omission of wider healthcare and social care costs.

4.3 Recommendations and link to evidence

Recommendations	
	1. Commissioners and providers of criminal justice service and healthcare services should provide all staff working in the criminal justice system, who provide direct care or supervision, a comprehensive induction, covering: • **the purpose of the service in which they work, and the role and availability of other related local services, including pathways for referral** • **the roles, responsibilities and processes of criminal justice, health and social care staff** • **legislation and local policies for sharing information with others involved in the person's care** • **protocols for dealing with mental health problems in the criminal justice system (for example, in-possession medicines, side-effects, withdrawal)**

	• **the importance of clear communication, including avoiding acronyms, and using consistent terminology.** 2. **Commissioners and providers of criminal justice service and healthcare services should educate all staff about:** • **the stigma and discrimination associated with mental health problems and associated behaviours, such as self-harm** • **the need to avoid judgemental attitudes** • **and the need to avoid using inappropriate terminology.** 3. **Provide multi-disciplinary and multi-agency training to increase consistency, understanding of ways of working, and promotion of positive working relationships for all staff who work in the criminal justice system on:** • **the prevalence of mental health problems in the criminal justice system, and why such problems may bring people into contact with the criminal justice system** • **the main features of commonly occurring mental health problems seen in the criminal justice system (for example, substance misuse, neurodevelopmental disorders, acquired cognitive impairment, personality disorder, depression, anxiety disorders, psychosis, post-traumatic stress disorder [PTSD]), and the impact these may have on behaviour and compliance with rules and statutory requirements** • **recognising and responding to mental health problems and communication problems that arise from, or are related to, physical health problems.** 4. **Give all staff involved in direct care, training and supervision to support them in:** • **dealing with critical incidents, including emergency life support** • **managing stress associated with working in the criminal justice system and how this may affect their interactions with people and their own mental health and wellbeing** • **the recognition, assessment, treatment, and management of self-harm and suicide** • **de-escalation methods to minimise the use of restrictive interventions** • **recognition of changes in behaviour, taking into account that these may indicate the onset of, or changes to, mental health problems.** • **knowledge of effective interventions for mental health problems** • **developing and maintaining safe, boundaries and constructive relationships**

	• **delivering interventions within the constraints of the criminal justice system (for example, jail craft training, formulation skills).**
Relative values of different outcomes	The GC were aware of the high prevalence of mental health problems and the poor recognition of these in the criminal justice system. Staff understanding of the nature of mental health disorders and their impact on functioning and possible links to offending behaviour was also discussed The GC placed considerable importance on the recognition of problems, access to appropriate care, the delivery of effective interventions to address mental health problems. The GC were also aware of the range of organisations involved in the delivery of care with very different cultures. Training knowledge and attitudes would therefore also need to consider matters of organisational culture. The GC were also concerned about the levels of self-harm particularly in the prison services and identified this as an important area to address in terms of training needs.
Trade-off between clinical benefits and harms	The GC discussed the potential benefits of comprehensive staff training relating to mental health and inter-agency working. Namely, higher detection rates, earlier intervention and better outcomes for service users. Additionally, the GC discussed the potential indirect benefits, such as the positive impact of a respectful and informed awareness of mental health which could result in better working relationships and a more positive culture generally within the criminal justice system. The GC noted that these benefits may extend to the staff, and that better awareness, greater self-efficacy and support may contribute to an improved working environment. The single RCT conducted in the United States reviewed for this question suggested some potential benefit of a training programme in terms of increased knowledge of substance misuse problems and referral to services but the GC were concerned about its applicability to care in the UK criminal justice system given the very different approaches adopted to substance misuse in the United States. In determining the particular content of each recommendation, the GC were guided by the key statements developed through the nominal group technique and their expert knowledge and experience of the criminal justice system. This informed not only the content of the recommendation but the overall structure of the recommendations focussing on multi-disciplinary and multi-agency training for all staff, the need for regular training and a focus on specific concerns such as critical incidents and the need to ensure all staff have a basic understanding of the delivery of mental health interventions. The GC considered the possible harms, for example those arising from a false positive identification of mental health problems including unnecessary anxiety, inappropriate assessment or intervention. The GC did not consider the harms to out-weigh the benefits The GC also discussed the benefits of supervision and support for practitioners delivering interventions to treat mental health problems in service users. The GC were of the opinion that these included benefits for the service user in terms of improved clinical outcomes as well as benefits for the staff providing the interventions in terms of feeling competent and well-supported. The GC were not able to identify any potential harms associated with improved staff training and support.
Trade-off between net health benefits	There was very limited economic evidence on staff working in the criminal justice system training.

and resource use	The GC considered that all staff working in the criminal justice system already receive an induction. So offering a comprehensive induction (as outlined in the recommendation 1.8.1) would not incur significant extra resource implications. Very limited economic evidence suggests that 1-day police officer training to improve interaction with people who might have mental health problems is potentially cost-effective. The 1-day training was associated with an improvement in staff ability: to recognise better mental health problems, to communicate with the public about mental health problems, to verbally de-escalate difficult situation. There was also an improvement in the levels of empathy with the public, police officers became more efficient and confident in their interactions with mentally ill, and there was a reduction in the use of police force. There was no evidence on training for other staff working in the criminal justice system, nor was there evidence on sustained training. However, the GC considered limited existing evidence and concluded that training for staff working in the criminal justice system has the potential to significantly change their behaviour in meaningful and positive ways, ensure better recognition of mental health problems and assessment of need (facilitating timely and appropriate treatment), and make their interactions more efficient, and as a result lead to the overall cost-savings. It must be noted that the economic evidence came from a Canadian study which was only partially applicable to the NICE decision-making context. It was characterised by potentially serious limitations including its before-after design, it's relatively short time horizon (7 months), the inclusion of intervention costs only, the lack of consideration of outcomes (including health outcomes) for people with mental health problems and the omission of wider healthcare and social care costs. In developing the recommendations for training the GC drew on their knowledge and experience and considered that the current poor recognition and lack of uptake of currently available interventions had a significant and negative effect on the overall use of resource in the criminal justice system, in particular in the prison service, and contributed to a higher prevalence of critical incidents which resulted in unplanned use of additional resources. The intention in this recommendation is not that all staff should be trained in delivering the mental health interventions but that staff should be aware of the nature and outcomes of such interventions.
Quality of evidence	Little high quality evidence was identified for this question, with a single RCT were the evidence was of moderate quality but the GC were considered that the very different approaches to the treatment of substance misuse in the United States and the UK meant that it had limited applicability. Therefore the GC used a formal consensus method (the nominal group technique) to further inform development of the recommendations in this area. High consensus (80%) was reached by the GC for the majority of the statements (22 out of 23).
Other considerations	The GC drew on their expert knowledge and experience of the multi-agency and multidisciplinary nature of mental health care in the criminal justice system and, in particular, their awareness of the frequent movement of staff and therefore decided that recommendations should consider both multi-disciplinary and refresher training to account for the frequent movement of staff. The GC were aware that the introduction of case identification or

	recognition tools or methods alone would not lead to improved recognition. They were of the view that all staff using such tools or methods should be trained in their use and interpretation. However there were no established methods for training across the CJS and so the GC decided to make a recommendation for further research into training methods in this area.

4.321 Research recommendations

1. What staff training models improve identification of mental health problems and clinical outcomes for adults in contact with the criminal justice system?

There is limited evidence on the effective models for the training and supervision of practitioners working in the criminal justice system which could best support the identification of mental health problems in the CJS. A series of studies are required to assess the best methods to improve the recognition of the full range of mental health programmes. These studies should be of adequate size and cover the range of health, social and criminal justice staff.

There is insufficient evidence to determine the best methods to deliver effective training to improve the identification of mental health problems in the criminal justice system. Lack of adequate training leads to under-recognition and sub-optimal treatment. Programmes need to be designed and evaluated which are specially developed with the needs of those working in the criminal justice system in mind. There is good evidence that the provision of training alone is unlikely to bring about substantial changes in staff behaviours without adequate service style change and the provision of high quality supervision. The nature of service style changes and the supervision training should also be evaluated.

Important outcomes could include:

- Staff competence
- Improved recognition of mental health problems
- Improved access to and uptake of mental health interventions.

5 Recognition and assessment

5.1 Introduction

There are many barriers to recognition of mental health problems for people in contact with the criminal justice system. Adults in contact with the criminal justice system may have greater difficulties in making use of services that are available, partly due to problems with establishing trusting relationships with those who can be perceived as authority figures, partly due to a tendency to increased social transience and in some cases difficulties with personal organisation. Staff working in the criminal justice system may lack awareness of the prevalence of mental health problems and how these problems can present themselves. Often, even serious mental illness is lost sight of as a cause or contributor to offending. Stigma against both those who offend and those who are mentally ill may also play a part.

Appropriate recognition of these issues is important for a number of reasons. Contact with the CJS may be an important opportunity to identify and address the needs of individuals who are disengaged from other services and thus would otherwise have their needs unmet. For a minority their poor mental health may contribute to their risk of harm to themselves or others, for example by exacerbating or triggering urges to self-harm or attempt suicide, exposing vulnerable people to the potential for bullying. Early identification, for example through liaison and diversion services, enables appropriate support to be offered at pivotal points in their journey through the CJS.

It is important for case identification and assessments to be timely, appropriate, done with reference to any existing medical and social care records and assessments and for them to have a positive impact on the individual's pathway through the CJS. Information sharing can be problematic even where local protocols exist. Too often assessments are episodic, with multiple similar assessments completed in various settings which contribute to disengagement, information overload, and with no positive outcome for the individual. In developing the recommendation set out in this chapter the GC were mindful of the varying skills and experience of staff including police and prisoner officers who have often to assess for immediate risks, the setting in which identification and assessment can take place (e.g. on the street or in prison reception), the capacity of the non-health care system to direct people into effective assessment and the interface between the health care and criminal justice systems.

Outside to the prison services currently there are no well-established and routinely used case identification or assessment tools and procedures. The distinctive nature and patterns of presentation in this guideline's target groups makes it desirable that specific assessment tools and processes be identified that could offer advantages over generic approaches.

5.2 Review question: What are the most appropriate tools for the recognition of mental health problems, or what modifications are needed to recognition tools recommended in existing NICE guidance, for adults:

- in contact with the police?
- in police custody?
- for the court process?
- at reception into prison?

- at subsequent time points in prison?
- in the community (serving a community sentence, released from prison on licence or released from prison and in contact with a community rehabilitation company [CRC] or the probation service)?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 15. A complete list of review questions and full review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 15: Clinical review protocol summary for the review of the most appropriate tools for the recognition of mental health problems

Component	Description
Population	Adults (aged 18 and over) with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Index test(s)	Any formal recognition and assessment tools considered appropriate and suitable for use
Reference standard	Diagnosis Statistical Manual (DSM) or International Classification of Diseases (ICD) diagnosis
Outcomes	Sensitivity: the proportion of true positives of all cases diagnosed with the target condition in the population Specificity: the proportion of true negatives of all cases not-diagnosed with the target condition in the population Reliability (for instance, inter-rater or test-retest reliability or internal consistency) Validity (for instance, criterion or construct validity)
Study design	Systematic reviews of diagnostic test accuracy studies, diagnostic cross-sectional studies

5.2.1 Clinical evidence

The literature search yielded 8948 articles overall for the review questions about the most appropriate tools:

- for the recognition of mental health problems,
- to support or assist in the assessment of mental health problems
- to support or assist in risk assessment

Scanning titles or abstracts identified 954 articles potentially relevant to the above review questions.

The GC agreed that for a tool to be considered appropriate and suitable for use for recognition it should: 1) have ≤28 items, 2) take ≤5 minutes to administer, 3) be able to be completed by a non-expert, and 4) be free to use where possible. Further, the decision was made for all three review questions to only review tools targeting disorders covered by existing NICE guidance if there was a substantial evidence base for tools for such disorders in the criminal justice system or when assessed by criminal justice specific tools that intend to assess multiple mental health issues. This decision was made for two reasons: 1) referring into existing guidance for specific disorders could provide a stronger evidence base than the limited number of studies for a given disorder in the criminal justice system, and 2) it was considered more practical to recommend a tool that was applicable to multiple mental health problems than recommending the use of multiple tools that are disorder specific.

After further inspection of the full articles, 926 studies did not meet one or more eligibility criteria outlined above. An additional seven studies forwarded by stakeholders, three studies identified by handsearching, and one study identified by another literature search for this guideline also did not meet the inclusion criteria. The most common reasons for exclusion were that: there was no appropriate reference standard, the population was not relevant (individuals cared for in hospital, not in contact with the criminal justice system, or aged under 18 years), or sensitivity and specificity were not presented (or sufficient information to allow for their calculation). This resulted in 10 articles representing 11 studies that were included for review question 2.1, one study that was included for review question 2.2 and 17 articles representing 18 studies that were included for review question 2.3.

There were two additional studies (McKinnon & Grubin, 2014; McKinnon et al., 2015) forwarded by stakeholder which met the inclusion criteria resulting in a total of 12 articles, representing 13 studies, that provided sufficient data to be included in the evidence synthesis for review question 2.1: (Baksheev et al., 2012; Ford et al., 2007; Ford et al., 2009; Harrison & Rogers, 2007; Louden et al., 2013; McKinnon & Grubin, 2014; McKinnon et al., 2015; Sacks et al., 2007a; Sacks et al., 2007b; Steadman et al., 2007; Steadman et al., 2005; Teplin & Swartz, 1989).

All studies were published in peer-reviewed journals between 1989 and 2015. Of, these eligible studies, five reported on the Brief Jail Mental Heal Screen (BJMHS; Steadman et al., 2005) or the revised version of the BJMHS (BJMHS-R; Steadman et al., 2007), four reported on the Referral Decision Scale (RDS; Teplin & Swartz, 1989) or its subscales, two reported on the Co-occurring Disorders Screening Instrument for Mental Disorders (CODSI-MD; Sacks et al., 2007a), two reported on the Co-occurring Disorders Screening Instrument for Severe Mental Disorders (CODSI-SMD; Sacks et al., 2007a), two reported on the Correctional Mental Health Screens for Men (CMHS-M; Ford et al., 2007) and Women (CMHS-W; Ford et al., 2007), two reported on the HELP-PC (McKinnon & Grubin, 2014) and two reported on the Custody Risk Assessment Form (Baksheev et al., 2012). Characteristics of these recognition tools can be found in Table 16.

Further information about both included and excluded studies and the full methodological checklists can be found in Appendix K. A summary of the methodological quality of the studies is presented in Table 17. If data was presented in sufficient detail for analysis, the data are presented using forest plots and summary ROC curves in Appendix O.

Table 16: Characteristics of tools included in the review of the most appropriate tools for the recognition of mental health problems

Tool	Target disorder	Intended population/setting	Scale information	Recommended cut-off	Format	Administration & qualifications	Cost/restrictions
BJMHS/ BJMHS-R	Serious mental illness	Prison	BJMHS: 8 items BJMHS-R: 12 items	≥2 from section 1 or ≥1 from section 2	Questionnaire administered by staff	Administration time: 2-3 minutes Administered by criminal justice service professionals following training.	Freely available from: http://www.prainc.com/wp-content/uploads/2015/10/bjmhsform.pdf
CODSI-MD/ CODSI-SMD	CODSI-MD: general mental health CODSI-SMD: serious mental illness	Prison substance abuse treatment programs	CODSI-MD: 6 items CODSI-SMD: 3 items	CODSI-MD: ≥3 CODSI-SMD: ≥2	Questionnaire administered by staff	Administration time: Unclear as they have only been administered as part of a test battery No specialist training required	Freely available from: http://www.ndri.org/manuals-and-instruments.html
CMHS-M	General mental health	Prison	12 items	≥6	Questionnaire administered by staff	Administration time: 3-5 minutes Administered by criminal justice or healthcare staff	Freely available from: http://www.asca.net
CMHS-W	General mental health	Prison	8 items	≥5	Questionnaire administered by staff	Administration time: 3-5 minutes Administered by criminal justice or healthcare staff	Freely available from: http://www.asca.net
Custody Risk Assessment	Risk	Police custody	Total number of items NR	≥1	Completed by police officer	Administration time: unclear	Unclear. Appears to be a local form used

Form			Depressed/ suicidal: 1 item Mental illness: 1 item				by one police station.
HELP-PC	General mental health and learning disabilities	Police custody	Embedded in wider assessment Mental health subscale: number of items not reported Learning disabilities subscale: 4 items (3 questions and 1 observation)	≥1	Interview and observation	Administration time: Median time by end of pilot 7.75 minutes Administered by custody officers. Details of training not reported.	Does not appear to be available outside of the London MET Police
RDS	Serious mental illness (Depression, bipolar, schizophrenia)	Prison	Total: 14 items[1] Bipolar subscale: 5 items Depression subscale: 5 items Schizophrenia subscale: 5 items	≥2 on depression or schizophrenia subscales, or ≥3 on bipolar subscale	Questionnaire administered by staff	Administration time: 5 minutes Training: may be used by laypersons but reliability/validity are only assured if users receive extensive training	Unclear

Note.
[1] One item contributes to both the depression and bipolar subscales.

1
2

Table 17: QUADAS II quality assessment of studies included in the review of the most appropriate tools for the recognition of mental health problems.

Study ID	Index test	Risk of bias				Applicability concerns		
		Participant selection	Index test	Reference standard	Flow and timing	Participant selection	Index test	Reference standard
Baksheev 2012	BJMHS/BJMHS-R, Custody Risk Assessment Form	Unclear	High[a] Unclear[b]	High[a] Unclear[b]	Low	Low	Low	Low
Ford 2007	BJMHS/BJMHS-R, CMHS-M, CMHS-W, RDS	Low	Unclear	Low	High	Low	Low	Low
Ford 2009	CMHS-M, CMHS-W	High	Unclear	Low	High	Low	Low	Low
Harrison 2007	RDS	High	Unclear	Low	Low	Low	Low	Low
Louden 2013	BJMHS/BJMHS-R	Unclear	Low	Unclear	High	Low	Low	Low
McKinnon 2014	HELP-PC	Low	Unclear	Unclear	High	Low	Low	Low
McKinnon 2015	HELP-PC	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Sacks 2007a	CODSI	Unclear	Unclear	Low	High	Low	Unclear	Low
Sacks 2007b	CODSI	Unclear	Low	Low	High	Low	Unclear	Low
Steadman 2005	BJMHS/BJMHS-R	Unclear	Unclear	Low	High	Low	Low	Low
Steadman 2007	BJMHS/BJMHS-R	Unclear	Unclear	Low	High	Low	Low	Low
Teplin 1989a	RDS	Low	High	Unclear	Unclear	Low	Unclear	Low
Teplin 1989b	RDS	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low

Note. [a] BJMHS/BJMHS-R; [b] Custody Risk Assessment Form

5.2.111 Tools without acceptable sensitivity and specificity

Due to the number of identified tools and reported cut-off points, the GC agreed to only review tools and cut-off points with acceptable sensitivity and specificity, which was determine by a relatively conservative threshold of ≥0.70 for both values.

Therefore, evidence relating to the following tools was not considered by the GC: Brief Jail Mental Health Screen (BJMHS)/Brief Jail Mental Health Screen - Revised (BJMHS-R), Co-occurring Disorders Screening Instruments (CODSI) and Custody Risk Assessment Form. An overview of the studies examining these tools can be found in Table 18.

Table 18: Study information table for the review of the most appropriate tools for the recognition of mental health problems – studies not presented to the GC

	BJMHS/BJMHS-R	CODSI	Custody Risk Assessment Form
Total no. of studies (N)	5 (1422)	2 (280)	1 (150)
Study ID	(1) Baksheev 2012 (2) Ford 2007 (3) Louden 2013 (4) Steadman 2005 (5) Steadman 2007	(1) Sacks 2007a (2) Sacks 2007b	(1) Baksheev 2012
Study design	(1,2,3,4,5) cross-sectional study	(1,2) cross-sectional study	(1) cross-sectional study
Country	(1) Australia (2 – 5) USA	(1, 2) USA	(1) Australia
Target Condition(s)	(1, 4, 5) Serious mental illness (1, 3) Axis-I disorder (Exc. Substance misuse) (2) Affective disorder, (2) Anxiety disorder (2) Axis-I disorder (2) Axis-I or Axis-II disorder	(1, 2) General mental health (1, 2) Serious mental illness	(1) Serious mental illness (1) Axis-I disorder (Exc. Substance misuse)
Reference Standard(s)	(1 – 5) DSM-IV	(1, 2) DSM-IV	(1) DSM-IV
Setting	(1) Police custody (2, 4, 5) Reception into prison (3) Community	(1, 2) Subsequent time points in prison	(1) Police custody
Age (mean)	(1) 30 (2, 5) Not reported (3) 34 (4) 32	(1) Not reported (2) 35	(1) 30
Sex (% female)	(1) 9 (2, 3) 33 (4) 41 (5) 56	(1) 25 (2) 41	(1) 9
Ethnicity (% Caucasian)	(1) 81 (2) 43	(1) Not reported (2) 52	(1) 81

	BJMHS/BJMHS-R	CODSI	Custody Risk Assessment Form
	(3) 39 (4, 5) Not reported		
Note. N = total number of participants			

5.2.112 Depression

Three studies examined the sensitivity and specificity of recognition tools for depression (N = 1249): (Harrison & Rogers, 2007; McKinnon & Grubin, 2014; Teplin & Swartz, 1989).

An overview of the trials included in this review can be found in Table 19. Summary of findings can be found in Table 20. Summary ROC curves are in Appendix O.

Table 19: Study information table for the review of the most appropriate tools for the recognition of mental health problems – depression

	HELP-PC	RDS: Depression subscale
Total no. of studies (N)	1 (323)	2 (926)
Study ID	(1) McKinnon 2014	(1) Harrison 2007 (2) Teplin 1989a
Study design	(1) cross-sectional study	(1,2) cross-sectional study
Country	(1) UK	(1, 2) USA
Reference Standard(s)	(1) Unclear	(1) DSM-IV (2) DSM-III
Setting	(1) Police custody	(1) Subsequent time points in prison (2) Reception into prison
Age (mean)	(1) 32	(1) 34 (2) 25
Sex (% female)	(1) 10	(1, 2) 0
Ethnicity (% Caucasian)	(1) 57	(1) Not reported (2) 12
Note. N = total number of participants		

Table 20: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – depression

Tool	Cut-off	Total no. of studies (N)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (range)	NPV (range)	Quality[1]
HELP-PC	Not reported	1 (323)	0.75 (0.55, 0.89)	0.80 (0.75, 0.84)	0.26	0.97	Low
RDS: Depression subscale	2	2 (828)	0.86 (0.34,0.99)	0.77 (0.20,1.00)	0.20-0.71	0.96-1.00	Very low

.N = total number of participants; PPV = positive predictive value; NPV = negative predictive value
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.113 Bipolar disorder

One study examined the sensitivity and specificity of recognition tools for bipolar disorder (N = 728): (Teplin & Swartz, 1989).

An overview of this trial can be found in Table 21. Summary of findings can be found in Table 22. Summary ROC curves are in Appendix O.

Table 21: Study information table for the review of the most appropriate tools for the recognition of mental health problems – bipolar disorder

	RDS: Bipolar subscale
Total no. of studies (N)	1 (728)
Study ID	(1) Teplin 1989a
Study design	(1) cross-sectional study
Country	(1) USA
Reference Standard(s)	(1) DSM-III
Setting	(1) Reception into prison
Age (mean)	(1) 25
Sex (% female)	(1) 0
Ethnicity (% Caucasian)	(1)12
Note. N = total number of participants	

Table 22: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – bipolar disorder

Tool	Cut-off	Total no. of studies (N)	Sensitivity (95% CI)	Specificity (95% CI)	PPV	NPV	Quality[1]
RDS: Bipolar subscale	1	1 (728)	1.00 (0.86,1.00)	0.87 (0.84,0.89)	0.21	1.00	Low
	2	1 (728)	0.92 (0.73,0.99)	0.98 (0.97,0.99)	0.61	1.00	Low
	3	1 (728)	0.83 (0.63,0.95)	1.00 (0.99,1.00)	1.00	0.99	Low

.N = total number of participants; PPV = positive predictive value; NPV = negative predictive value
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.124 Affective disorder

One study examined the sensitivity and specificity of recognition tools for affective disorder (N = 302): (Ford et al., 2007).

An overview of this trial can be found in Table 23. Summary of findings can be found in Table 24. Summary ROC curves are in Appendix O.

Table 23: Study information table for the review of the most appropriate tools for the recognition of mental health problems – affective disorder

	CMHS-M	CMHS-W
Total no. of studies (N)	1 (302)	1 (302)
Study ID	(1) Ford 2007	(1) Ford 2007
Study design	(1) cross-sectional study	(1) cross-sectional study
Country	(1) USA	(1) USA
Reference Standard(s)	(1) DSM-IV	(1) DSM-IV
Setting	(1) Reception into prison	(1) Reception into prison
Age (mean)	(1) Not reported	(1) Not reported
Sex (% female)	(1) 33	(1) 33
Ethnicity (% Caucasian)	(1) 43	(1) 43
Note. N = total number of participants		

Table 24: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – affective disorder

Tool	Cut-off	Total no. of studies (N)	Sensitivity (95%CI)	Specificity (95%CI)	PPV	NPV	Quality[1]
CMHS-M (All men)	7	1 (201)	0.83 (0.63,0.95)	0.73 (0.66,0.79)	0.30	0.97	Low
CMHS-M (Caucasian men)	7	1 (98)	0.94 (0.73,1.00)	0.78 (0.67,0.86)	0.47	0.98	Low
CMHS-M (Black men)	7	1 (69)	1.00 (0.29,1.00)	0.70 (0.57,0.80)	0.13	1.00	Low
CMHS-W	5	1 (100)	0.73 (0.54,0.87)	0.70 (0.58,0.81)	0.55	0.84	Low

.N = total number of participants; PPV = positive predictive value; NPV = negative predictive value
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.175 Learning disabilities

One study examined the sensitivity and specificity of recognition tools for learning disabilities (N = 351): (McKinnon et al., 2015).

An overview of this trial can be found in

Table 25. Summary of findings can be found in Table 26. Summary ROC curves are in Appendix O.

Table 25: Study information table for the review of the most appropriate tools for the recognition of mental health problems – learning disabilities

	HELP-PC
Total no. of studies (N)	1 (351)
Study ID	(1) McKinnon 2015
Study design	(1) cross-sectional study
Country	(1) UK
Reference Standard(s)	(1) Unclear
Setting	(1) Police custody
Age (mean)	(1) Not reported
Sex (% female)	(1) Not reported
Ethnicity (% Caucasian)	(1) Not reported
Note. N = total number of participants	

Table 26: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – learning disabilities

Tool	Cut-off	Total no. of studies (N)	Sensitivity	Specificity	PPV	NPV	Quality[1]
HELP-PC	1	1 (351)	0.83 (0.36,1.00)	0.88 (0.84,0.91)	0.11	1.00	Low

.N = total number of participants; PPV = positive predictive value; NPV = negative predictive value
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.116 Schizophrenia

One study examined the sensitivity and specificity of recognition tools for schizophrenia (N = 728): (Teplin & Swartz, 1989).

An overview of this trial can be found in Table 27. Summary of findings can be found in Table 28. Summary ROC curves are in Appendix O.

Table 27: Study information table for the review of the most appropriate tools for the recognition of mental health problems – schizophrenia

	RDS: Schizophrenia subscale
Total no. of studies (N)	1 (728)
Study ID	(1) Teplin 1989a
Study design	(1) cross-sectional study
Country	(1) USA
Reference Standard(s)	(1) DSM-III
Setting	(1) Reception into prison
Age (mean)	(1) 25
Sex (% female)	(1) 0
Ethnicity (% Caucasian)	(1)12
Note. N = total number of participants	

Table 28: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – schizophrenia

Tool	Cut-off	Total no. of studies (N)	Sensitivity (95%CI)	Specificity (95%CI)	PPV	NPV	Quality[1]
RDS: Schizophrenia subscale	1	1 (728)	0.88 (0.68,0.97)	0.96 (0.94,0.97)	0.43	1.00	Low

.N = total number of participants; PPV = positive predictive value; NPV = negative predictive value
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.117 Psychosis

One study examined the sensitivity and specificity of recognition tools for psychosis (N = 323): (McKinnon & Grubin, 2014).

An overview of this trial can be found in Table 29. Summary of findings can be found in Table 30. Summary ROC curve is in Appendix O.

Table 29: Study information table for the review of the most appropriate tools for the recognition of mental health problems – psychosis

	HELP-PC
Total no. of studies (N)	1 (323)
Study ID	(1) McKinnon 2014
Study design	(1) cross-sectional study
Country	(1) UK
Reference Standard(s)	(1) Unclear
Setting	(1) Police custody
Age (mean)	(1) 32
Sex (% female)	(1) 10
Ethnicity (% Caucasian)	(1) 57
Note. N = total number of participants	

Table 30: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – psychosis

Tool	Cut-off	Total no. of studies (N)	Sensitivity (95%CI)	Specificity (95%CI)	PPV	NPV	Quality[1]
HELP-PC	Not reported	1 (323)	0.93 (0.76,0.99)	0.81 (0.76,0.86)	0.32	0.99	Low

.N = total number of participants; PPV = positive predictive value; NPV = negative predictive value
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.118 Axis-I or Axis-II disorder

Two studies examined the sensitivity and specificity of recognition tools for Axis-I or Axis-II disorder (N = 508): (Ford et al., 2007; Ford et al., 2009).

An overview of this trial can be found in Table 31. Summary of findings can be found in Table 32. Summary ROC curves are in Appendix O.

Table 31: Study information table for the review of the most appropriate tools for the recognition of mental health problems – Axis-I or Axis-II disorder

	CMHS-M	CMHS-W	RDS
Total no. of studies (N)	2 (508)	1 (206)	1 (302)
Study ID	(1) Ford 2007 (2) Ford 2009	(1) Ford 2009	(1) Ford 2007
Study design	(1,2) cross-sectional study	(1) cross-sectional study	(1) cross-sectional study
Country	(1, 2) USA	(1) USA	(1) USA
Reference Standard(s)	(1, 2) DSM-IV	(1) DSM-IV	(1) DSM-IV
Setting	(1, 2) Reception into prison	(1) Reception into prison	(1) Reception into prison
Age (mean)	(1, 2) Not reported	(1) Not reported	(1) Not reported
Sex (% female)	(1) 33 (2) 49	(1) 49	(1) 33
Ethnicity (% Caucasian)	(1) 43 (2) Not reported	(1) Not reported	(1) 43
Note. N = total number of participants			

Table 32: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – Axis-I or Axis-II disorder

Tool	Cut-off	Target condition(s)	Total no. of studies (N)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (range)	NPV (range)	Quality[1]
CMHS-M (All men)	5	Axis-I or Axis-II disorder, excluding ASPD	1 (106)	0.80 (CI not reported)	0.78 (CI not reported)	0.74	0.84	Very low
	6	Axis-I or Axis-II disorder, excluding ASPD	2 (307)	0.69 (0.17,0.96)	0.76 (0.26,0.98)	0.60-0.76	0.78-0.85	Very low
CMHS-M (Caucasian men)	6	Axis-I or Axis-II disorder, excluding ASPD	1 (97)	0.82 (0.65,0.93)	0.78 (0.66,0.87)	0.66	0.89	Very low
CMHS-M (Black men)	6	Axis-I or Axis-II disorder, excluding ASPD	1 (69)	0.80 (0.56,0.94)	0.71 (0.57,0.83)	0.53	0.90	Very low

Tool	Cut-off	Target condition(s)	Total no. of studies (N)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (range)	NPV (range)	Quality[1]
CMHS-W	4	Axis-I or Axis-II disorder	1 (100)	0.74 (0.61,0.84)	0.84 (0.67,0.95)	0.91	0.61	Low
CMHS-W	4	Axis-I or Axis-II disorder, excluding ASPD	1 (100)	0.74 (0.61,0.84)	0.72 (0.55,0.85)	0.81	0.64	Low
RDS	3	Axis-I or Axis-II disorder, excluding ASPD	1 (27)	0.73 (0.45,0.92)	0.83 (0.52,0.98)	0.85	0.71	Low

.N = total number of participants; PPV = positive predictive value; NPV = negative predictive value
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.119 Current prison reception health screen

There were no studies that met our inclusion criteria that examined the prison reception health screen developed by Grubin et al. (2002). As this tool has been widely adopted in UK prisons, the GC decided that it was important to review evidence regarding its sensitivity and specificity to provide some context in which to interpret the performance of the included recognition tools.

Therefore, two studies identified by the search strategy described above, but that did not meet our inclusion criteria, were presented to the GC (N = 1442): (Evans et al., 2010; Grubin et al., 2002). These studies were both initially excluded because they did not use an appropriate reference standard; further, Evans et al. (2010) weighted sensitivity and specificity and therefore the results could not be included in a pooled analysis.

An overview of these trials can be found in Table 33. Summary of findings can be found in Table 34. The summary ROC curve is in Appendix O.

Table 33: Study information table for the review of the most appropriate tools for the recognition of mental health problems – current prison reception health screen

	Prison reception health screen
Total no. of studies (N)	2 (680)
Study ID	(1) Evans 2010 (2) Grubin 2002
Study design	(1,2) cross-sectional study
Country	(1) New Zealand (2) UK
Reference Standard(s)	(1) MINI (2) SADS-L
Setting	(1, 2) Reception into prison
Age (mean)	(1, 2) Not reported
Sex (% female)	(1) 0 (2) 20
Ethnicity (% Caucasian)	(1, 2) Not reported
Note. N = total number of participants	

Table 34: Summary of findings table for the review of the most appropriate tools for the recognition of mental health problems – current prison reception health screen

Tool	Cut-off	Total no. of studies (N)	Sensitivity (range)	Specificity (range)	PPV	NPV	Quality[2]
Prison reception health screen	1	2 (680)	0.42-0.97	0.75-0.83	0.60[1]	0.99[1]	Low

Note. N = total number of participants; PPV = positive predictive value; NPV = negative predictive value;
[1]It was only possible to extract PPV and NPV from one of the studies
[2]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II domains.

5.2.2 Economic Evidence

No economic evidence on the tools for the recognition of mental health problems for adults who are in contact with the criminal justice system was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

5.2.3 Clinical evidence statements

5.2.3.1 Depression

There was low quality evidence from one study (n=323) that the HELP-PC (cut-off not reported) has acceptable diagnostic accuracy with sensitivity of 75% (95%CI 55-89%) and specificity of 80% (95%CI: 75-84%) for the recognition of depression.

There was very low quality evidence from two studies (n=828) that the RDS: Depression Subscale with a cut-off of 2 has acceptable diagnostic accuracy with sensitivity of 86% (95%CI: 34-99%) and specificity of 77% (95%CI: 2-100%) for the recognition of depression.

5.2.32 Bipolar disorder

There was low quality evidence from one study (n=728) that, for the recognition of bipolar disorder, the RDS: Bipolar Subscale

- with a cut-off of 1 has acceptable diagnostic accuracy with sensitivity of 100% (95% CI: 86-100%) and specificity of 87% (95% CI: 84-89%)
- with a cut-off of 2 has acceptable diagnostic accuracy with sensitivity of 92% (95% CI: 73-99%) and specificity of 98% (95%CI: 97-99%).
- with a cut-off of 3 has acceptable diagnostic accuracy with sensitivity of 83% (95% CI: 63-95%) and specificity of 100% (95%CI: 99-100%)

5.2.33 Affective disorder

There was low quality evidence from one study (n=201) that the CMHS-M with a cut-off of 7 has acceptable diagnostic accuracy with sensitivity of 83% (95%CI: 63-95%) and specificity of 73% (95%CI: 66-79%) for the recognition of affective disorders. The subgroup analyses indicated that the tool can detect affective disorders among Caucasian men (n=98) with sensitivity of 94% (95%CI: 73-100%) and specificity of 78% (95%CI: 67-86%) and among Black men (n=69) with sensitivity of 100% (95%CI 29-100%) and specificity of 70% (95% CI: 57-80%).

There was low quality evidence from one study (n=100) that the CMHS-W with a cut-off of 5 has acceptable diagnostic accuracy with sensitivity of 73% (95%CI: 54-87%) and specificity of 70% (95%CI: 58-81%) for the recognition of affective disorders.

5.2.34 Learning disabilities

There was low quality evidence from one study (n=351) that the HELP-PC with a cut-off of 1 has acceptable diagnostic accuracy with sensitivity of 83% (95%CI: 36-100%) and specificity of 88% (95%CI: 84-91%) for the recognition of learning disabilities.

5.2.35 Schizophrenia

There was low quality evidence from one study (n=728) that the RDS: Schizophrenia Subscale with a cut off of 1 has acceptable diagnostic accuracy with sensitivity of 88% (95%CI: 68-97%) and specificity of 96% (95%CI: 94-97%) for the recognition of schizophrenia.

5.2.36 Psychosis

There was low quality evidence from one study (n=323) that the HELP-PC (cut-off not reported) has acceptable diagnostic accuracy with sensitivity of 93% (95%CI: 76-99%) and specificity of 81% (95%CI: 76-86%) for the recognition of psychosis.

5.2.37 Axis-I or Axis-II disorder

There was very low quality evidence from two studies (n=307) that the CMHS-M with a cut-off of 6 has acceptable diagnostic accuracy with sensitivity of 69% (95%CI: 17-96%) and specificity of 76% (95%CI: 26-98%) for the recognition of Axis-I or Axis-II disorders, excluding Anti-Social Personality Disorder (ASPD). The subgroup analyses indicated that the tool can detect the disorders among Caucasian men with sensitivity of 82% (95%CI: 65-93%)

and specificity of 78% (95%CI: 66-87%) whereas among Black men with sensitivity of 80% (95%CI 56-94%) and specificity of 71% (95% CI: 57-83%).

There was low quality evidence from one study (n=100) that the CMHS-W with a cut-off of 4 has acceptable diagnostic accuracy with sensitivity of 74% (95%CI: 61-84%) and specificity of 84% (95%CI: 67-95%) for the recognition of Axis-I or Axis-II disorders.

There was low quality evidence from one study (n=100) that the CMHS-W with a cut-off of 4 has acceptable diagnostic accuracy with sensitivity of 74% (95%CI: 61-84%) and specificity of 72% (95%CI: 55-85%) for the recognition of Axis-I or Axis-II disorders, excluding ASPD.

There was low quality evidence from one study (n=27) that the RDS with a cut off of 3 has acceptable diagnostic accuracy with sensitivity of 73% (95%CI: 45-92%) and specificity of 83% (95%CI: 52-98%) for the recognition of Axis-I or Axis-II disorders, excluding ASPD.

5.2.3.8 Prison reception health screen

There was low quality evidence from two studies (n=680) that the current prison reception health screen with a cut-off of 1 has acceptable diagnostic accuracy with sensitivity of 42-97% and specificity of 75-83% for the recognition of mental health disorders.

5.2.4 Economic evidence statements

No economic evidence on tools for the recognition of mental health problems for adults who are in contact with the criminal justice system is available.

5.3 Review question: What are the most appropriate tools to support or assist in the assessment of mental health problems, or what modifications are needed to assessment tools recommended in existing NICE guidance, for adults:

- in contact with the police?
- in police custody?
- for the court process?
- at reception into prison?
- at subsequent time points in prison?
- in the community (serving a community sentence, released from prison on licence or released from prison and in contact with a community rehabilitation company [CRC] or the probation service)?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 35. A complete list of review questions and full review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 35: Clinical review protocol summary for the review of the most appropriate tools for the assessment of mental health problems

Component	Description
Population	Adults (aged 18 and over) with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Index test	Any formal recognition and assessment tool considered appropriate and suitable for use
Reference standard	Diagnosis Statistical Manual (DSM) or International Classification of Diseases (ICD) diagnosis
Outcomes	**Critical**: Sensitivity: the proportion of true positives of all cases diagnosed with the target condition in the population Specificity: the proportion of true negatives of all cases not-diagnosed with the target condition in the population Reliability (for instance, inter-rater or test-retest reliability or internal consistency) Validity (for instance, criterion or construct validity) **Important:** Feasibility for use – time taken, burden on user or individual
Study design	Systematic reviews of diagnostic test accuracy studies, diagnostic cross-sectional studies

5.3.41 Clinical evidence

There was only one study that that provided sufficient data to be included in the evidence synthesis for this review question. (Mokros et al., 2012). The study was published in a peer-reviewed journal and reported on the Severe Sexual Sadism Scale (SSSS; Nitschke et al., 2009).

The SSSS did not have acceptable sensitivity and specificity (of 70% or greater); therefore, the above study was not considered by the GC. An overview of this study can be found in Table 36.

Table 36: Study information table for the review of the most appropriate tools for the assessment of mental health problems

	SSSS
Total no. of studies (N)	1 (105)
Study ID	(1) Mokros 2012
Study design	(1) cross-sectional study
Country	(1) Austria
Target Condition(s)	(1) Sexual Sadism
Reference Standard(s)	(1) DSM-IV-TR
Setting	(1) Prison
Age (mean)	(1) 33
Sex (% female)	(1) 0
Ethnicity (% Caucasian)	(1) Not reported
Note. N = total number of participants	

5.3.2 Economic evidence

No economic evidence on the tools for the assessment of mental health problems for adults who are in contact with the criminal justice system was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

5.3.3 Clinical evidence statements

There was no clinical evidence considered by the GC for this review question as the only study that met the inclusion criteria did not report any evidence for a tool with acceptable sensitivity and specificity. However, the group decided this was an important issue and therefore agreed that this question should be considered as part of the nominal group technique used to address review question 2.4.

5.3.4 Economic evidence statements

No economic evidence on the tools for the assessment of mental health problems for adults who are in contact with the criminal justice system is available.

5.4 Recommendations and link to evidence

Recommendations	**5. Be vigilant for the possibility of unidentified or emerging mental health problems in people in contact with the criminal justice system, and review available records for any indications of a mental health problem.** **6. Ensure all staff working in criminal justice settings are aware of the potential impact on a person's mental**

health of being in contact with the criminal justice system

First-stage health assessment at reception into prison

This subsection covering what happens when a person first arrives into prison is taken from the NICE guideline on physical health in prisons. It does not apply to other criminal justice system settings.

This material, was developed jointly by NICE's physical health in prisons and mental health in the criminal justice system committees has already been consulted on as part of the development of the physical health in prisons guideline. It is therefore not open to consultation.

The final, amended version of this section will appear for the first time when the physical health in prisons guideline publishes in November 2016. This amended version will also appear in the final version of mental health in the criminal justice system guideline when it is publishes in October 2017.

7. A healthcare professional (or trained healthcare assistant under the supervision of a registered nurse) should carry out a health assessment for every person on their first reception into prison. This should be done before the person is allocated to their cell. It should include identifying:

- **any issues that may affect the person's immediate health and safety before the second-stage health assessment**
- **priority health needs to be addressed at the next clinical opportunity.**

8. The first-stage health assessment should include the questions and actions in table 1. It should cover:

- **physical health**
- **alcohol use**
- **drug use**
- **mental health**
- **self-harm and suicide.**

9. Take into account any communication needs or difficulties the person has, and follow the principles in NICE's guideline on patient experience in adult NHS services.

Table 1 Questions for first-stage prison health assessment

Topic questions	Actions
11 Status	
Has the person been charged with murder or manslaughter?	Yes: refer for urgent mental health assessment by the prison mental health in-reach team.

	Ensure that the person is seen by the GP while they are in reception. No: record no action required.
22 Physical health	
32.1 Prescribed medicines	
Is the person taking any prescribed medicines, including preparations such as creams or drops, and if so: • what are they? • what are they for? • how do they take them?	Yes: make a note of any current medicines being taken and generate a medicine chart. Refer the person to the GP for appropriate medicines to be prescribed and continued. If medicines are being taken check that the next dose has been provided. No: record no action required.
42.2 Physical injuries	
Has the person received any physical injuries over the past few days, and if so: • what were they? • how were they treated?	Yes: assess severity of injury, any treatment received and record any head, abdominal injuries or fractures. Refer the person to the GP at reception. In very severe cases, or after GP assessment, the person may need to be transferred to an external hospital. Liaise with prison staff to transfer the person to the hospital emergency department by ambulance. Document any bruises or lacerations observed. If the person has made any allegations of assault, record negative observations as well (for example, no physical evidence of injury). No: record no action required.
52.3 Head injuries or loss of consciousness	
Has the person ever suffered a head injury or lost consciousness, and if so: • how many times has this happened? • have they ever been unconscious for more than 20 minutes? • do they have any problems with their memory or concentration?	Yes: refer the person to the GP at reception. No: record no action required.
62.4 Other physical health conditions	
Does the person have any of the following: • allergies, asthma, diabetes, epilepsy or fits • chest pain, heart disease	Ask about each illness listed. Yes: make short notes on any details of the person's condition or management. For example, 'Asthma – on Ventolin one puff daily'.

	• tuberculosis, sickle cell disease • hepatitis B or C virus, HIV, other sexually transmitted infections • learning disabilities • neurodevelopmental disorders • physical disabilities?	Make appointments with relevant clinics or specialist nurses if specific needs have been identified. No: record no action required.
	2.5 Are there any other physical health problems the person is aware of, that have not been reported?	Yes: record the details and check with the person that no other physical health complaint has been overlooked. No: record no action required.
	2.6 Are there any other concerns about the person's physical health?	Make a note of any other concerns about physical health. This should include any health-related observations about the person's physical appearance (for example, weight, pallor, jaundice, gait). As with recent injuries, both negative and positive signs are relevant. Yes: refer the person to the GP at reception. No: note 'Nil'.
	7 2.7 Additional questions for women	
	Ask the woman if she has reason to think she is pregnant.	Yes: refer the person to the GP at reception and to a midwife. No: record response.
	Ask if she would like a pregnancy test.	Yes: if requested, provide a pregnancy test. Record the outcome and if positive make an appointment for the person to see the GP. No: record response.
	8 2.8 Independent living and diet	
	Ask the person if they need help to live independently.	Yes: note any needs. Liaise with the prison disability lead in reception about: the location of the person's cell further disability assessments the prison may need to carry out. No: record response.
	Ask if they use any equipment or aids (for example, walking stick, hearing aid, glasses).	Yes: remind prison staff that all special equipment and aids the person uses should follow them from reception to their cell. No: record response.
	Ask if they need a special medical diet.	Yes: note the medical diet the person needs and send a request to catering. No: record response.
	9 2.9 Past or future medical appointments	
	Ask the person if they have seen a doctor or other healthcare	Yes: note details of any recent medical contact. Arrange a

professional in the past few months, and if so what this was for.	contact letter to get further information from the person's doctor. Note any ongoing treatment the person needs and make appointments with relevant clinics, specialist nurses, GP or other healthcare staff. No: record no action required.
Ask if they have any outstanding medical appointments, who they are with, and the dates.	Yes: note future appointment dates. Ask healthcare administrative staff to manage these appointments or arrange for new dates and referral letters to be sent if the person's current hospital is out of the local area. No: record no action required.
03 Alcohol and drug use	
3.1 Ask the person if they drink alcohol, and if so: • how much they normally drink • how much they drank in the week before coming into custody.	Urgently refer the person to the GP at reception or the drug services team if: • they drink more than 15 units of alcohol daily or • they are showing signs of withdrawal. No: record response.
13.2 Type and frequency of drug use	
Ask the person if they have used drugs in the last month. If yes, ask about frequency of use, and last use of, for example: • heroin • methadone • benzodiazepines • amphetamine • cocaine or crack • novel psychoactive substances.	Ask about use of different drugs including those listed. Yes: refer the person to drug services if there are concerns about their immediate clinical management and they need immediate support. Take into account whether: • they have taken drugs intravenously • they have a positive urine test for drugs • their answers suggest that they use drugs more than once a week. Refer the person to the GP at reception if there are any physical health concerns. No: record response.
23.3 Intravenous drugs	
Ask the person if they have taken any drugs intravenously.	Yes: check injection sites. Refer the person to drug services if there are concerns about their immediate clinical management and they need immediate support. Refer the person to the GP at reception if there are any physical health concerns. No: record response.

33.4 Prescription drugs	
Ask the person if they have used prescription or over-the-counter medicines in the past month that: • were not prescribed or recommended for them, or • for purposes or at doses that were not prescribed. If yes, ask what this medicine was and how they used it (frequency and dose).	Yes: refer the person to drug services if there are concerns about their immediate clinical management and they need immediate support. Refer the person to the GP at reception if there are any physical health concerns. No: record response.
44 Mental health	
54.1 Previous contact with mental health services	
Ask the person if they have ever seen a health professional or service about a mental health problem (including a psychiatrist, GP, psychologist, counsellor, community mental health team or learning disability team). If yes, ask: • who they saw • the nature of the problem.	Yes: consider referring the person for mental health assessment by the prison mental health in-reach team) if they have received care for mental health problems. Refer the person to the GP at reception. If the person has been in contact with learning disability services refer them to the GP in reception No: record response.
Ask the person if they have ever been admitted to a psychiatric hospital. If yes, ask them: • the date of their most recent discharge • the name of the hospital • the name of their consultant.	Yes: refer the person for mental health assessment by the prison mental health in-reach team if they have received inpatient care for mental health problems. Refer the person to the GP at reception. No: record response.
64.2 Medicine for mental health problems	
• Ask the person if they have ever been prescribed medicine for any mental health problems. If yes, ask: • what the medicine was • when they received it • what the current dose is (if they are still taking it).	Yes: consider referring the person for mental health assessment if they have received medicine for mental health problems. Refer the person to the GP at reception. No: record response.
75 Self-harm and suicide	
85.1 History of self-harm or suicide attempts	
Ask the person if they have ever tried to harm themselves. If yes, ask: • whether this was inside or outside prison • what the most recent incident was • what the most serious incident was.	Yes: consider referring the person for a mental health assessment if they have ever tried to harm themselves. No: record response.
Ask the person if they: • have a history of previous	Yes: refer the person for an urgent mental health

	suicide attempts • are currently thinking about or planning to harm themselves or attempt suicide.	assessment. Open an Assessment, Care in Custody and Teamwork (ACCT) plan if there are: • serious concerns raised in response to questions about self-harm, including thoughts, intentions, or plans • a history of previous suicide attempts. Refer the person to the GP at reception. No: record response.

Identification throughout the care pathway and second stage health assessment in prisons)

10. **Consider using the Correctional Mental Health Screen for Men (CMHS-M) or Women (CMHS-W) to identify possible mental health problems if[g]:**
    - **the person's history, presentation or behaviour suggest they may have a mental health problem**
    - **the person's responses to the first-stage health assessment suggest they may have a mental health problem**
    - **the person has a chronic physical health problem with associated functional impairment**
    - **concerns have been raised by other agencies about the person's abilities to participate in the criminal justice process[g].**

11. **When using the CMHS-M or CMHS-W with a transgender person, use the measure that is in line with their preferred gender identity.**

12. **If a man scores 6 or more on the CMHS-M, or a woman scores 4 or more on the CMHS-W, or there is other evidence supporting the likelihood of mental health problems, practitioners should:**
    - **conduct a further assessment if they are competent to perform assessments of mental health problems, or**
    - **refer the person to an appropriately trained professional for further assessment if they are not competent to perform such assessments themselves[g].**

[g]This recommendation applies both throughout the care pathway and to second stage health assessment in prisons. Consultation on this recommendation (in the context of second stage health assessment in prisons) has already happened as part of the consultation on the physical health in prisons guideline.

g This recommendation, applies both throughout the care pathway and to second stage health assessment in prisons. Consultation on this recommendation (in the context of second stage health assessment in prisons) has already happened as part of the consultation on the physical health in prisons guideline.

Relative values of different outcomes	When assessing tools for recognition and assessment of mental health problems the GC agreed that preference should be given to tools that could identify or be helpful in assessing a range of mental health problems, as opposed to recommending the use of multiple tools which could detect only single disorders. Sensitivity and specificity were selected as the critical outcomes for case recognition tools and reliability and validity for assessment tools
Trade-off between clinical benefits and harms	**Case recognition** When considering whether or not to recommend a case identification tool, the GC were mindful of the benefits associated with the identification of mental health problems in the criminal justice population and the prison population, in particular, as the prevalence of mental disorders is known to be significantly higher in this population They were also aware of the under-recognition of mental health populations in this area and therefore the sub-optimal treatment received. The GC also considered the potential harms (e.g. increased anxiety or stigma) or inappropriate use of resources (e.g. unnecessary treatment) that may arise from false positives. The GC therefore did not consider scales that did not meet a pre-determined level of sensitivity and specificity. In addition to the properties of particular scales the GC were also aware that initial screening or case recognition may be undertaken by staff with limited experience and skills in dealing with mental health problems. This meant that the GC had to identify questions or measures that could be delivered and interpreted by staff with this level of experience. The GC used informal consensus methods to inform any changes to the initial prison screen. A number of instruments were identified in low quality studies which when considered for single disorders (for example schizophrenia or depression) suggested that they had reasonable sensitivity. The only instrument that the GC identified which covered the full range of mental disorders only was the CMHS-M/CMHS-W which also had good psychometric properties. The structure of the tool did not support its use in the initial prison reception assessment but did support its use as a case identification tool in a second stage assessment in the prison system or for use a case identification tool in other areas of the criminal justice system. **Assessment** For assessment the GC were concerned with tools that improved the performance of the overall assessment process (e.g. more accurate diagnosis) and did not prove over burdensome for the individual being assessed or the person doing the assessment. Only one tool was identified (the SSSS) which did not have adequate psychometric properties and was considered by the GC not sufficient to support a recommendation.
Trade-off between net health benefits and resource use	**Case recognition** The GC were aware that all prisoners on first reception to prison are given an initial, brief health assessment which is expected to cover all physical and mental health problems with a focus on immediate management of acute problems and the identification of any associated risks. The intention of the initial assessment in prison is also to identify people who would need further assessment. The GC were also aware of current practice in prisons and so developed simple identification criteria which were compatible with current procedures and did not demand significant additional time or training and thereby had limited impact on costs. For non-prison populations the GC were also mindful of the time and skills required (and potentially associated costs) of any

	screening instruments. The choice of instruments was therefore guided by a set of principles which focused on brief, copy-free instruments which required limited training to deliver and score. The GC agreed that the use of a recognition tool (such as the CMHS-M/CMHS-W) which could be administered by a non-expert in five minutes or less would be the most effective way to limit the impact of this assessment on resources. The CMHS-M/W has good sensitivity when compared with standard care. This would result in a significant reduction in the rate of false negatives. Assuming similar specificity rates between CMHS-M/W and standard care, there is a clear cost advantage of using this tool given that it takes only 5 minutes to administer and reduces the number of false negatives by approximately 200 per 1,000 prisoners screened. The GC was aware of a wide range of alternative methods used in the criminal justice system and considered that the addition of this measure would impose limited additional cost burden on the system, and, given the clinical evidence, may very likely produce better outcomes. Its use as a) a further case identification method in the prison service and b) as a primary case recognition tool in other parts of the criminal justice system was supported. **Assessment** The GC identified no tools which had sufficient validity or reliability to support a recommendation.
Quality of evidence	The quality of the evidence ranged from moderate to low. The most common reasons that studies were marked down in terms of quality were that the flow and timing of the study, the conduct or interpretation of the index test and the relevance of the population included. There was very low or low quality evidence from two studies that the CMHS-M/CMHS-W had good sensitivity such that they were preferred as recognition tools by the GC. The RDS performed well psychometrically but the GC were informed by the developer of the RDS that the tool was validated against an outdated standard (the Diagnostic Interview Schedule) and the decision was therefore made to recommend only the CMHS-M/CMHS-W. Evidence for other 'single disorder' case recognition tools was essentially confined to single studies of low or very low quality. Given the GC's preference for a multi-disorder tool the GC agreed that there was insufficient evidence to recommend an alternative to, the current prison reception health screen. The GC agreed that using an adapted version of this assessment for the reception screen should be adopted which included additional items including those learning disability and changes to the level of alcohol consumption required to trigger a further assessment. These decision were informed by their expert knowledge and experience. The GC agreed that the Correctional Mental Health Screen for Men (CMHS-M) or Women (CMHS-W) should form part of the second stage of prison health screening and be used as a case identification tool in other parts of the criminal justice system. The evidence on sensitivity and specificity for the instrument met the GC's predetermined criteria.
Other considerations	The GC used informal consensus drawing on their knowledge and expertise to suggest amendments to the current prison reception health screen in the following areas: drugs and alcohol use (including that the threshold of 20 units per day for urgent referral regarding alcohol withdrawal be lowered to 15 units in line with NICE CG 115), contact with previous mental health services, self-harm and suicide, learning disabilities, assessor's impression of the service user. These amendments were informed by a review of relevant NICE guideline, for example the current prison recommendation for drug and alcohol use was not in agreement with current NICE guidance or the need for

	increasing awareness on the part of prison staff of the mental health needs or prisoners with learning disabilities. The GC, drawing on their expert knowledge and experience were also aware that it may not always be possible to use a formal measure, however brief, and so developed a recommendation by informal consensus on the need for staff to be vigilant for possible mental disorder. They were also aware of how communication difficulties could mask the identification of mental health problems and also developed a recommendation of taking these into account when considering the presence of a mental disorder. The GC were aware of the particular difficulties faced by transgender people in prison and the appropriate identification of mental health problems in this group. Therefore they decided, based on their knowledge and experience to recommend that the choice of which CMHS scale is used should be determined by the gender that the individual identifies with. The GC were aware of the high level of co-morbidity in the criminal justice population and, in particular, the challenges in identifying mental health problems in individuals with acquired cognitive impairment and neuro-developmental disorders. Given the absence of specific evidence on case identification tools in these populations the GC therefore decided to make a research recommendation.

5.4.2 Research recommendations

2. What are the reliable and valid tools to identify cognitive impairment among people in contact with the criminal justice system (focusing on people with trauma, neurodevelopmental disorders and acquired cognitive impairment as well as veterans and older people)?

There is limited evidence that interventions can reduce the cognitive or functional impairments associated with acquired cognitive impairment. Acquired cognitive impairment is common in criminal justice population. Moreover, people with acquired cognitive impairment have high risk of self-harm. Acquired cognitive impairment may arise as result of a traumatic brain injury, or a stroke. Experts in this area have suggested that early identification of deficits and prompt management strategies could be important in ameliorating the long-term impact of acquired cognitive impairment. However, there is lack of evidence on reliable and valid case identification tools and methods. It is important that research is developed to assist the staff in criminal justice pathway to facilitate identification of acquired cognitive impairment and support better understanding and management of acquired cognitive impairment.

5.5 Review question: What are the most appropriate tools to support or assist in risk assessment, for adults with mental health problems:

- in police custody?
- for the court process?
- at reception into prison?
- at subsequent time points in prison?
- in the community (serving a community sentence, released from prison on licence or released from prison and in contact with a community rehabilitation company [CRC] or the probation service)?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 37. A complete list of review questions and full review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 37: Clinical review protocol summary for the review of the most appropriate tools to support or assist in risk assessment for adults with mental health problems

Component	Description
Population	Adults (aged 18 and over) with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Intervention(s)	Any formal recognition and assessment tool considered appropriate and suitable for use by the guideline committee
Comparison	Reference standard
Outcomes	**Critical** Offending (including sexual offences), self-harm, attempted suicide and completed suicide. Reliability (for instance, inter-rater or test-retest reliability or internal consistency); Validity (for instance, criterion or construct validity) **Important** Practicality/Feasibility for use in routine care
Study design	Systematic reviews of risk assessment studies, diagnostic cross-sectional studies, cohort studies or case-control studies

5.5.1 Clinical evidence

For this review question the GC agreed that studies would be only be included if they examined predictive validity against a behavioural outcome (i.e. not simply measuring ability to predict risk level as assigned by other risk assessment tools). The GC also agreed to exclude studies if they only assessed recidivism for violent offending, general offending, or driving while intoxicated. This decision was made as such behaviours may not be linked to mental health problems and therefore would be outside of the scope of this guideline. Further, the decision was made to only include studies examining risk for sexual reoffending where ≥80% of the sample had a paraphilia to ensure offending behaviour was associated with a mental health problem.

The literature search for review questions 2.1 – 2.3 yielded 8948 articles overall. Scanning titles or abstracts identified 954 articles potentially relevant to the above review questions. After further inspection of the full articles, 926 studies did not meet one or more of the eligibility criteria. An additional 7 studies forwarded by stakeholders, 3 studies identified by handsearching, and 1 study identified by another literature search for this guideline also did not meet the inclusion criteria. This resulted in 17 articles representing 18 studies that were included for review question 2.3. An additional study identified by hand-searching -Wichmann 2000(Wichmann et al., 2000)also met the inclusion criteria resulting in a total of 18 articles, representing 19 studies, that provided sufficient data to be included in the evidence synthesis for review question 2.3:(Beggs & Grace, 2010; Frottier et al., 2009; Hanson et al., 2010; Hanson & Thornton, 2000; Helmus et al., 2015; Horton et al., 2014; Ivanoff & Jang, 1991; Kingston et al., 2010; Naud & Daigle, 2010; Perry & Gilbody, 2009; Perry & Olason, 2009; Seto et al., 2004; Sjostedt & Grann, 2002b; Spurgeon et al., 2000; Thomas et al., 2014; Wakeling et al., 2011a; Wichmann et al., 2000).

All but one of the studies were published in peer-reviewed journals between 1991 and 2015; the remaining study (Wichmann et al., 2000) was published by Canada's correctional service in 2000. These studies report on tools which can be categorised as assessing risk of sexual reoffending, self-harm and/or suicide and relapse into substance misuse. Characteristics of these risk assessment tools can be found in Table 38.

Of the eligible studies reporting on risk of self-harm or suicidal behaviour:

- 3 each reported on the Beck Hopelessness Scale (BHS); (Ivanoff & Jang, 1991; Perry & Gilbody, 2009) and the Suicide and Self-harm Concerns about Offenders in Prison Environment (SCOPE); (Perry & Gilbody, 2009; Perry & Olason, 2009)
- 2 reported on the Suicide Probability Scale (SPS; (Naud & Daigle, 2010; Naud & Daigle, 2013))
- 1 each reported on the Prison Screening Questionnaire (PriSnQuest; (Horton et al., 2014)), the Self-Harm Inventory (SHI; (Horton et al., 2014)), the Suicide Potential Scale (Wichmann et al., 2000) and the Viennese Instrument for Suicidality in Correctional Institutions (VISCI; (Frottier et al., 2009))

Of the eligible studies reporting on risk of sexual reoffending:

- 3 reported on the Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR; (Hanson & Thornton, 2000; Seto et al., 2004; Sjostedt & Grann, 2002b))
- 2 each reported on the Screening Scale for Paedophilic Interests (Helmus et al., 2015; Seto et al., 2004), the Sex Offender Risk Appraisal Guide (SORAG; (Kingston et al., 2010; Seto et al., 2004)) and the Static-2002 and it's revised version (Hanson et al., 2010; Helmus et al., 2015)
- 1 each reported on the Offender Group Reconviction Scale (OGRS; (Wakeling et al., 2011a)), the Risk Matrix 2000 (RM2000; (Wakeling et al., 2011a)), the Stable 2007 (Helmus et al., 2015), the Structured Anchored Clinical Judgment (SACJ/SACJ-Min; (Hanson & Thornton, 2000)), the Violence Risk Scale: Sex Offender Version (VRS:SO; (Beggs & Grace, 2010)) and the VRS:SO Deviance subscale (Beggs & Grace, 2010)

Of the eligible studies reporting on risk of relapse into substance misuse:

- 1 each reported on the Alcohol Use Disorders Inventory Test (AUDIT; (Thomas et al., 2014)) and the Relapse Screening Questionnaire (RSQ; (Spurgeon et al., 2000))

Further information about included and excluded studies can be found in Appendix K. A summary of the methodological quality of the studies is presented in Table 40. If data was presented in sufficient detail for analysis, the data are presented using forest plots and summary ROC curves in Appendix O.

Table 38: Characteristics of risk assessment tools with acceptable diagnostic accuracy (sensitivity and specificity ≥ 70%).

Tool	Target disorder/ behaviour	Intended population/ setting	Scale information	Recommended cut-off	Format	Administration and qualifications	Cost/ restrictions
Offender Group Reconviction Scale (OGRS)	Offender reconviction	Previous offenders	6 items	Higher score = higher probability of reconviction	Assessment scored using official records	Administration Time: n/r Training: Not required Administered by general probation staff	Available through OASys
PriSnQuest	Mental Illness	Criminal Justice System	8 items	n/r	n/r	Administration Time: n/r Training: n/r Administered by general prison staff	Unclear
RRASOR	Sexual recidivism	Sex offenders	4 items Score: 0-6	n/r	Assessment scored using official records	Administration/ Scoring time: n/r Training: no clinical expertise required	Unclear
SACJ/SACJ-Min	Sexual and violent recidivism	Adult male sex offenders	3 stage assessment	Risk of sexual and violent recidivism: <2 = low risk 2-3 = medium risk ≥4 = high risk	Stage One: initial actuarially based screening Stage Two: a more in-depth analysis of aggravating factors		Unclear

Tool	Target disorder/ behaviour	Intended population/ setting	Scale information	Recommended cut-off	Format	Administration and qualifications	Cost/ restrictions
					Stage Three: careful monitoring of offender performance over time to note the impact of treatment on risky dispositions		
SCOPE	Suicide risk and deliberate self-harm	Prison	27 items across 2 domains (protective social networks and optimism) Score: 0-162	Risk of suicide or deliberate self-harm: >38 on domain 1, or >30 on domain 2	Self-report Likert-type questionnaire	Administration Time: <5 minutes Training: n/r Administered by general prison staff	Freely available as online assessment
Static-2002/Revised	Sexual and violent recidivism	Adult male sex offenders	14 items Score: 0-14	Risk of sexual and violent recidivism: 0-2 = low risk 3-4 = low-moderate risk 5-6 = moderate risk 7-8 = moderate-high risk 9+ = high risk	Assessment scored using official records – can be supplemented by self-report	Administration/ Scoring Time: n/r Training: One day training from certified trainer recommended	Freely available
VISCI	Suicide	Prison	22 items but not all items are	Risk of suicide Pre-trial:	Dichotomous (yes/no)	Administration Time: n/r	Unclear

Tool	Target disorder/ behaviour	Intended population/ setting	Scale information	Recommended cut-off	Format	Administration and qualifications	Cost/ restrictions
			scored Scored items: 7 for pre-trial offenders 8 for sentenced offenders	≥3.12 risk present ≥6.89 high risk Risk of suicide after Sentenced: ≥1.93 risk present ≥5.45 high risk	questionnaire	Training: Not required Administered by general prison staff	

Table 39: Characteristics of risk assessment tools with unacceptable diagnostic accuracy (sensitivity or specificity < 70%)

Tool	Target disorder/ behaviour	Intended population/ setting	Scale information	Recommended cut-off	Format	Administration and qualifications	Cost/ restrictions
Alcohol Use Disorders Inventory Test (AUDIT)	Hazardous alcohol consumption	General	10 items	Risk of hazardous alcohol consumption: 0-7 = low risk 8-15 = moderate-low risk 16-19 = moderate-high risk 20-40 = high risk	Self-report Likert-type questionnaire	Administration/ Scoring Time: n/r Training: primary health care	Freely available
Beck Hopelessness Scale (BHS)	Suicide risk	General, Adults 17 – 80	20 item	n/r	Self-report inventory	Administration Time: n/r Training: Can be administered by general prison staff – to be interpreted by	Manual: $83 Record forms: $58 Scoring key: $10.50

Tool	Target disorder/ behaviour	Intended population/ setting	Scale information	Recommended cut-off	Format	Administration and qualifications	Cost/ restrictions
						mental health clinician	
Risk Matrix 2000 (RM2000)	Sexual and violent recidivism	Adult male sex offenders	Consists of 3 scales: RM2000/S for sexual offending. RM2000/V for non-sexual violence engaged RM2000/C is a combination of both	n/r	Dynamic Assessment scored using official records	Administration Time: n/r Training: n/r	Unclear
RSQ	Substance misuse relapse	Adults on probation	23 items	Risk of substance misuse relapse: 0-39 = low risk 40-69 = moderate 70-89 = high 90-99 = severe	Self-report Likert-type questionnaire	Administration Time: n/r Training: Not required Administered by general probation staff	Unclear
Screening Scale for Paedophilic Interests (SSPI)	Paraphilic recidivism	Paraphilic sex offenders	4 items	Higher score = higher probability of sexual recidivism	Dichotomous (present/absent) questionnaire	Administration/ Scoring time: brief Training: no clinical expertise required	Unclear
Self-Harm Inventory (SHI)	Self-harm	General	22 items	n/r	Dichotomous (yes/no)	Administration	Freely available

Tool	Target disorder/ behaviour	Intended population/ setting	Scale information	Recommended cut-off	Format	Administration and qualifications	Cost/ restrictions
					questionnaire	Time: <5 minutes Training: n/r	
Stable 2007	Sexual and violent recidivism	Adult male sex offenders	13 items	n/r	Dynamic Assessment scored using official records	Administration Time: n/r Training: Not required Administered by general probation staff	Freely available
Suicide Potential Scale (SPS)	Suicide	Outpatient	6 items	>0 = should be considered at risk of suicide	Dichotomous (yes/no) questionnaire	Administration Time: n/r Training: n/r	Unclear
VRS:SO	Sexual offending risk and change in risk as a function of intervention	Sex offenders	24 items Score: 0-72	Risk of sexual offence: 0-20 = low risk 21-30 = moderate-low risk 31-40 = moderate-high risk 41-72 = high risk	Assessment scored using official records	Administration/ Scoring Time: n/r Training: two-day workshop from certified trainers recommended Administered by qualified health/social care staff	Manual: $50 Score-sheets: $1

1

Table 40: Quality assessment of studies included in the review of the most appropriate tools for assessment of risk

Study ID	Index test	Risk of bias				Applicability concerns		
		Participant selection	Index test	Reference standard	Flow and timing	Participant selection	Index test	Reference standard
Beggs 2010	VRS: SO	Unclear	Low	Unclear	Unclear	Low	Low	Low
Frottier 2009	VISCI	High	Unclear	Low	Unclear	Low	Low	Low
Hanson 2000	RRASOR	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low
Hanson 2000	SACJ/SACJ-Min	Unclear	Low	Unclear	Unclear	Low	Unclear	Low
Hanson 2010	Static-2002/Static-2002R	Unclear	Low	Unclear	Unclear	Low	Unclear	Low
Helmus 2015	SSPI	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Helmus 2015	Stable 2007	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Helmus 2015	Static-2002/Static-2002R	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Horton 2014	PriSnQuest	Unclear	Unclear	Unclear	High	Low	High	Unclear
Horton 2014	SHI	Unclear	Unclear	Unclear	High	Low	High	Unclear
Ivanoff 1991	BHS	High	Unclear	Unclear	Unclear	Low	Low	High
Kingston 2010	SORAG	Unclear	Low	Unclear	Unclear	Low	Low	Low
Naud 2010	SPS	Unclear	Low	Unclear	High	Low	Low	Low
Naud 2013	SPS	Unclear	Unclear	Unclear	Low	Low	Low	Low
Perry 2009a	BHS	Unclear	Unclear	Unclear	High	Low	Low	High
Perry 2009a	SCOPE	Unclear	Unclear	Unclear	High	Low	Low	High
Perry 2009b	BHS	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Perry 2009b	SCOPE	Unclear	Unclear	Unclear	Unclear	Low	Low	Low
Perry 2009c	SCOPE	Unclear	Unclear	Unclear	High	Low	Low	High

Study ID	Index test	Risk of bias				Applicability concerns		
		Participant selection	Index test	Reference standard	Flow and timing	Participant selection	Index test	Reference standard
Seto 2004	SSPI	Low	Unclear	Low	Low	Low	High	Low
Seto 2004	RRASOR	Low	Unclear	Low	Low	Low	Low	Low
Seto 2004	SORAG	Low	Unclear	Low	Low	Low	Low	Low
Sjostedt 2002	RRASOR	Low	Unclear	Unclear	Unclear	Low	Low	Low
Spurgeon 2000	RSQ	Unclear	Low	Unclear	High	Low	Low	Low
Thomas 2014	AUDIT	Low	Low	Unclear	High	Low	Low	Low
Wakeling 2011a	OGRS	Unclear	Unclear	Low	High	Low	Unclear	Low
Wakeling 2011a	RM2000	Unclear	Low	Low	High	Low	Low	Low
Wichmann 2000	SPS	High	Unclear	Unclear	Unclear	Low	Low	Low

Note.
AUDIT= Alcohol Use Disorders Inventory Test; BHS=Beck Hopelessness Scale; OGRS=Offender Group Reconviction Scale; PriSnQuest=Prison Screening Questionnaire; RM2000= Risk Matrix 2000; RRASOR= Rapid Risk Assessment for Sexual Offense Recidivism; RSQ SACJ/SACJ-Min=Structured Anchored Clinical Judgment; SCOPE= Self-harm Concerns about Offenders in Prison Environment; SHI=Self-Harm Inventory; SORAG=Sex Offender Risk Appraisal Guide; SPS=Suicide Probability Scale; SSPI=Screening Scale for Pedophilic Interests; VISCI=Viennese Instrument for Suicidality in Correctional Institutions; VRS: SO=Violence Risk Scale: Sexual Offender Version.

5.5.11 Tools without acceptable sensitivity and specificity

The GC agreed to only review tools and cut-off points with acceptable sensitivity and specificity, which was determine by a relatively conservative threshold of ≥0.70 for both values. In the absence of values for sensitivity and specificity, tools with AUC values ≥0.75 were considered to have acceptable performance.

Therefore, evidence relating to the following tools were not considered by the GC: Offender Group Reconviction Scale (OGRS), Risk Matrix 2000 (RM2000), Screening Scale for Paedophilic Interests, Sex Offender Risk Appraisal Guide (SORAG), Stable 2007, Structured Anchored Clinical Judgment (SACJ/SACJ-Min), Violence Risk Scale: Sex Offender Version (VRS:SO), Beck Hopelessness Scale (BHS), Prison Screening Questionnaire (PriSnQuest), Self-Harm Inventory (SHI), Suicide Potential Scale, Suicide Probability Scale (SPS), Alcohol Use Disorders Inventory Test (AUDIT) and Relapse Screening Questionnaire (RSQ). An overview of studies examining these tools can be found in Table 38 and Table 38, for those tools with acceptable and unacceptable accuracy respectively.

5.5.12 Risk of self-harm and/or suicidal behaviour

2 included studies examined the sensitivity and specificity of 2 risk assessment tools for self-harm and/or suicidal behaviour (N = 1331): Perry 2009a and Frottier 2009.

These tools are the SCOPE (1 cohort study) and the VISCI (1 case-control study). 2 further cohort studies were not considered by the GC, one because it examined only individual subscales of the SCOPE (Perry 2009c) and the other as the cut-off used for the SCOPE resulted in unacceptably low sensitivity and specificity. An overview of the studies included in this review can be found in Table 41. Summary of findings can be found in Table 42.

Table 41: Study information table for the review of the most appropriate tools for risk assessment of self-harm and/or suicidal behaviour

	SCOPE	VISCI
Total no. of studies (N[1])	1 (1166)	1 (165)
Study ID	Perry 2009a	Frottier 2009
Study design	(1) cohort study	(1) case-control study
Country	UK	Austria
Reference Standard(s)	Self-report	Official records
Setting	Prison	Prison
Age (mean)	23.8 years	n/r
Sex (% female)	40%	n/r
Ethnicity (% Caucasian)	87%	n/r
Note. [1]N = total number of participants		

Table 42: Summary of findings table for the review of the most appropriate tools for risk assessment of self-harm and/or suicidal behaviour

Tool	Cut-off	Total no. of studies (N)	Sensitivity	Specificity	PPV	NPV	Quality[1]
SCOPE	76-78	1 (681)	0.72-0.76	0.70-0.74	NR	NR	Low
VISCI	3.38	1 (75)	0.72	0.82	0.67	0.85	Low

Note. N = total number of participants who provided data; NR = not reported; PPV = positive predictive value; NPV = negative predictive value

[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall

Tool	Cut-off	Total no. of studies (N)	Sensitivity	Specificity	PPV	NPV	Quality[1]
assessment of the risk of bias and applicability QUADAS-II (adapted) domains.							

5.5.123 Risk of sexual reoffending

Four studies examined the performance of tools to support or assist in risk assessment for sexual reoffending (N = 2625): Hanson 2010, Helmus 2015, Seto 2004 and Sjostedt 2002b. None of these studies reported sensitivity and specificity. Instead AUC values were reported.

One further cohort study (Beggs 2010) examined only a single subscale of the VRS:SO and so was not considered by the GC. An overview of the studies included in this review can be found in Table 43. Summary of findings can be found in Table 44.

Table 43: Study information table for the review of the most appropriate tools for risk assessment of sexual reoffending

	RRASOR	Static-2002
Total no. of studies (N[1])	2 (1416)	2 (608)
Study ID	(1) Seto 2004 (2) Sjostedt 2002	(1) Hanson 2010 (2) Helmus 2015
Study design	(1,2) cohort study	(1,2) cohort study
Country	(1) Canada (2) Sweden	(1) UK (2) Canada
Reference Standard(s)	(1,2) Official records	(1,2) Official records
Setting	(1,2) Prison	(1) Community (2)Various
Age (years, mean)	(1) 43.0 (2) 41.0	(1) 43.0 (2) 42.8
Sex (% female)	(1,2) 0%	(1,2) 0%
Ethnicity (% Caucasian)	(1,2) not reported	(1,2) not reported
Note. N[1] = total number of participants;		

Table 44: Summary of findings table for the review of the most appropriate tools for risk assessment of sexual reoffending

Tool	Cut-off	Total no. of studies (N)	AUC (95% CI) range	PPV	NPV	Quality[1]
RRASOR	n/a	2 (585)	0.75 (0.65–0.85) – 0.83 (73–0.93	NR	NR	Moderate to high
Static-2002	n/a	2 (1401)	0.77 (0.70–0.85) –0.79 (0.70–0.88)	NR	NR	Low to moderate
Note. N = total number of participants who provided data; NR = not reported; PPV = positive predictive value; NPV = negative predictive value						
[1]Studies were assigned a quality rating for use in clinical evidence statements according to an overall assessment of the risk of bias and applicability QUADAS-II (adapted) domains.						

5.5.132 Economic evidence

No economic evidence on the tools for risk assessment for adults with mental health problems who are in contact with the criminal justice system was identified by the systematic

search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

5.5.3 Clinical evidence statements

5.5.3.1 Risk of self-harm and/or suicidal behaviour

There was low quality evidence from one study (N = 681) that the SCOPE with cut-off points 76, 77 and 78 has clinically useful (≥70%) sensitivity and specificity for the identification of individuals with self-harm and/or suicidal behaviour. Sensitivity was optimised at 76% with a cut-off of 76; specificity is optimised at 74% with a cut-off of 78.

There was low quality evidence from one study (N = 75) that the VISCI with a cut-off of 3.38 has clinically useful (≥70%) sensitivity and specificity for the identification of individuals who complete suicide.

5.5.3.2 Risk of sexual reoffending

There was moderate-high quality evidence from two studies (N = 1401) that the RRASOR has a clinically useful (>.75) AUC value for the prediction of sexual recidivism.

There was moderate-low quality evidence from two studies (N = 585) that the Static 2002 has a clinically useful (>.75) AUC value for the prediction of sexual recidivism.

5.5.4 Economic evidence statements

No economic evidence on tools for the risk assessment for adults with mental health problems who are in contact with the criminal justice system is available.

5.6 Recommendations and link to evidence

Recommendations	No recommendations were made about what tools to use to undertake risk assessment
Relative values of different outcomes	The GC agreed that the most important outcomes in risk assessment within the criminal justice system related to self-harm or suicide risk, risk of sexual reoffending and risk of relapse as these have the greatest potential for benefit or harm for both the service user and the general public.
Trade-off between clinical benefits and harms	When considering whether or not to recommend a case identification tool, the GC were mindful of the benefits associated with the identification of mental health problems in the prison population (which are known to be significantly higher than in the general population) but also considered the potential harm or inappropriate use of resources that may arise from false positives. For this reason, the GC were careful to evaluate both the sensitivity and specificity of the measures reviewed. The GC agreed that risk assessment tools can be helpful aids in clinical decision making. However, they also considered the importance of high sensitivity and specificity of risk assessment tools, and for this reason set conservative thresholds for both of these. They agreed that this was particularly important when such tools may inform decisions about treatment and the most appropriate setting in which this should take place, including decisions about continued detention.

	The GC considered the benefits of risk assessment tools against the potential for false negatives or for the tools to be misused or misinterpreted. In particular, the GC were concerned that such tools should not be considered in isolation when making a determination about the extent of a risk and should not be seen as a substitute for a comprehensive approach to decision making drawing on a number of sources of data to inform a decision. Decisions made solely on the basis of a rating scale could do harm by leading to an under/over estimation of the risk. Evidence was available for two measures, the RRASOR for which moderate quality evidence indicated good sensitivity and the Static 2002 which had lower quality evidence and good sensitivity.
Trade-off between net health benefits and resource use	There was no evidence on the cost-effectiveness of tools to support or assist in risk assessment for adults with mental health problems. The GC considered the time it takes to administer risk assessment tools and the consequence associated with self-harm. The GC noted that risk assessments, in particular for those with mental health problems in contact with the criminal justice system is a routine part of all assessments. Therefore offering risk assessment based on a set of key principles that the GC developed from the formal consensus methods for Review protocol 5.6 would not result in significant extra resource implications.
Quality of evidence	The was low quality evidence for self-harm and suicide risk assessment tools based on one small and one medium sized study undertaken only in prison settings. Given the limited evidence, the GC did not consider it sufficient to recommend a particular tool. The GC agreed that the evidence for the RRASOR, ranging from low to high quality, may be sufficient to consider making a recommendation, although it had only been evaluated in a prison setting. The evidence for the Static 2002 was of lower quality and the GC did not think it sufficient to support a recommendation.
Other considerations	The GC noted that existing NICE recommendations (Self-harm in over 8s: short-term management and prevention of recurrence and Self-harm in over 8s: long-term management) advise against the use of structured risk assessment tools to predict future self-harm or suicide, or to determine who should be offered treatment, but that such tools could be considered to structure a risk assessment. The low quality evidence identified in this review did not support recommending any specific tool developed for use in the criminal justice system. The GC therefore did not recommend the use of any tools when developing the recommendations for the assessment of risk or as part of any screening process, for example a reception into prison assessment. Despite identifying some low to high quality evidence for the RRASOR to predict paraphilic reoffending, the GC were concerned about recommending the tool as the scores are largely determined by key historical events, which could lead to no changes in the prediction of risk of re-offending even if there were other significant changes that might suggest a

	significant change in the level of risk. Therefore, the GC decided not to make a recommendation for use. In the GC's correspondence with the developer of this tool, the developer's view was that the instrument should no longer be used in routine practice. The GC noted that other risk assessment tools were used within criminal justice settings (in particular, for the assessment of violent offending), but were not considered as they were outside the scope of the guideline. The GC considered making a research recommendation for the recognition of risk to self but decided that priority should be given to research which focused on the factors associated with suicide (which the GC made a research recommendation for). The output of this research could then inform the development of future recognition tools, without such knowledge there is a danger that less than optimal tools will be developed.

5.7 Review question: What are the key components of, and the most appropriate structure for a comprehensive assessment of mental health problems for adults:

- in police custody?
- for the court process?
- at reception into prison?
- at subsequent time points in prison?
- in the community (serving a community sentence, released from prison on licence or released from prison and in contact with a community rehabilitation company [CRC] or the probation service)?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 45. A complete list of review questions and full review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 45: Clinical review protocol summary for the review of tools and methods to support a comprehensive assessment

Component	Description
Population	Adults (aged 18 and over) with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Index test	Any formal recognition and assessment tools considered appropriate and suitable for use
Reference standard	Assessment of mental health problems by an experienced clinician
Outcomes	Critical - Reliability (for instance, inter-rater or test-retest reliability or internal consistency); Validity (for instance, criterion or construct validity), Improved assessment of need Important – Practicality/Feasibility for use in routine care; improved care planning and organisation of care
Study design	Not applicable (group consensus was used).

5.7.1 Group consensus for the key components of, and the most appropriate structure for, a comprehensive assessment of mental health problems for adults within the criminal justice system

When agreeing the review protocol the GC decided, based upon the initial scoping searches undertaken for the guideline; their existing knowledge of the evidence base and from a consideration of published NICE mental health guidelines, that searching for published evidence on this topic it would not be a good use of time and resource. Additionally, they agreed that given the criminal justice system comprises a number of very varied settings, including some with very specific characteristics (e.g. prison) that it would not be productive to consider indirect evidence (e.g. from other mental health guidelines). The GC therefore decided to develop a set of principles to inform assessment methods for use with this population using a modified form of the nominal group technique. The method used for the nominal group technique is described in full within the methods section in Chapter 3.

Key issues related to comprehensive assessment within this population were identified through a range of sources and from discussions within the GC meetings. These issues were used to generate nominal statements covering a range of areas that had been identified as important by the GC. These included ensuring that assessments were rigorous, how others should be involved in the assessment, and who these individuals should be and the importance of being clear about the intended product of the assessment. These statements were grouped together into 6 areas each with its' own questionnaire; principles, purpose, structure, outcomes, risk management and additional considerations, and these 6 questionnaires were then distributed to the GC to be rated. Examples of statements that were rated highly by the committee are 'A comprehensive assessment should include all services involved in the care of the service user' and 'Staff conducting a comprehensive assessment should be able to appraise the reliability and validity of data sources'.

The risk management questionnaire was completed by 14 of the 19 GC members, but the other 5 questionnaires were completed by 16 of the 19 members (round 1). Percentage consensus values were calculated, and comments collated, for each statement. The rankings and comments were then presented to the GC members, and used to inform a structured discussion within the GC meeting. Generally, there was high agreement among the GC members however where there were statements with lower agreement these were re-drafted to account for comments from the GC members and re-distributed in questionnaire form (round 2). This was completed by 14 of the 19 GC members. Discussions following each round of ratings led to the development of recommendations in this area. A brief summary of the outcome of this process is depicted in Table 46 below. The full list of statements and ratings can be found in Appendix V and blank copies of the questionnaires used can be found in Appendix U.

Table 46: Summary of nominal group technique process followed for the development of recommendations on the structure and key components of a comprehensive assessment of mental health problems for adults within the criminal justice system

Round 1		**Round 2**		**No. of recommendations generated**
Principles of comprehensive assessment				21 recommendations
Level of agreement	*Statements N (total=24)*	*Level of agreement*	*Statements N (total=4)*	
High	20	High	2	
Moderate	4	Moderate	2	
Low	0	Low	0	
Purpose of comprehensive assessment				

Round 1		Round 2		No. of recommendations generated
Level of agreement	*Statements N (total=19)*	*Level of agreement*	*Statements N (total=4)*	
High	15	High	4	
Moderate	4	Moderate	0	
Low	0	Low	0	
Structure of a comprehensive assessment				
Level of agreement	*Statements N (total=12)*	*Level of agreement*	*Statements N (total=2)*	
High	7	High	2	
Moderate	4	Moderate	0	
Low	1	Low	0	
Outcomes from a comprehensive assessment				
Level of agreement	*Statements N (total=20)*	*Level of agreement*	*Statements N (total=3)*	
High	14	High	1	
Moderate	4	Moderate	2	
Low	2	Low	0	
Risk management				
Level of agreement	*Statements N (total=11)*	*Level of agreement*	*Statements N (total=0)*	
High	10	High	n/a	
Moderate	1	Moderate	n/a	
Low	0	Low	n/a	
Additional considerations during a comprehensive assessment				
Level of agreement	*Statements N (total=13)*	*Level of agreement*	*Statements N (total=0)*	
High	12	High	n/a	
Moderate	1	Moderate	n/a	
Low	0	Low	n/a	

<Insert Note here>

5.7.2 Economic evidence

No studies assessing the cost effectiveness of methods for the assessment of mental health problems in people who are in contact with the criminal justice system were identified by the systematic search of the literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

5.8 Clinical evidence statements based upon formal consensus ratings

Regarding the principles of a comprehensive assessment

The GC agreed that assessments should:

- be understood and relate to a particular context,
- be reviewed as appropriate,
- should identify service user strengths,

- should consider the impact of the physical environment on psychological distress
- should be followed by a feedback appointment where possible

They agreed that assessments should

- be collaborative and maximise everyone's contribution
- include all relevant services and agreed family members or carers.

They agreed that a formulation should clearly acknowledge factors the service user considers pertinent and that differences between service user and staff views should be acknowledged.

They agreed that assessments should be paced and structured according to the service user's comprehension, adjustments should be made for any learning disabilities and an appropriate adult or specialist should be involved where appropriate.

They agreed staff should be competent in a range of relevant communication skills and that the assessment should be responsive to new information.

They decided it was important for a clear and detailed record of the outcome to be kept.

They also agreed that the assessment should aim to understand the relationship between offending behaviour and mental health, and develop alternative adaptive strategies.

They decided that it was important for the comprehensive assessment to integrate with other care plans.

There was moderate agreement for involving a person from the service user's network where this is appropriate, and to consider using validated tools that are relevant to the disorder under assessment.

There was also moderate agreement for agreeing a preferred format for feedback from the assessment in advance.

Regarding the purpose of a comprehensive assessment

The GC decided that it is important to obtain an understanding of the person's problem, including the nature and severity of these problems and identify adaptations to interventions or the environment that the service user requires.

The GC decided that

- the purpose of the assessment should be made clear in advance.
- the assessment should assess multiple areas of need, take into account symptom severity, service user understanding and assess for coexisting problems.
- risk to self and others should be assessed, as well as potential triggers and probability of risky events.
- risk assessment should result in a risk management plan and that this should identify interventions and factors that may reduce risk.
- a formulation should provide a shared understanding of the problem, including its development and maintenance, the focus and impact of interventions, barriers to engagement and the impact of the social and physical environment.

There was moderate agreement for the assessment to consider the impact of mental health problems on treatment planning, to obtain a diagnosis or problem specification and to systematically assess a range of factors during risk assessment.

Regarding the structure of a comprehensive assessment

The GC decided that they should be multidisciplinary, with a named lead person and organisation, and that assessing staff should know about diagnostic classifications and their limitations.

The GC agreed that staff should be trained and competent in the use of a range of assessment and outcome monitoring measures, preferably using those developed or adapted for the criminal justice system, and able to appraise the reliability of data sources.

They decided that the assessment should integrate information from multiple sources, corroborating information from other informants with that of the service user and reviewing past history and behaviour.

There was moderate agreement for assessments to consider the views of others relevant to the service user, and for staff to select assessment tools based upon their utility, cost and availability.

Regarding the outcomes of a comprehensive assessment

The GC decided that it is important to identify realistic and optimistic goals, the steps needed to achieve these and to prioritise areas most amenable to change.

They agreed that staff should ensure service users are aware of the need to monitor and report risk behaviours.

They agreed that a care plan should result from the assessment and initial formulation produced during that assessment, and that this should be developed and communicated both verbally and in writing to relevant parties as soon as possible.

They agreed that crisis and risk management plans should be incorporated into the care plan, that the care plan should be multidisciplinary and developed collaboratively.

They also agreed that the care plan should identify appropriate evidence-based interventions and referral options, include a profile of the service-user's needs, and take into account the needs of families and carers.

They agreed that referrers should ensure they provide sufficient information to allow the referral to proceed.

They also agreed that symptoms and functioning should be monitored regularly and that there should be an agreement on when both the assessment and progress will be reviewed.

There was moderate agreement for goals being agreed with the service user, and for outcomes to be explicitly linked to goals and intended targets of interventions.

Regarding risk management

The GC agreed that risk management plans should be written to take into account the setting in which they will be implemented and applicable policies or statutory responsibilities.

They agreed that risk management plans should

- be shared appropriately with other involved agencies and should clearly specify the procedure for review.
- include interventions to reduce risk and minimise harm, and should be individual to the service user.
- enable service users themselves to actively participate in risk management and to appreciate that risk levels will fluctuate over time.

There was moderate agreement for risk management plans to include proactive interventions

Regarding additional considerations during a comprehensive assessment,

The GC agreed that staff should clearly set out the boundaries of confidentiality for the service user.

They agreed that staff should be aware of the potential for the service user to have negative expectations based upon their previous experiences and counter this by maintaining a manner that is empathic and non-judgemental and discussing difficulties in a way engenders hope.

They also agreed that assessments should be undertaken in a suitably private environment.

They agreed that staff should share both pre-existing information and assessment outcome with other agencies according to local procedures and policies, and that routine systems for this should be developed.

There was moderate agreement for the need for staff to be aware of the potential for service users to either feign or minimise mental health problems.

5.7.4 Economic evidence statements

No evidence on the cost effectiveness of methods for the comprehensive assessment of mental health problems in people who are in contact with the criminal justice system is available.

5.8 Recommendations and link to evidence

Recommendations	**13. Use this guideline with the NICE guidelines on service user experience in adult mental health and patient experience in adult NHS services to improve the experience of care for people with learning disabilities and mental health problems.** **14. Obtain, evaluate and integrate all available and reliable information about the person when assessing or treating people in contact with the criminal justice system. For example:** • **person escort record (PER)** • **pre-sentence report** • **primary and secondary medical records** • **custody reports** • **Offender Assessment and Sentence Management (OASys).** **Take into account how up to date the information is and how it was gathered.** **15. Work with a family member, partner, carer, advocate or legal representative when possible in order to get relevant information and support the**

	person, help explain the outcome of assessment, and help them make informed decisions about their care. Take into account: • **the person's wishes** • **the nature and quality of family relationships** • **any statutory or legal considerations that may limit family and carer involvement.** **16. Carry out assessments:** • **in a suitable environment that is safe and private** • **in an engaging, empathic and non-judgemental manner.** **17. When assessing a person, make appropriate adjustments to assessment that take into account any suspected neurodevelopmental disorders, cognitive impairments, or physical disabilities. Seek advice or involve specialists if needed.** **18. Service providers should ensure that a practitioner who is competent and has experience of working with people in contact with the criminal justice system who have mental health problems, undertakes the mental health assessment and where necessary coordinates the input of other professionals into the assessment.** **19. If there are concerns about a person's mental capacity, practitioners should:** • **perform a mental capacity assessment if they are competent to do this (or refer the person to a practitioner who is)** • **consider involving an advocate to support the person.** **20. All practitioners should discuss rights to confidentiality with people and explain:** • **what the assessment is for, and how the outcome of the assessment may be used** • **how consent for sharing information with named family members, carers and other services should be sought** • **that the assessor may have a legal or ethical duty to disclose information relating the safety of the person or others, or to the security of the institution.** **21. All practitioners should ensure assessment is a collaborative process that:** • **involves negotiation with the person, as early as possible in the assessment process, about how**

	information about them will be shared with others involved in their care • makes the most of the contribution of everyone involved, including the person, those providing care or legal advice, and families and carers • engages the person in an informed discussion of treatment, support and care options • allows for the discussion of the person's concerns about the assessment process. 22. Ensure all practitioners carrying out mental health assessments are competent to assess common presenting problems, with an understanding of the context and setting in which they are undertaken. They should: • tailor the content, structure and pace of an assessment to the person's needs and adjust the assessment as new information emerges • take into account the person's understanding of the problem • have knowledge and awareness of diagnostic classification systems and their limitations • appraise the reliability and validity of all available health and criminal justice systems records • identify and take into account the reasons for any significant differences between the assessor's views and those of the person, and other agencies involved in their care • use validated tools relevant to the disorders or problems being assessed • take into account the views of practitioners from other services involved in the person's care. 23. All practitioners carrying out mental health assessment should take into account the following when conducting an assessment of suspected mental health problems for people in contact with the criminal justice system: • the nature and severity of the presenting problems (including substance misuse) and their development and history • coexisting mental health problems • coexisting physical health problems • social and personal circumstances • social care, educational and occupational needs • people's strengths that may help engagement with interventions • previous care, support and treatment, including how the person responded to these

	• offending history, and how this may interact with mental health problems. 24. When assessing people in contact with the criminal justice system all practitioners should: • recognise potential barriers to accessing and engaging in interventions and methods to overcome these • discuss mental health problems and treatment options in a way that gives rise to hope and optimism by explaining that change is possible and attainable • be aware that people may have negative expectations based on earlier experiences with mental health services, the criminal justice system, or other relevant services. 25. All practitioners should share the outcomes of an assessment, in accordance with local policies and legislation, with: • the person and when possible with family members and carers • all staff involved in the direct development and implementation of the plan, • other staff agencies (as needed) not directly involved in the development and implementation of the who could support the effective implementation and delivery of the plan. 26. Practitioners should review and update assessments: • if new information is available about the person's mental health problem • if there are significant differences between the views of the person and the views of the family, carers or staff that cannot be resolved through discussion. • when major legal or life events occur • when the person is transferred between, or out of, criminal justice services • if a person experiences a significant change in care or support (for example, stopping an Assessment, Care in Custody and Teamwork [ACCT] plan). 27. When updating assessments, practitioners should consider: • reviewing demographic, psychological, social, personal historical and criminological factors • assessing multiple areas of need, including social and personal circumstances, physical

	health, occupational rehabilitation, education and previous and current care and support • **developing an increased understanding of the function of the offending behaviour and its relationship with mental health problems** • **covering any areas not fully explored by the initial assessment.** **28. Undertake a risk assessment for all people in contact with the criminal justice system when a mental health problem occurs or is suspected.** **29. All practitioners should include the following in risk assessments for people in contact with the criminal justice system:** • **risk to self, including self-harm, suicide, self-neglect, risk to own health and degree of vulnerability to exploitation or victimisation** • **risk to others that is linked to mental health problems, including aggression, violence and sexual offending and predation** • **causal and maintaining factors** • **the likelihood, imminence and severity of the risk** • **the impact of their social and physical environment** • **protective factors that may reduce risk.** **30. During risk assessment the practitioner undertaking the assessment should explain to the person that their behaviours may need to be monitored. For example, behaviours that may indicate a risk to self or others, or if monitoring will help the person to identify, anticipate and prevent high-risk situations.** **31. The practitioner undertaking the assessment should develop a risk management plan for a person when indicated by their risk assessment. This should:** • **integrate with or be consistent with the mental health assessment and plan** • **take an individualised approach to each person and recognise that risk levels may change over time** • **set out the interventions to reduce risk at the individual, service or environmental level** • **take into account any legal or statutory responsibilities which apply in the setting in which they are used** • **be shared with appropriate parties (including**

	families and carers) and services • be reviewed regularly by those responsible for implementing the plan and adjusted if risk levels change. 32. All practitioners should ensure that management of the risks of self-harm and suicide, the risk of harm to others, the risk of exploitation by others and the risk of self-neglect is: • informed by the assessments and interventions in relevant NICE guidance for the relevant mental health disorders including the NICE guidelines on self-harm in over 8s: short-term management and prevention of recurrence and self-harm in over 8s: long-term management. • implemented in line with agreed protocols for safeguarding and appropriate adults • implemented in line with agreed protocols in police custody, prisoner escort services, prison, community settings and probation service providers • integrated with and recorded in the relevant information systems (for example, the ACCT procedure in prisons, the Offender Assessment and Sentence Management (OASys), and SystmOne and Multi Agency Risk Assessment Conference (MARAC) and multi-agency public protection arrangements (MAPPA). 33. Develop a mental health plan of care in collaboration with the person and, when possible, their family, carers and advocates. All practitioners developing the plan, should ensure it is integrated with care plans from other services, and includes: • a profile of the person's needs, identifying agreed goals and the means to progress towards goals • identification of the roles and responsibilities of those practitioners involved in delivering the plan • a clear strategy to access all identified interventions and services • agreed outcome measures and timescale to evaluate and review the plan • a risk management and a crisis plan if developed • an agreed process for communicating the plan to all relevant agencies, the person, and their families and carers. 34. Give people the opportunity to discuss the outcomes and implications of their assessment and the content of their plan of care with the

	practitioner undertaking the assessment. **35. When developing or implementing a plan of care all practitioners should take into account:** • **the ability of the person to take in and remember information** • **the need to provide extra information and support to help with the understanding and implementation of the plan of care** • **the need for any adjustment to the social or physical environment** • **the need to adjust the structure, content, duration or frequency of any intervention** • **the need for any prompts or cognitive aids to help with delivery of the intervention.**
Relative values of different outcomes	The GC were concerned to develop recommendations which would produce a reliable and valid comprehensive assessment of need and facilitate the development of a care plan. The GC were interested particularly in factors which differentiate the assessment of service users within the criminal justice system different from those in general mental health services. They agreed that the context changes the way in which clinicians interact with individuals in order to engage them in an assessment, and that the higher risk of self-harm makes accurate identification of those with mental health problems, facilitated by engaging the service user in the assessment, crucial. They also noted the importance of considering the impact of offending behaviour on mental health and vice versa, the importance of inter-professional and inter-agency coordination, the interaction between physical and mental health, and the need to revisit assessments as circumstances change as a person moves long the criminal justice pathway.
Trade-off between clinical benefits and harms	The GC agreed that the benefits of a comprehensive assessment were the reliable and valid identification of the needs of individuals who require mental health interventions with the potential for timely intervention, improved clinical outcomes and potentially lower rates of reoffending as a result. The benefits of a risk assessment were similar, but included the potential to change the intensity of supervision or input when people are identified as being at high risk of self-injury, and the associated avoidance of serious self-harm and suicide, reduced risk of exploitation and harm to others. The potential harms of a comprehensive assessment and risk assessment relate predominantly to assessments carried out by staff lacking relevant skills or without experience of the criminal justice system, being unaware of the culture of the criminal justice system, or where important information could not be obtained or key staff consulted resulting in misdiagnosis, inadequate care plans or inaccurate risk assessment, possibly leading to avoidable harm or unnecessary detention.
Trade-off between net health benefits and resource use	There was no evidence on the cost effectiveness of methods for the comprehensive assessment of mental health problems in people in contact with the criminal justice system. However, the GC expressed their view that if such assessment leads to a timely identification and appropriate treatment of mental health problems then the additional costs associated with undertaking such

	assessment are likely to be outweighed by the improvements in the short-term provision of more effective care and improved mental health outcomes in the longer term with potential future cost savings to the healthcare system (delays in treatment exacerbate symptoms) and criminal justice system (improvement in mental health may prevent future reoffending).
Quality of evidence	The GC used a formal consensus method (NGT) which although characterised by high levels of agreement across all areas constitutes low quality evidence.
Other considerations	The GC were aware of the specific recommendations about assessment in other NICE mental health guidelines. The GC would expect practitioners to consider these recommendations and the recommendations from NICE guidelines on the experience of care to inform the assessment of specific mental disorders to which this guideline relates. The GC noted that there was limited evidence available in this area and considered making a recommendation for further research. However they agreed that other areas were a higher priority for research and so did not make a research recommendation in this instance.

1

2
3

6 Interventions

6.1 Introduction

It is widely acknowledged that there is a high prevalence of often complex mental health problems experienced by people in contact with the Criminal Justice System. Unfortunately, specific research on the mental health needs of people in contact with criminal justice system and on the effectiveness of interventions for this population has been very limited. There are considerable challenges to be faced when researching interventions with this population such as difficulties in engagement and challenges in delivering interventions in Criminal Justice settings due to environmental constraints that are not always conducive to therapeutic interventions.

Fortunately for many people in contact with the criminal justice system existing NICE guidance for specific conditions may well be applicable in most, if not all cases. What is not well understood, is where guidance may well not apply, how interventions may need to be adapted to be delivered effectively in the criminal justice environment. For example, do psychological interventions need be flexible to take account of difficulties of the prison environment?

Personality disorder is very common in the criminal justice system as a primary or co-morbid diagnosis. People with personality disorders should not be excluded from health interventions because of personality disorder although interventions may need to be modified in duration or intensity. Interventions should facilitate learning and develop new behaviours in: problem solving; emotion regulation and impulse control; managing interpersonal relationships; and self-harm.

Certain issues in the prison environment that all prescribers should be aware of are the risks of overdose or diversion associated with in-possession medications, problems with administration times of not in possession (NIP) medications particularly last dispensing times often being in early evenings (e.g. sedative anti-psychotics and anti-depressants), and the availability of medications in first 48 hours in custody and on release. Polypharmacy, for mental health and physical conditions, is common in people in contact with the criminal justice system and should also be guarded against where at all possible. There may also be particular difficulties with medications that are open to abuse such as hypnotics and medications for chronic pain.

However, despite all these cautions, the fundamental challenge remains in effectively identifying mental problems and ensuring that people in contact with the criminal justice system are offered or referred to effective mental health treatments and are supported in accessing these treatments. It remains the case that, many if not the majority of, such people are not accessing treatment with negative effects on their mental and physical health and a potential increase in the likelihood of re-offending.

6.2 Review question: What are the most effective interventions to promote mental health and wellbeing in adults in contact with the criminal justice system (including environmental adaptations and individual- and population-based psychoeducational interventions)?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 47. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 47 Clinical review protocol summary for the review of effective interventions to promote mental health and wellbeing in adults in contact with the criminal justice system

Component	Description
Population	Adults (aged 18 and over) with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Intervention(s)	• Psychological and social interventions • Pharmacological interventions • Combined psychological or social and pharmacological interventions • Support and education interventions aimed at promoting mental health and wellbeing (including environmental adaptations and individual- and population-based psychoeducational interventions)
Comparison	• Treatment as usual • No treatment • Waitlist control • Placebo (including attention control) • Any alternative management strategy
Outcomes	• Critical – Improvement in mental health and well-being • Important – Improvement in knowledge and awareness about mental health problems; Improvement in uptake and access to mental health services
Study design	Systematic reviews of RCTs and RCTs

6.2.1 Clinical evidence

6.2.1.1 Parent training for parent-child attachment for women with sub-threshold symptoms of depression

Two RCTs (N = 308) met the eligibility criteria for this review. Sleed et al. (2013) evaluated Better Start, a manualized intervention including group parent training sessions and home visits for women following release from specialized mother and baby units within prisons in England and Wales. Menting et al. (2014) was a Netherlands study evaluating New Beginnings, a manualized attachment-based group intervention developed specifically for mothers and babies in prison. An overview of the trials included in the analysis can be found in Table 48. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 49. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 48 Study information for trials included in the analysis of parent training for parent-child attachment for women with sub-threshold symptoms of depression

	Parent training versus treatment as usual
Total no. of studies (N[1])	2 (308)
Study ID	(1) Menting 2014 (2) Sleed 2013
Study design	(1,2) RCT
Country	(1) Netherlands (2) UK

	Parent training versus treatment as usual
Diagnosis	(1) Diagnostic status unclear but paper states “mothers reported high levels of maternal distress, including depression” (2) Sub-threshold symptoms of depression (CES-D=15)
Age (mean)	(1) Mothers: NR; Children: 6.4 years (2) Mothers:26.8 years; Babies:4.7 years
Gender (% female)	(1) Mother: 100%; Children:51% (2) Mother:100%; Babies:61%
Ethnicity (% white)	(1) not reported (2) Mothers: 55%; Babies: 51%
IQ (mean)	not reported
Offence	(1) 57.5% drug-related offences (2) not reported
Expected treatment length (weeks)	(1) 48.1 weeks (2) not reported
Intervention	(1) Better Start, a manualized intervention including group parent training sessions and home visits for women recently released from prison (2) New Beginnings, a manualized attachment-based group intervention developed specifically for mothers and babies in prison
Comparison	(1) Treatment as usual (usual services and help in finding adequate services when needed) (2) Treatment as usual (access to health and social care provision as provided by the prison service)
Format	(1) Individual and group (2) Group
Dose/intensity (hours)	(1) 30(2/first 12 weeks, 0.4/next 17 weeks) (2) 16(4/week)
Intervention setting	(1)83% of group sessions in community centres and 17% (1/6 groups) in prison. Intervention also involved home visits. (2) Prison (7 specialized MBUs in England and Wales)
Length of treatment received (weeks)	(1) 30 (2) 4
Continuation phase (length and inclusion criteria)	(1) 0 (2) 8(2-month post-intervention follow-up but small Ns available for follow-up did not allow for data analysis)
Notes. N= total number of participants; [1]Number randomised. NR-Not reported	

Table 49: Summary of findings table for parent training versus treatment as usual for parent-child attachment for women with sub-threshold symptoms

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Parent training (95% CI)
Depression (CES-D) (Scale from 0 to 60; lower better)	115 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 15.3 (SD 11.8)	MD 1.70 lower (5.65 lower to 2.25 higher)
Number of participants	115 (1 RCT)	⊕⊕◯◯ LOW [1,2]	RR 0.79 (0.51 to	472 per 1,000	99 fewer per 1,000 (231 fewer to 99 more)

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Parent training (95% CI)
with symptoms of depression (CES-D=>16)			1.21)		
Mother-child attachment: Reflective functioning (PDI) (Scale from -1 to 9; higher better)	109 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 3.15 (SD 1.33)	MD 0.39 higher (0.15 lower to 0.93 higher)
Mother-child interaction: Dyadic attunement (behavioural observation; scale from 11 to 55; higher better)	88 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 38.06 (SD 7.3)	MD 3.08 lower (6.39 lower to 0.23 higher)
Mother-child interaction: Parent positive engagement (behavioural observation; scale from 5 to 25; higher better)	88 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 19.9 (SD 3.2)	MD 0.17 lower (1.44 lower to 1.10 higher)
Mother-child interaction: Child involvement (behavioural observation; scale from 6 to 30; higher better)	103 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 16.99 (SD 5)	MD 0.37 lower (2.19 lower to 1.45 higher)
APQ: inconsistent discipline (Scale from 6 to 30; lower better)	102 (1 RCT)	⊕⊕◯◯ LOW[3]	-	Mean 15.88 (SD 3.79)	MD 3.02 lower (4.72 to 1.33 lower)
APQ: positive parenting (Scale from 6 to 30; higher better)	102 (1 RCT)	⊕⊕◯◯ LOW[3,2]	-	Mean 27.28 (SD 2.51)	MD 2.23 lower (3.49 lower to 0.97 lower)
APQ:	102	⊕⊕◯◯	-	Mean	MD 0.47 lower

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Parent training (95% CI)
involvement (Scale from 10 to 50; higher better)	(1 RCT)	LOW[3,2]		31.21 (SD 6.49)	(3.29 lower to 2.35 higher)
APQ: poor monitoring/ supervision (Scale from 10 to 50; lower better)	102 (1 RCT)	⊕⊕◯◯ LOW[3,2]	-	Mean 10.48 (SD 2.04)	MD 0.72 higher (0.22 lower to 1.67 higher)
APQ: corporal punishment (Scale from 3 to 15; lower better)	102 (1 RCT)	⊕⊕◯◯ LOW[3,2]	-	Mean 4.84 (SD 2.08)	MD 0.29 lower (1.21 lower to 0.63 higher)
Drop-out (all cause)	308 (2 RCTs)	⊕⊕◯◯ LOW [1,2,3]	RR 1.12 (0.76 to 1.64)	246 per 1,000	30 more per 1,000 (59 fewer to 157 more)

1.Sleed (2013) - no blinding
2,95% CI includes both no effect and clinically significant harm or benefit
3.Menting (2014) - unclear randomisation method and no blinding

1

6.2.122 Yoga for promoting mental health and wellbeing

3 One RCT (N = 167) met the eligibility criteria for this review. Bilderbeck et al. (2013) was a
4 study of the impact of a ten-week yoga course on the psychological wellbeing of prisoners.
5 Prisoners diagnosed with psychiatric illness were excluded. An overview of the trial included
6 in the analysis can be found in Table 50. Further information about both included and
7 excluded studies can be found in Appendix L.

8 Summary of findings can be found in Table 51. The full GRADE evidence profiles and
9 associated forest plots can be found in Appendices N and O, respectively.

10 **Table 50 Study information for studies included in the analysis of Yoga versus Waitlist**
11 **Control**

	Yoga versus Waitlist Control
Total no. of studies (N[1])	1 (167)
Study ID	Bilderbeck 2013
Study design	RCT
Country	UK
Diagnosis	No MH problems reported
Age (mean)	36.1
Gender (% female)	7
Ethnicity (% white)	80
IQ (mean)	Not reported
Offence	Not reported
Treatment length (weeks)	Not reported

	Yoga versus Waitlist Control
Intervention	Yoga classes consisting of a standardised set of hatha yoga postures and stretches and relaxation breathing exercises during the final 10-20 minutes of each class
Comparison	Waitlist
Format	Group
Dose/intensity (hours)	10(2/week)
Intervention setting	Prison
Length of treatment (weeks)	10 weeks
Continuation phase (length and inclusion criteria)	1 week (post-intervention assessments conducted 1 week after completion of course)
Notes. N= total number of participants; [1]Number randomised. NR- Not reported	

Table 51: Summary of findings table for yoga versus waitlist control for promoting mental health and wellbeing

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with waitlist control	Risk difference with Yoga
Positive affect (PANAS) (Scale from 10 to 50; higher better)	100 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 31.22 (SD 7.56)	MD 5.94 higher (2.91 higher to 8.97 higher)
Negative affect (PANAS) (Scale from 10 to 50; lower better)	100 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 19.15 (SD 8.16)	MD 4.13 lower (6.80 lower to 1.46lower)
Perceived stress (PSS) (Scale from 0 to 40; lower better)	100 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 16.07 (SD 7.79)	MD 4.67 lower (7.65 lower to 1.69 lower)
Psychological distress (BSI) (Scale from 0 to 212; lower better)	100 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	Mean 37.09 (SD 29.44)	MD 12.60 lower (22.82 lower to 2.38 lower)
Drop-out (all cause)	167 (1 RCT)	⊕⊕◯◯ LOW [1,2,3]	RR 1.54 (1.04 to 2.28)	313 per 1,000	169 more per 1,000 (13 more to 400 more)
1.Bilderbeck (2013) - no blinding, attrition bias (significantly higher dropout with yoga) *2.Study was an exploratory trial - without sample size calculation* *3.95% CI includes the possibility that the benefit is less than the minimum important difference*					

6.2.143 Meditation for promoting mental health and well-being

One RCT (N = 33) met the eligibility criteria for this review. Sumter et al. (2009) was a US study of meditation in a residential detention facility for nonviolent female probationers. An overview of the trial included in the meta-analysis can be found in Table 52. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 53. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 52 Study information table for studies included in the meta-analysis of meditation

	Meditation versus Treatment as usual
Total no. of studies (N[1])	1(33)
Study ID	Sumter 2009
Study design	RCT
Country	US
Diagnosis	No MH problem reported
Age (mean)	Not reported
Gender (% female)	100
Ethnicity (% white)	58
IQ (mean)	Not reported
Offence	Not reported
Treatment length (weeks)	Not reported
Intervention	Meditation exercise involving basic instruction and guidance in the importance of posture (erect spine), counting in breaths and out breaths, repeating a phrase or mantra (which was self-selected), walking meditation and moving meditation (simple yoga postures). There were also group discussions.
Comparison	Treatment as usual (detainees at the facility were not allowed to talk unless permission was granted. During the time when the experimental group practiced meditation, the control group continued with their regular daily activities. These activities typically consisted (at the time in the afternoon) of free time that could be used for exercise, reading, or being outside in the yard. Otherwise, both groups experienced a similar daily routine.)
Format	Group
Dose/intensity (hours)	17.5 (2.5/week)
Intervention setting	Residential detention facility for probationers
Length of treatment (weeks)	7
Continuation phase (length and inclusion criteria)	0

Notes. N= total number of participants; [1]Number randomised.

Table 53: Summary of findings table for meditation for promoting mental health and well-being

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Meditation[3]
Desire to throw things or hit people within past month (study-specific measure; scale from 0 to 4; lower better)	33 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	-	SMD 1.01 lower (1.73 lower to 0.28 lower)
Being bothered by nail biting within past month	33 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	-	SMD 1.18 lower

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Meditation[3]
(study-specific measure; scale from 0 to 4; lower better)					(1.91 lower to 0.44 lower)
Feelings of guilt within past month (study-specific measure; scale from 0 to 4; lower better)	33 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	-	SMD 0.42 lower (1.11 lower to 0.27 higher)
Feelings of hopelessness within past month (study-specific measure; scale from 0 to 4; lower better)	33 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	-	SMD 0.06 lower (0.74 lower to 0.63 higher)
Being bothered by sleeping difficulties within past month (study-specific measure; scale from 0 to 4; lower better)	33 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	-	SMD 0.28 lower (0.96 lower to 0.41 higher)

1. Sumter (2009) - no blinding, unclear allocation concealment
2. Imprecision – 95% CI includes both no difference and clinically important harm or benefit
3. It was not possible to calculate MD, so SMD is reported.

6.2.114 Physical exercise programmes versus exercise as usual for promoting mental health and well-being

One RCT (N =75) met the eligibility criteria for this review. Battaglia et al. (2015) was a three arm trial comparing two different physical exercise programs with exercise as usual in an Italian prison. An overview of the trial can be found in Table 54. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 55**.** The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 54 Study information table for studies included in the analysis of physical exercise programmes versus exercise as usual

	Physical exercise programmes versus exercise as usual
Total no. of studies (N[1])	1 (75)
Study ID	Battaglia 2015
Study design	RCT
Country	Italy
Diagnosis	No MH problem reported
Age (mean)	32 years
Gender (% female)	0%
Ethnicity (% white)	Not reported
IQ (mean)	Not reported
Offence	Not reported
Intervention	Cardiovascular plus resistance training or high-intensity strength training
Comparison	Exercise as usual
Format	Face to face

	Physical exercise programmes versus exercise as usual
Dose/intensity (hours)	1 hour session twice a week
Intervention setting	Prison
Length of treatment (weeks)	39 weeks
Continuation phase (length and inclusion criteria)	0
Notes. N= total number of participants; [1]Number randomised.	

Table 55: Summary of findings table for physical exercise programmes versus exercise as usual

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with exercise as usual	Risk difference with Physical exercise programme
Change in Symptom Checklist-90-Revised (SCL-90-R) Global Severity Index (GSI) - CRT or HIST exercise programme versus exercise as usual between baseline and 39 weeks (Scale from 0 to 4; lower better)	64 (1 RCT)	⊕⊕◯◯ LOW 1	-	The mean change in Symptom Checklist-90-Revised (SCL-90-R) Global Severity Index (GSI) was 0.03 higher (SD 0.06)	MD 0.17 lower (0.21 lower to 0.12 lower)
Change in Symptom Checklist-90-Revised (SCL-90-R) Positive Symptom Total (PST) - CRT or HIST exercise programme versus exercise between baseline and 39 weeks (Scale from 0 to 90; lower better)	64 (1 RCT)	⊕⊕◯◯ LOW 1	-	The mean change in Symptom Checklist-90-Revised (SCL-90-R) Positive Symptom Total (PST) was 1 higher (SD 3.19)	MD 7.08 lower (9.15 lower to 5 lower)
Change in Symptom Checklist-90-Revised (SCL-90-R) Positive Symptom Distress Index (PSDI) - CRT or HIST exercise programme versus exercise as usual between baseline and 39 weeks (Scale from 0 to 4; lower better)	64 (1 RCT)	⊕⊕◯◯ LOW 1	-	The mean change in Symptom Checklist-90-Revised (SCL-90-R) Positive Symptom Distress Index (PSDI) was 0.07 higher (SD 0.12)	MD 0.33 lower (0.41 lower to 0.25 lower)
1. Battaglia 2015 - unclear allocation concealment, no blinding, per-protocol analysis					

6.2.2 Economic evidence

No studies assessing the cost effectiveness of interventions to promote mental health and wellbeing in adults in contact with the criminal justice system (including environmental

adaptations and individual- and population-based psychoeducational interventions) were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.2.3 Clinical evidence statements

6.2.3.1 Parent training for parent-child attachment for women with sub-threshold symptoms

Low quality evidence from a single study (N=115) found no clinically important effects of parent training on depression as measured by mean scores on the Center for Epidemiological Studies Depression Scale (CES-D; SMD -0.16 [-0.53, 0.21]) or the number of participants with symptoms of depression (CES-D=>16; RR 0.79 [0.51, 1.21]) for women with sub-threshold symptoms of depression resident in specialized mother and baby units in prison.

Low quality evidence from single study data (N=88-109) also showed no clinically important effects of parent training on:

- measures of mother-child attachment (SMD 0.27 [-0.11, 0.64])
- mother-child interaction (SMD -0.38 [-0.81, 0.05]; SMD -0.06 [-0.48, 0.37; SMD -0.08 [-0.46, 0.31],; for dyadic attunement, parent positive engagement and child involvement respectively)
- or maternal perceptions of their child (SMD 0.44 [-0.04, 0.91]; SMD -0.12 [-0.58, 0.35]; SMD -0.29 [-0.74, 0.16]; SMD 0.04 [-0.41, 0.49]; for positive and negative perceptions, and intensity and impact of problem behaviour respectively)

A single study (N=102-3) found low quality inconsistent evidence for effects of parent training on maternal perceptions of parenting as measured by the Alabama Parenting Questionnaire (APQ) for women recently released from prison.

- For the Inconsistent discipline sub-scale, a clinically important effect was found in favour of parent training (SMD -0.81 [-1.27, -0.34]).
- However, for the Positive parenting subscale there was a clinically important effect in favour of the treatment as usual control (SMD -0.66 [-1.12, -0.20]).
- While for the other three subscales no clinically significant effects were found (SMD -0.08 [-0.53, 0.37]; SMD 0.33 [-0.13, 0.78]; SMD -0.15 [-0.60, 0.30]; for Involvement, Poor monitoring/supervision and Corporal Punishment respectively).

Low quality evidence from two studies (N=308) did not indicate a clinically significant difference in drop-out between the parent training and treatment as usual conditions (RR 1.12 [0.76, 1.64]) for women in prison or recently released from prison.

6.2.3.2 Yoga for promoting mental health and wellbeing

A single study (N=100) found very low quality evidence for clinically significant effects of yoga on increasing positive affect (SMD 0.77 [0.36, 1.18]), decreasing negative affect (SMD -0.58 [-0.98, -0.18]) and reducing perceived stress (SMD -0.61 [-1.01, -0.21]), and a smaller effect for reducing psychological distress (SMD -0.47 [-0.87, -0.07) for participants in prison.

However, this study (N=167) also found that there was a clinically significant number of participants who dropped out of the study in the yoga condition (RR 1.54 [1.04, 2.28]).

6.2.3.3 Meditation for promoting mental health and well-being

A single study (N=33) found very low quality evidence for a clinically important effect of meditation for non-violent female probationers in a residential detention facility on reducing

the desire to throw things or hit people (SMD -1.01 [-1.73, -0.28]) and on being bothered by nail biting (SMD -1.18 [-1.91, -0.44]) as measured by a study-specific scale

However, confidence in these effect estimates was very low due to serious risk of bias (unclear randomisation method and allocation concealment, lack of blinding and lack of a valid and reliable outcome measure) and very serious imprecision).

This same study found no evidence for clinically significant effects of meditation on feelings of guilt (SMD -0.42 [-1.11, 0.27]) or hopelessness (SMD -0.06 [-0.74, 0.63]) or being bothered by sleeping difficulties (SMD -0.28 [-0.96, 0.41]).

6.2.3.4 Physical exercise programmes for promoting mental health and well-being

Low quality evidence from one randomised controlled trial (N=75) indicated that a 9 month physical exercise programme had a clinically important beneficial effect on psychological wellbeing as measured by the Symptom Checklist-90-Revised (SCL-90-R) when compared to exercise as usual.

6.2.4 Economic evidence statements

No evidence on the cost effectiveness of interventions to promote metal health and wellbeing for adults who are in contact with the criminal justice system is available.

6.3 Recommendations and link to evidence

Recommendations	No recommendation made.
Relative values of different outcomes	Critical outcomes for this question included improvement in mental health and well-being (data on parenting outcomes could be seen as indicators of improved well-being) but there was no data on large scale population of service level interventions which are commonly the focus of health promotion interventions and the GC were uncertain of the value of the available data to inform decisions about population (i.e. those in the criminal justice system) or service level interventions.
Trade-off between clinical benefits and harms	The GC were of the view that most health promotion programmes were unlikely to have any significant harms associated with their use but there was a concern that the delivery of such programmes may increase the threshold or reduce the likelihood of interventions being offered to people with established mental health problems. There were 6 trials of parenting interventions focussed on attachment difficulties. The studies were all of low quality and there were inconsistent indication of benefit e.g. on maternal mental health or mother-child attachment. One very low quality study of yoga which had high attrition reported some evidence of benefit on positive affect. One very low quality reported no positive benefits associated with meditation. One small low quality study reported a significant benefit on mental health symptoms following a physical exercise programme.
Trade-off between net health benefits and resource use	There was no data available on the cost effectiveness of health promotion interventions. However in the absence of good quality evidence for the effectiveness of these interventions the GC were concerned that it may lead to inappropriate use of resource.
Quality of evidence	The quality of the evidence ranged from low to very low. This was due to lack of blinding, inadequate randomisation, attrition bias and imprecision in the effect estimates of the included randomised trials. The GC considered

	that RCT evidence was required for this question given the cost implications of implementing mental health promotion across the entire criminal justice system. It was noted that confidence in the effect estimates was very low for the study looking at the effects of parent training on maternal perceptions of parenting as measured by the Alabama Parenting Questionnaire (APQ) for women recently released from prison. This was due to a very serious risk of bias (randomization was temporarily suspended for 24.7% of participants, participants and intervention administrators were non-blind, and unclear reliability and validity of the outcome measure) and very serious risk of imprecision (N<400 and wide confidence intervals). It was also noted that the confidence in the effect estimates was very low for the study looking at yoga for promoting mental health and well-being. This was due to a serious risk of bias (due to unclear allocation concealment, non-blind participants and intervention administrators and high risk of attrition bias) and serious imprecision.
Other considerations	The GC considered the evidence insufficient to make recommendations for mental health promotion or well-being interventions; nor was the evidence considered sufficient to support any specific recommendations on parent training, yoga, acupuncture, meditation or physical exercise programmes. The GC were aware of a number of mental health promotion programmes which had been developed outside of the criminal justice system. They did not consider that the sufficient evidence drawing on their knowledge and expertise to make an extrapolation from existing data on mental health promotion. They therefore decided not to make a recommendation. In view of the very limited evidence on health promotion and the absence of any specific high quality data on suicide prevention, the GC, being aware of the high prevalence of self-harm and completed suicides in the criminal justice system decided to make a research recommendation for the evaluation of different models on suicide prevention in the criminal justice system.

6.3.11 Research recommendations

3. What factors are associated with suicide attempts and completed suicides?

There is high prevalence of suicide attempts among people in contact with criminal justice system. While considering interventions on the prevention of self-harm among these population, it is important to examine the factors related to successful suicide. A retrospective analysis of observational studies of suicidal attempts and completed suicides using suicide as a definitive and measurable outcome should be performed to identify the prognostic factors for successful prevention.

6.4 Review question: What interventions are effective, or what modifications are needed to psychological, social, pharmacological or physical interventions recommended in existing NICE guidance, for adults in contact with the criminal justice system who have:

- alcohol-use disorders?
- antenatal or postnatal mental health problems [for women]?
- antisocial personality disorder?

- attention deficit hyperactivity disorder?
- autism?
- bipolar disorder?
- borderline personality disorder?
- challenging behaviour or mental health problems [for adults with learning disabilities]?
- delirium?
- dementia?
- depression (with or without a coexisting chronic physical health problem)?
- eating disorders?
- generalised anxiety disorder and panic disorder (with or without agoraphobia)?
- obsessive-compulsive disorder and body dysmorphic disorder?
- post-traumatic stress disorder?
- psychosis (with or without coexisting substance misuse) or schizophrenia?
- self-harmed (self-harming)?
- social anxiety disorder?
- substance misuse disorders?
- violent and aggressive behaviour [for adults with mental disorders]?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 56. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 56: Clinical review protocol summary for the review of psychological, social, pharmacological or physical interventions that are effective for adults in contact with the criminal justice system

Component	Description
Population	Adults with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Intervention(s)	• Psychological and social interventions • Pharmacological interventions • Combined psychological or social and pharmacological interventions • Support and education interventions aimed at promoting mental health and wellbeing
Comparison	• Treatment as usual • No treatment • Waitlist control • Placebo • Any alternative management strategy
Outcomes	1) Substance misuse disorders • Critical – Abstinence and reduction in drug or alcohol use; Offending and re-offending outcomes • Important – Mental health outcomes, Adaptive functioning (for example, employment status, development of daily living and interpersonal skills and quality of life); Service utilisation (e.g. hospital admission, engagement with services; Self-harm and suicide 2) Mental health problems

Component	Description
	• Critical – Remission, relapse, symptomatology; Offending and re-offending • Important –Adaptive functioning (for example, employment status, development of daily living and interpersonal skills and quality of life); Service utilisation (e.g. hospital admission, engagement with services); Self-harm and suicide
Study design	Systematic reviews and RCTs

6.4.11 Clinical evidence

6.4.12.1 Substance misuse

6.4.1.13.1 Psychological interventions

23 studies met the criteria for this review: Alemi 2010 (Alemi et al., 2010), Annis 1979 (Annis, 1979), Binswanger 2015 (Binswanger et al., 2015), Brown 1980 (Brown, 1980), Carroll 2006 (Carroll et al., 2006), Carroll 2012 (Carroll et al., 2012), Crane 2015b (Crane et al., 2015b), Easton 2000 (Easton et al., 2000), Easton 2007c (Easton et al., 2007c), Forsberg 2011 (Forsberg et al., 2011B), Gordon 2008 (Gordon et al., 2008), Gordon 2014 (Gordon et al., 2014), Kinlock 2007 (Kinlock et al., 2007), Kinlock 2009 (Kinlock et al., 2009), McKenzie 2012 (McKenzie et al., 2012), Miller 1975 (Miller, 1975)M, Proctor 2012 (Proctor et al., 2012), Sinha 2003 (Sinha et al., 2003), Stuart 2013 (Stuart et al., 2013), Villagara-Lanza 2013 (Villagra Lanza & Menendez, 2013), Villagara-Lanza 2014 (Lanza et al., 2014), Witkiewitz 2014 (Witkiewitz et al., 2014) and Zlotnick 2009 (Zlotnick et al., 2009).

The interventions studied included acceptance and commitment therapy (ACT), cognitive behavioural therapy (CBT), other cognitive and behavioural therapies, mindfulness-based approaches, counselling, motivational interviewing techniques, self-help and psychoeducation.

Cognitive behavioural therapy versus active intervention for substance misuse

3 RCTs (N=254) met the eligibility criteria for this review: Carroll 2012, Easton 2007c and Zlotnick 2009(Carroll et al., 2012; Easton et al., 2007c; Zlotnick et al., 2009).

An overview of the trials can be found in Table 57. Further information about both included and excluded studies can be found in Appendix N.

Summary of findings can be found in Table 58. The full evidence profiles and associated forest plots can be found in Appendices O and P.

The Zlotnick 2009 and Easton 2007c studies both describe 2-arm trials. The Zlotnick 2009 trial compared a variation of CBT tailored specifically to substance misuse (seeking safety) with treatment based upon the 12-step recovery model, whilst the Easton 2007c trial compared a variation of CBT designed for co-occurring substance misuse and domestic violence with the 12-step model. Finally, the Carroll 2012 trial was a 4-armed trial in which service users were allocated to one of the following conditions; standard CBT alone, CBT plus contingency management for adherence, contingency management for abstinence or CBT plus contingency management for abstinence. Only the CBT alone and CBT plus contingency management for adherence arms were included within this sub-review (CBT versus active intervention). The contingency interventions in Carroll 2012 were looked at under different sub-reviews (Contingency management versus Active intervention). The Carroll 2012 and Easton 2007c studies were both conducted in the community whilst the Zlotnick 2009 trial was conducted in a residential facility. The Easton 2007c trial intervention was delivered in a group setting whilst treatment in the other 2 studies was delivered individually (Carroll 2012) or a mixture of the two (Zlotnick 2009).

The evidence for this review was low to very low quality. No data was available for the outcomes of service utilisation, adaptive functioning or rates of self-injury.

Table 57: Study information table for trials included in the analysis of CBT versus active intervention for substance misuse

	CBT versus active intervention
Total no. of studies (N[1])	3 (254)
Study ID	(1) Carroll 2012 (2) Easton 2007c (3) Zlotnick 2009
Study design	RCT
Country	USA
Diagnosis	(1) Drug misuse (2) Alcohol misuse (3) Mixed
Age (mean) years	(1)25.7 (2)38 (3)34.6
Sex (% female)	(1)15.7 (2)0 (3)100
Ethnicity (% white)	(1)18.9 (3)47% (2)49%
Setting	(1, 2) Community (3) Residential
Coexisting conditions/other treatments received during study	(1, 2, 3) NA
Treatment length (weeks)	(1, 2) 12 weeks (3) 18-20 weeks
Intervention (mean dose; mg/day)	(1) CBT, 50 mins per week (2) CBT for substance abuse and domestic violence (SADV), 1.5 hours per week (3) Seeking Safety, group 90 mins 3x per week, individual 1 hour per week
Delivery method	(1) Individual (2) Group of up to 10 people (3) Mixed
Comparison	1) CBT plus contingency management for adherence (provision of prizes contingent upon session attendance and homework completion) (2, 3) 12 step programme
Note. N= total number of participants; NA = Not applicable [1]Number randomised;	

Table 58: Summary of findings table for the analysis of CBT versus active intervention for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control [mean(±SD)]	Risk difference with CBT versus active intervention (95% CI)

Days using cannabis (during treatment) - Self-report	95 (1 study)	⊕⊕⊖⊖ LOW[1]	-	31.9 (±38) days	MD 10.15 days higher (6.63 lower to 26.93 higher)
Days using cannabis (during treatment) - Urine test	95 (1 study)	⊕⊕⊖⊖ LOW[2]	-	57.1 (±37) days	MD 17.13 days higher (0.92 to 33.34 higher)
Days with positive urine test (during treatment)	75 (1 study)	⊕⊖⊖⊖ VERY LOW[3,2]	-	0.35 (±0.48) days	MD 0.3 days higher (2.23 lower to 2.15 higher)
Days with positive breathalyser test (during treatment)	75 (1 study)	⊕⊖⊖⊖ VERY LOW[3,1]	-	0.22 (±6) days	MD 0.04 lower (0.46 lower to 0.44 higher)
Days abstinent (during treatment) - Alcohol	71 (1 study)	⊕⊖⊖⊖ VERY LOW[3]	-	79.8 (±23.1) days	MD 10.40 higher (1.53 to 19.27 higher)
Days abstinent (during treatment) - Drugs	71 (1 study)	⊕⊖⊖⊖ VERY LOW[3]	-	96.1 (±14.5) days	MD 0.70 higher (0.41 lower to 6.12 higher)
Addiction Severity Index (ASI-6): alcohol composite score (Scale from 0 to 9; lower better)	44 (1 study) 26-38 weeks	⊕⊖⊖⊖ VERY LOW[4,1]	-	The mean addiction severity index (ASI-6): alcohol composite score in the control group was 0.2 (±0.23)	MD 0.10 lower (0.22 lower to 0.02 higher)
Addiction Severity Index (ASI-6): drug composite score (Scale from 0 to 9; lower better)	44 (1 study) 26-38 weeks	⊕⊖⊖⊖ VERY LOW[4]	-	The mean addiction severity index (asi-6): drug composite score in the control group was 0.18 (±0.11)	MD 0.02 lower (0.09 lower to 0.05 higher)
Weeks abstinent	44 (1 study) 26-38 weeks	⊕⊖⊖⊖ VERY LOW[4,1]	-	The mean weeks abstinent in the control group was 7.6 (±5.2) weeks	MD 1.30 weeks lower (4.4 lower to 1.8 higher)
Reincarceration	44 (1 study) 26-38 weeks	⊕⊖⊖⊖ VERY LOW[4,1]	RR 0.51 (0.2 to 1.27)	429 per 1000	210 fewer per 1000 (from 343 fewer to 116 more)

1 95% CI includes both no effect and the minimal important difference
2 95% CI includes the minimal important difference
3 high risk of performance bias. unclear risk for allocation concealment, detection, attrition, reporting and other bias
4 high risk of concealment bias, unclear risk on all other dimensions

Cognitive behavioural therapy versus wait-list control for substance misuse

1 RCT (N=27) met the eligibility criteria for this review: Villagara-Lanza 2014 (Lanza et al., 2014).

An overview of the trial can be found in Table 59. Further information about both included and excluded studies can be found in Appendix N.

Summary of findings can be found in Table 60. The full evidence profiles and associated forest plots can be found in Appendices O and P.

The Villagara-Lanza 2014 study was a 3-arm trial, with groups receiving CBT, ACT or no treatment (waitlist control). The comparison of CBT and waitlist control group is described here. Treatment was delivered in a group format within a prison setting.

The evidence for this review was low to very low quality. No data was available for the outcomes of service utilisation, adaptive functioning or rates of self-injury.

Table 59: Study information table for trials included in the analysis of CBT versus waitlist control

	CBT versus waitlist control
Total no. of studies (N[1])	1 (27)
Study ID	Villagara-Lanza 2014
Study design	RCT
Country	Spain
Diagnosis	Substance misuse disorder
Age (mean)	33.2 years
Sex (% female)	100%
Ethnicity (% white)	Not reported
Setting	Prison
Coexisting conditions/other treatments received during study	Educational programme provided as standard by the prison
Treatment length (weeks)	16 weeks
Intervention (mean dose; mg/day)	CBT, 1.5 hours per week
Delivery method	Group
Comparison	Waitlist control
Note. N= total number of participants; [1] Number randomised.	

Table 60: Summary of findings table for the analysis of CBT versus wait-list control for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control [mean(±SD)]	Risk difference with CBT versus waiting-list control (95% CI)
Addiction Severity Index (ASI-6): alcohol composite score (Scale from 0 to 9; lower better)	27 (1 study)	⊕⊖⊖⊖ VERY LOW[1]	-	0.42 (±0.06)	MD 0.01 lower (0.05 lower to 0.03 higher)
Addiction Severity Index (ASI-6): drug composite score (Scale from 0 to 9; lower better)	27 (1 study)	⊕⊖⊖⊖ VERY LOW[2]	-	0.44 (±0.04)	MD 0.03 lower (0.07 lower to 0.01 higher)
Abstinent in previous 3 months (6 month follow-up)	27 (1 study)	⊕⊕⊖⊖ LOW[1]	RR 1.38 (0.3 to 6.25)	182 per 1000	69 more per 1000 (from 127 fewer to 955 more)

1 high risk for performance bias, high risk for 'other bias'
2 95% CI includes the minimal important difference

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

Acceptance and commitment therapy versus cognitive behavioural therapy for substance misuse

1 RCT (N=30) met the eligibility criteria for this review: Villagara-Lanza 2014 (Lanza et al., 2014).

An overview of the trial can be found in Table 61. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 62. The full evidence profiles and associated forest plots can be found in Appendices N and O. respectively.

Villagara-Lanza 2014 was a 3-armed trial comparing Acceptance and Commitment Therapy (ACT), CBT and control. The comparisons of ACT versus control and CBT versus control are detailed elsewhere in this chapter. Treatment was delivered in groups in prison.

The evidence for this review was low to very low quality. No evidence was available for the outcomes of offending and reoffending, service utilisation, adaptive functioning or rates of self-injury.

Table 61: Study information table for trials included in the analysis of ACT versus CBT for substance misuse in adults within the criminal justice system

	ACT versus CBT
Total no. of studies (N[1])	1 (30)
Study ID	Villagara-Lanza 2014
Study design	RCT
Country	Spain
Diagnosis	Substance misuse disorders
Age (mean)	33.2 years
Sex (% female)	100
Ethnicity (% white)	Not reported
Setting	Prison
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	16
Intervention (mean dose; mg/day)	Acceptance and commitment therapy; 1.5 hour sessions once per week
Delivery method	Group
Comparison	CBT; 1.5 hour sessions once per week for 16 weeks
Note. N= total number of participants; [1] Number randomised	

Table 62: Summary of findings for the analysis of ACT versus CBT for substance misuse in adults within the criminal justice system

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT	Risk difference with ACT (95% CI)

Addiction Severity Index (ASI-6): alcohol composite score (Scale from 0 to 9; lower better)	30 (1 study)	⊕⊕⊖⊖ LOW[1,2]	-	The mean addiction severity index (asi-6): alcohol composite score in the control group was 0.41 (SD 0.05)	MD 0.04 lower (0.07 to 0.01 lower)
Addiction Severity Index: drug composite score (Scale from 0 to 9; lower better)	30 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3]	-	The mean addiction severity index: drug composite score in the control group was 0.41 (SD 0.07)	MD 0.01 lower (0.05 lower to 0.03 higher)
Abstinence from drugs	30 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW[1,3]	RR 1.71 (0.60 to 4.86)	250 per 1000	178 more per 1000 (from 100 fewer to 965 more)

[1] High risk of performance and detection bias, all other domains low risk
[2] optimal information size criterion not met
[3] confidence interval includes both clinically significant benefit and harm

Acceptance and commitment therapy versus waitlist for substance misuse

2 RCTs (N=61) met the eligibility criteria for this review: Villagara-Lanza 2013(Villagra Lanza & Menendez, 2013) and Villagara-Lanza 2014(Lanza et al., 2014).

An overview of the trials can be found in Table 63. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 64. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

The Villagara-Lanza 2013 study was a 2-armed trial comparing ACT and a waitlist control group. The Villagara-Lanza 2014 study was a 3-armed trial comparing ACT, CBT and waitlist control groups. The comparisons of ACT versus CBT and CBT versus control are included elsewhere within the chapter. Both trials delivered interventions in groups within a prison setting.

The evidence for this review was low to very low quality. No data were available for the outcomes of offending and reoffending, service utilisation, adaptive functioning or rates of self-injury.

Table 63: Study characteristics for the analysis of ACT versus waitlist control for substance misuse in adults within the criminal justice system

	ACT versus waitlist control
Total no. of studies (N[1])	2 (61)
Study ID	(1) Villagara-Lanza 2013 (2) Villagara-Lanza 2014
Study design	RCT
Country	(1, 2) Spain
Diagnosis	(1, 2) Substance misuse disorders
Age (mean)	(1) 32.0 years[2] (2) 33.2 years[3]
Sex (% female)	(1, 2) 100

	ACT versus waitlist control
Ethnicity (% white)	(1, 2) Not reported
Setting	(1, 2) Prison
Coexisting conditions/other treatments received during study	(1, 2) Not reported
Treatment length (weeks)	(1, 2) 16
Intervention (mean dose; mg/day)	(1, 2) Acceptance and commitment therapy; 1.5 hour sessions once per week
Delivery method	(1, 2) Group
Comparison	(1, 2) Waitlist control
Note. N= total number of participants; [1] Number randomised	

Table 64: Summary of findings for the analysis of ACT versus waitlist control for substance misuse in adults within the criminal justice system

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects Risk with Waitlist	Risk difference with ACT (95% CI)
Addiction Severity Index (ASI-6): alcohol composite score (Scale from 0 to 9; lower better)	56 (2 studies) 42 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3]	-	The mean addiction severity index: alcohol composite score in the control groups ranged from 0.40 to 0.42	SMD 0.60 SDs lower (1.72 SDs lower to 0.53 SDs higher)
Addiction Severity Index (ASI-6): drug composite score (Scale from 0 to 9; lower better)	52 (2 studies) 42 weeks	⨁⊖⊖⊖ VERY LOW[1,3]	-	The mean addiction severity index: drug composite score in the control groups ranged from 0.40 to 0.42	SMD 0.44 SDs lower (1.19 SDs lower to 0.3 SDs higher)
Abstinent from drugs in previous 3 months	25 (1 study) 42 weeks	⨁⨁⊖⊖ LOW[3,4]	RR 2.36 (0.59 to 9.48)	182 per 1000	247 more per 1000 (from 75 fewer to 1000 more)

[1] high risk of performance bias, unclear or mixed risk on three other facets
[2] I^2 =75%, random effects model used and outcome downgraded for inconsistency
[3] high risk of performance bias, unclear or mixed risk on two other facets
[4] confidence interval includes both clinically significant benefit and harm

Mindfulness-based relapse prevention (MBRP) versus Cognitive Behavioural Therapy (CBT) for substance misuse

1 RCT (N=105) met the eligibility criteria for this review: Witkiewitz 2014(Witkiewitz et al., 2014).

An overview of the trial can be found in Table 65. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 66. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

This was a 2-armed trial comparing mindfulness-based CBT with a CBT-based relapse prevention intervention. The authors hypothesised that the addition of a mindfulness-based component would help service users to identify when they were on 'automatic pilot' and

accordingly assist them in identifying triggers for cravings and help prevent relapse. The interventions were provided in a group format in a residential detention setting.

The evidence for this review was very low quality. No data were available for the outcomes of offending and reoffending, service utilisation, adaptive functioning or rates of self-injury.

Table 65: Study characteristics for the analysis of mindfulness-based relapse prevention versus active intervention for substance misuse in adults within the criminal justice system

	Mindfulness-based relapse prevention (MBRP) versus CBT
Total no. of studies (N[1])	1 (105)
Study ID	Witkiewitz 2014
Study design	RCT
Country	USA
Diagnosis	Substance misuse disorders
Age (mean)	34.1 years
Sex (% female)	100
Ethnicity (% white)	42
Setting	Residential detention facility
Coexisting conditions/other treatments received during study	NR
Treatment length (weeks)	8
Intervention (mean dose; mg/day)	MBCBT; 2x 50 min sessions per week
Delivery method	Group
Comparison	CBT; 2 x 50 min sessions per week
Note. N= total number of participants; NR=Not reported; RCT=randomised controlled trial [1] Number randomised	

Table 66: Summary of findings table for the analysis of MBRP versus CBT

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Active intervention (CBT)	Risk difference with Mindfulness-based relapse prevention (95% CI)
Drug-use days	54 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	The mean drug-use in the control group was 0.5 days	MD 0.46 lower (1.16 lower to 0.24 higher)
Short Inventory of Problems (SIP) follow-up (Scale from 0 to 45; lower better)	54 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	The mean short inventory of problems (SIP) follow-up in the control groups was 21.9	MD 7.30 lower (15.81 lower to 1.21 higher)
Addiction Severity Index: family-social composite score	54 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	The mean addiction severity index: family-social composite score in the control groups was 0.14	MD 0.01 lower (0.09 lower to 0.07 higher)
Addiction Severity Index: legal composite score	54 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	The mean addiction severity index: legal composite score in the control group was	MD 0.31 lower (0.45 to 0.17 lower)

			0.35	
Addiction Severity Index: medical composite score	54 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	The mean addiction severity index: medical composite score in the control group was 0.32	MD 0.20 lower (0.37 to 0.03 lower)
Addiction Severity index: psychiatric compose score	54 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	The mean addiction severity index: psychiatric compose score in the control group was 0.34	MD 0.11 lower (0.22 lower to 0.00 higher)

[1] high risk of bias from blinding and other factors, unclear risk of bias on 5 other domains
[2] optimal information size criterion not met

Contingency management versus active intervention for substance misuse

4 RCTs (N=461) met the eligibility criteria for this review: Carroll 2006(Carroll et al., 2006), Carroll 2012(Carroll et al., 2012), Prendergast 2015(Prendergast et al., 2015) and Sinha 2003(Sinha et al., 2003).

An overview of the trials can be found in Table 67. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 68. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively

The Carroll 2006 study was a 4-armed trial comparing CBT plus motivational enhancement and contingency management, motivational enhancement plus CBT only, drug counselling plus contingency management and drug counselling alone. Only the contingency management plus drug counselling and drug counselling alone arms are included within this review. The Carroll 2012 trial was a 4-armed trial in which service users were allocated to one of the following conditions; standard CBT alone, CBT plus contingency management for adherence, contingency management for abstinence or CBT plus contingency management for abstinence. Here the 3 contingency management arms are combined and compared with the CBT arm. By combing the 3 contingency management arms to create a single pair-wise comparison with CBT alone randomisation was preserved, the interventions being similar enough to not downgrade this evidence for indirectness. The Sinha 2003 study was a 2-armed trial, where participants received either contingency management plus motivational enhancement therapy, or motivational enhancement therapy alone. The Prendergast 2015 study was also a 2-armed trial comparing contingency management with a psychoeducational intervention called ‘attendance education group’.

The evidence for this review was low to very low quality. No data was available for the outcomes of adaptive functioning, offending and reoffending or rates of self-injury.

Table 67: Study information table for trials included in the analysis of contingency management versus active intervention for substance misuse

	Contingency management versus active intervention
Total no. of studies (N[1])	4 (461)
Study ID	(1) Carroll 2006 (2) Carroll 2012 (3) Prendergast 2015 (4) Sinha 2003
Study design	RCT
Country	(1 to 4) USA

	Contingency management versus active intervention
Diagnosis	(1, 2, 3) Drug misuse (4) Substance misuse
Age (mean)	(1) 21.0 years (2) 25.7 years (3) 20.6 years (4) 43.6 years
Sex (% female)	(1) 10 (2) 15.7 (3) 7 (4) Not reported
Ethnicity (% white)	(1, 3) Not reported (2) 18.9 (4) 13.4
Setting	(1) Not reported (2, 3) Community (4) Community and inpatient
Coexisting conditions/other treatments received during study, if any	(1) Drug counselling (2) CBT in 2 arms (3) Motivational enhancement therapy
Treatment length (weeks)	(1) 8 weeks (2) 12 weeks (3) Not reported (4) 22 weeks
Intervention (mean dose; mg/day)	Contingency management: (1, 2) weekly (3, 4) Not reported
Delivery method	(1, 2) Individual (3, 4) Not reported
Comparison	(1) Drug counselling (2) CBT (3) Psychoeducation (4) Motivational enhancement therapy
Note. N= total number of participants; [1] Number randomised	

Table 68: Summary of findings table for the analysis of contingency management versus active intervention for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with active intervention	Risk difference with Contingency management versus active intervention (95% CI)
Days using cannabis (during treatment) - Self-report	263 (2 studies)	⨁⨁⨀⨀ LOW[1,2]	-	Mean 0.72 (SD 0.32)	SMD 0.01 higher (0.24 lower to 0.26 higher)
Days using cannabis (during treatment) - Urine test	136 (1 study)	⨁⨀⨀⨀ VERY LOW[3,4]	-	Mean 0.7 (SD 0.41)	SMD 0.23 lower (0.57 lower to 0.11 higher)
Addiction Severity Index	65	⨁⨀⨀⨀	-	Mean 0.25 (SD	SMD 0.18 higher

(ASI): marijuana composite score - Post-treatment	(1 study)	VERY LOW[5,6]		0.25)	(0.32 lower to 0.67 higher)
Addiction Severity Index (ASI): marijuana composite score - Follow-up	65 (1 study)	⊕⊖⊖⊖ VERY LOW[6,7]	-	Mean 0.21 (SD 0.17)	SMD 0.11 higher (0.38 lower to 0.6 higher)
Days cannabis use per month - Post-treatment	65 (1 study)	⊕⊖⊖⊖ VERY LOW[6,7]	-	Mean 6.08 days (SD 7.21)	MD 4.89 days more (0.43 to 9.35 days more)
Days cannabis use per month - Follow-up	86 (1 study)	⊕⊖⊖⊖ VERY LOW[6,7]	-	Mean 8.32 days (SD 8.76)	MD 2.13 days more (2.05 days fewer 6.31 days more)
Participants still in treatment at follow-up	165 (1 study) 52 weeks	⊕⊖⊖⊖ VERY LOW[6,7]	RR 0.81 (0.47 to 1.39)	268 per 1000	51 fewer per 1000 (from 142 fewer to 105 more)
No. of days in treatment	165 (1 study) 52 weeks	⊕⊖⊖⊖ VERY LOW [6,7]	-	The mean no. of days in treatment in the control group was 82 days	MD 3.00 lower (21.01 lower to 15.01 higher)

[1] One study high risk of bias, unclear for selection and reporting bias. Other study high risk for performance and unclear for allocation concealment and reporting bias
[2] Optimal information size criterion not met (N<400)
[3] high risk of bias, unclear for selection and reporting bias
[4] Optimal information size criterion not met (N<200) & CI includes both clinically significant harm and no effect
[5] performance bias is high risk, all other categories (except other) are unclear risk
[6] CI includes both clinically significant or harm and no effect
[7] high risk of blinding and outcome reporting bias, unclear risk of performance and concealment bias

Contingency management versus Treatment as usual (TAU) for substance misuse

1 RCT (N=20) met the eligibility criteria for this review: Miller 1975(Miller, 1975).

An overview of the trial can be found in Table 69. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 70. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

The Miller 1975 study was a 2-arm trial, with groups receiving contingency management or no treatment. Contingency management consisted of the provision of goods and services in exchange for sobriety. This trial was conducted in the community with treatment provided on an individual basis.

The evidence for this review was low quality. No data was available for the outcomes of mental health, service utilisation, adaptive functioning or rates of self-injury.

Table 69: Study information table for trials included in the analysis of contingency management versus treatment as usual substance misuse

	Contingency management versus treatment as usual
Total no. of studies (N[1])	1 (20)
Study ID	Miller 1975
Study design	RCT

	Contingency management versus treatment as usual
Country	USA
Diagnosis	Alcohol misuse
Age (mean) years	48.8
Sex (% female)	Not reported
Ethnicity (% white)	Not reported
Setting	Community
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	Not reported
Intervention	Contingency management (frequency and duration not reported)
Delivery method	Individual
Comparison	TAU (Participants in the TAU group had the same goods and services available to them as participants in the experimental group, but reinforcers were not provided for on a contingent basis)

Note. N= total number of participants; TAU = treatment as usual
[1] Number randomised

Table 70: Summary of findings table for the analysis of contingency management versus TAU for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Contingency management versus control (95% CI)
Arrests for public drunkenness	20 (1 study)	⊕⊕⊖⊖ LOW[1]	-	Mean 15.3 arrests (SD 11.8)	MD 1.70 fewer arrests (5.65 fewer to 2.25 more)

[1] Optimal information size criterion not met (N<200); 95% CI of effect includes both clinically significant benefit and no effect

Motivational enhancement therapy versus active intervention for substance misuse

3 RCTs (N=362) met the eligibility criteria for this review: Carroll 2006(Carroll et al., 2006), Easton 2000(Easton et al., 2000) and Stuart 2013(Stuart et al., 2013).

An overview of the trial can be found in Table 71. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 72. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Easton 2000 and Stuart 2013 were both 2-armed trials comparing motivational enhancement with psychoeducation. Carroll 2006 was a 4-armed trial with participants randomly allocated to receive one of the following: motivational enhancement therapy plus CBT and contingency management, motivational enhancement therapy plus CBT without contingency management, drug counselling plus contingency management or drug counselling alone. The drug counselling comparisons are described elsewhere within this chapter. The interventions in the Carroll 2006 trial were delivered individually.

The available data for this review were of very low quality. No data were available for the outcomes of offending and reoffending, adaptive functioning, rates of self-injury or service utilisation.

Table 71: Study characteristics table for the comparison of motivational enhancement therapy versus active intervention

	Motivational enhancement versus active intervention
Total no. of studies (N[1])	3 (362)
Study ID	(1) Carroll 2006 (2) Easton 2000 (3) Stuart 2013
Study design	(1,2,3) RCT
Country	(1,2,3) USA
Diagnosis	(1, 2) Drug misuse (3) Alcohol misuse
Age (mean) years	(1) 21.0 (2) 36.2 (3) 31.5
Sex (% female)	(1)10 (2, 3) 0
Ethnicity (% white)	(1) NR (2) 29.0 (3) 90.5
Setting	(1 to 3) NR
Coexisting conditions/other treatments received during study	(1) CBT + contingency management (2, 3) Psychoeducation
Treatment length (weeks)	(1)8 (2)10 (3) NR
Intervention	Motivational enhancement: (1, 2) 1 session (3) NR
Delivery method	(1) Individual (2, 3) NR
Comparison	(1) CBT plus contingency management (2, 3) Psychoeducation
Note. N= total number of participants; NR=Not reported; [1] Number randomised	

Table 72: Summary of findings for the analysis of motivational enhancement versus active intervention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with active intervention	Risk difference with Motivational enhancement therapy versus active intervention (95% CI)
Percentage of days abstinent from alcohol (self-report) - 3 month follow-	238 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Mean 65.1 % (SD 30.8)	MD 9.5 % more (2.51 to 16.49 % more)

up					
Percentage of days abstinent from alcohol (self-report) - 6 month follow-up	214 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	-	Mean 67.9 % (SD 29.2)	MD 4.8 % more (2.50 % fewer to 12.10 % more)
Percentage of days abstinent from alcohol (self-report) - 12 month follow-up	190 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3]		Mean 67.9 % (SD 29.2)	MD 0.8 % more (8.37 % fewer to 6.77 % more)
Percentage of days abstinent from alcohol and drugs - 3 month follow-up	238 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]		Mean 50.9 % (SD 37.1)	MD 9.7 % more (0.7 % more to 18.63 % more)
Percentage of days abstinent from alcohol and drugs - 6 month follow-up	214 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]		Mean 54.6 % (SD 35.1)	MD 5.2 % more (4.05 % fewer to 14.45 % more)
Percentage of days abstinent from alcohol and drugs - 12 month follow-up	190 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3]		Mean 58.6 % (SD 37.1)	MD 9.7 % more (0.7 % more to 18.63 % more)
Drinks per drinking days - 3 month follow-up	238 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]		Mean 9 drinks (SD 10.3)	MD 1.7 drinks fewer (3.75 fewer to 0.35 more)
Drinks per drinking days - 6 month follow-up	214 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]		Mean 7.3 drinks (SD 5.6)	MD 0.70 drinks more (0.93 fewer to 2.33 more)
Drinks per drinking days - 12 month follow-up	192 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3]		Mean 7.1 drinks (SD 5.1)	MD 0.30 drinks fewer (1.90 fewer to 1.33 more)
Percentage of days with cannabis use (during treatment)	136 (1 study)	⊕⊖⊖⊖ VERY LOW[3,4]		Mean 0.73% (SD 0.48)	SMD 0.1 lower (0.44 lower to 0.24 higher)
Percentage of urine tests positive for cannabis use (during treatment)	136 (1 study)	⊕⊖⊖⊖ VERY LOW[3,4]		Mean 0.7% (SD 0.5)	SMD 0.91 lower (1.27 to 0.56 lower)
Self-reported motivation to take steps to change substance abuse scores	27 (1 study)	⊕⊖⊖⊖ VERY LOW[2,5]		The mean self-reported motivation to take steps to change substance abuse scores in the control groups was 21	MD 4.10 higher (5.77 lower to 13.97 higher)

[1] High performance bias + unclear for 4 other bias types).
[2] Optimal information size criterion not met (N < 400)
[3] Attrition bias (more than 50% of sample)

[4] High performance bias + high attrition bias + unclear on 3 other types of bias.
[5] High risk of performance, detection and other bias, unclear selection and attrition bias

Motivational interviewing (MI) versus control or treatment as usual (TAU) for substance misuse

4 RCTs (N=492) met the eligibility criteria for this review: Alemi 2010(Alemi et al., 2010), Crane 2015b(Crane et al., 2015b), Davis 2003(Davis et al., 2003) and Forsberg 2011(Forsberg et al., 2011B).

3 trials were 2-armed (Alemi 2010, Crane 2015b and Davis 2003) and compared motivational interviewing with control or no treatment. Forsberg 2011 was a 3-armed trial that compared two different forms of motivational interviewing with treatment as usual. These two forms (with workshop training only or with peer group supervision in addition) have been combined here. All trials delivered the intervention of interest individually. Crane 2015b conducted their trial in the community whilst Davis 2003 and Forsberg 2011 conducted trials in prison settings.

An overview of the trials can be found in Table 73. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 74. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

The evidence for this review was low to very low quality. No data were available for the outcomes of adaptive functioning or rates of self-injury.

Table 73: Study characteristics table for the analysis of motivational interviewing or motivational feedback compared with control or treatment as usual

	MI versus control/TAU
Total no. of studies (N[1])	4 (492)
Study ID	(1) Alemi 2010 (2) Crane 2015b (3) Davis 2003 (4) Forsberg 2011
Country	(1, 2, 3) USA (4) Sweden
Diagnosis	(1 to 4) Substance misuse
Age (mean) years	(1, 4) NR (2) 33.1 (3) 45.7
Sex (% female)	(1) 62.0 (2) 0 (3) 2.7 (4) NR
Ethnicity (% white)	(1) 11.0 (2) 50.0 (3) 49.3 (4) NR

	MI versus control/TAU
Setting	(1) NR (2) Community (3) Initiated in prison, continued in the community (4) Prison
Coexisting conditions/other treatments received during study	(1 to 4) NR
Treatment length (weeks)	(1 to 4) Single session
Intervention (mean dose; mg/day)	Motivational interviewing: (1, 2) NR (3, 4) once per week
Delivery method	Individual: (1) online, or via telephone if online impossible (2, 3, 4) face-to-face
Comparison	(1) Treatment as usual (not specified) (2) No treatment (3) No motivational feedback (4) Usual planning interviewing
Note. N= total number of participants; NR=Not reported [1] Number randomised	

1

2 **Table 74: Summary of findings for the analysis of motivational interviewing or**
3 **motivational feedback versus control or treatment as usual**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Motivational interviewing/Motivational feedback versus control/TAU (95% CI)
Self-reported drug use - 1 month follow-up	79 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2,3]	RR 1.3 (0.86 to 1.95)	475 per 1000	142 more per 1000 (from 66 fewer to 451 more)
Self-reported days with drug use in past 30 days (10 month follow-up)	114 (1 study)	⨁⊖⊖⊖ VERY LOW[4,5]			SMD 0.04 higher (0.41 lower to 0.49 higher)
Urine test positive for drug use (during study period)	79 (1 study)	⨁⊖⊖⊖ VERY LOW[2,3,6]	RR 1.1 (0.62 to 1.96)	350 per 1000	35 more per 1000 (from 133 fewer to 336 more)
Self-reported alcohol use - 1 month follow-up	79 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2,3]	RR 1.3 (0.86 to 1.95)	475 per 1000	142 more per 1000 (from 66 fewer to 451 more)
Days with illegal activity in past 30 days (10 month follow-up)	103 (1 study)	⨁⊖⊖⊖ VERY LOW[4,5]			SMD 0.07 higher (0.4 lower to 0.53 higher)
Drop-out from subsequent treatment - binge drinking group	23 (1 study) 26 weeks	⨁⨁⊖⊖ LOW[7,8]	RR 0.27 (0.07 to 1.02)	667 per 1000	487 fewer per 1000 (from 620 fewer to 13 more)

Drop-out from subsequent treatment - no binge drinking group	35 (1 study) 26 weeks	⨁⨁⊖⊖ LOW[7,8]	RR 0.94 (0.3 to 2.91)	267 per 1000	16 fewer per 1000 (from 187 fewer to 509 more)
Number of subsequent treatment sessions attended - binge drinking group	19 (1 study) 26 weeks	⨁⨁⊖⊖ LOW[7,8]		The mean number of subsequent treatment sessions attended - binge drinking group in the control groups was 3.44	MD 11.16 higher (3.86 to 18.46 higher)
Number of subsequent treatment sessions attended - no binge drinking group	35 (1 study) 26 weeks	⨁⨁⊖⊖ LOW[7,8]		The mean number of subsequent treatment sessions attended - no binge drinking group in the control groups was 13	MD 1.65 lower (8.28 lower to 4.98 higher)
Speciality addiction clinic attendance	30 (1 study)	⨁⊖⊖⊖ VERY LOW[8,9]	RR 1.53 (0.59 to 3.99)	308 per 1000	163 more per 1000 (from 126 fewer to 920 more)

[1] high performance bias + high other bias + 3 unclear;
[2] very serious limitations (outcome)
[3] Optimal information size criterion not met (n = 79)
[4] high performance and detection bias.
[5] Optimal information size criterion not met (n = 114)
[6] high performance bias + high other bias + 3 unclear
[7] High risk of performance bias, unclear selection and detection bias
[8] Optimal information size criterion not met
[9] High risk of blinding, performance and detection bias, unclear selection and concealment bias

Group counselling versus treatment as usual for substance misuse

1 RCT (N=150) met the eligibility criteria for this review: Annis 1979(Annis, 1979).

An overview of the trial can be found in Table 75. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 76.The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

This was a 3-armed trial with service users being allocated to group counselling either with or without video-feedback, or treatment as usual. Outcomes with and without video feed-back are combined here. In the analysis the data for the two group counselling arms were pooled together and compared to the treatment as usual arms, preserving randomisation. When appraising quality it was considered that the two counselling arms used very similar interventions - differing only in the use of video feedback.

The available data for this review were of very low quality. No data were available for the outcomes of adaptive functioning, rates of self-injury or service utilisation.

Table 75: Study information table for trials included in the analysis of group counselling versus treatment as usual for substance misuse

	Group counselling versus TAU
Total no. of studies (N[1])	1 (150)
Study ID	Annis 1979
Study design	RCT

	Group counselling versus TAU
Country	Canada
Diagnosis	Substance misuse
Age (mean) years	24.5
Sex (% female)	0
Ethnicity (% white)	89
Setting	Prison
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	8 weeks
Intervention (mean dose; mg/day)	Group counselling; 9 hours 4 times per week
Delivery method (number per group)	Groups method (5/group)
Comparison	TAU (not specified)
Note. N= total number of participants; TAU=treatment as usual [1] Number randomised	

1

2 **Table 76: Summary of findings for the analysis of group counselling versus treatment**
3 **as usual**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Group counselling versus treatment as usual (95% CI)
Rearrest (12 month follow-up)	128 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.97 (0.7 to 1.35)	558 per 1000	17 fewer per 1000 (from 167 fewer to 195 more)
Number of reconvictions (12 month follow-up)	149 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	-	Mean 1 reconviction (SD 1.7) per participant	MD 0.10 fewer reconvictions (0.68 fewer to 0.48 more)
Reincarceration (12 month follow-up)	128 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 0.87 (0.5 to 1.5)	326 per 1000	42 fewer per 1000 (from 163 fewer to 163 more)
Days incarcerated (12 month follow-up)	149 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,3]	-	Mean 47.3 days (SD 85.7)	MD 0.30 days more (28.9 fewer to 29.5 more)
Self-reported drug use (12 month follow-up) - Marijuana	128 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 0.65 (0.44 to 0.96)	558 per 1000	195 fewer per 1000 (from 22 fewer to 313 fewer)
Self-reported drug use (12 month follow-up) - LSD	128 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 0.79 (0.37 to 1.67)	209 per 1000	44 fewer per 1000 (from 132 fewer to 140 more)
Self-reported drug use (12 month follow-up) - Speed	128 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 1.77 (0.62 to 5.05)	93 per 1000	72 more per 1000 (from 35 fewer to 377 more)
Self-reported drug use (12 month follow-up) - Heroin	128 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 1.18 (0.32 to 4.34)	70 per 1000	13 more per 1000 (from 47 fewer to 233 more)

[1] high risk of performance and detection bias. Unclear risk of remaining categories (other than 'other' bias)
[2] Imprecision: optimal information size criterion not met
[3] Confidence interval of effect includes both clinically significant benefit and harm

Self-help versus control for substance misuse

1 RCT (N=183) met the eligibility criteria for this review: Proctor 2012(Proctor et al., 2012).

An overview of the trial can be found in Table 77. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 78. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

The RCT had 2 arms, with service users randomly allocated either to either complete a self-help journal or to receive no intervention. The purpose of the journal was to assist service users to make a connection between their substance misuse and criminal activity, and was based upon the trans-theoretical model of change.

The data for this review was of a low quality. No data were available for the outcomes of mental health, service utilisation, adaptive functioning or rates of self-injury.

Table 77: Study information table for trials included in the analysis of self-help versus control for substance misuse

	Self-help versus control
Total no. of studies (N[1])	1 (183)
Study ID	Proctor 2012
Study design	RCT
Country	USA
Diagnosis	Drug misuse
Age (mean) years	36.6
Sex (% female)	0
Ethnicity (% white)	73
Setting	Prison
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	Not reported
Intervention (mean dose; mg/day)	Not reported
Delivery method	Individual
Comparison	No treatment
Note. N= total number of participants; [1] Number randomised	

Table 78: Summary of findings for the analysis of self-help versus control for substance misuse

Outcomes	No of Participants (studies)	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Self-help versus control (95% CI)

Follow up					
Subsequent bookings (12 month follow-up)	183 (1 study)	⊕⊕⊖⊖ LOW[1]	RR 0.76 (0.59 to 0.97)	659 per 1000	158 fewer per 1000 (from 20 fewer to 270 fewer)

[1] 183 participants were randomised but is unclear how many were assessed for eligibility

6.4.1.122 Pharmacological interventions

Opioid antagonists

These drugs bind to opioid receptors without activating them, preventing the body from responding to opioids and endorphins in the same way as they would otherwise. Naloxone is also used as an antidote drug in instances of opioid overdose, whilst Naltrexone can help reverse the long-term neurochemical after-effects of opioid misuse, which is hypothesised to help prevent relapse.

5 RCTs (N=394) met the eligibility criteria for this review: Cornish 1997(Cornish et al., 1997), Coviello 2010(Coviello et al., 2010), Hanlon 1977(Hanlon et al., 1977), (Lee et al., 2016; Lee et al., 2015) and Lobmaier 2010(Lobmaier et al., 2010).

Naloxone versus placebo

1 RCT (N=154) met the eligibility criteria for this review: Hanlon 1977

An overview of the trial can be found in Table 79. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 80.The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

The RCT by Hanlon 1977 had 2 arms and was conducted in a community setting.

The data for this review was of very low quality. No data were available for the outcomes of offending and reoffending, service utilisation, adaptive functioning and rates of self-injury.

Table 79: Study information table for trials included in the analysis of Naloxone versus placebo for drug misuse

	Naloxone versus placebo
Total no. of studies (N[1])	1 (154)
Study ID	Hanlon 1977
Study design	RCT
Country	USA
Diagnosis	Drug (opiate) misuse
Age (mean) years	26.3
Sex (% female)	0
Ethnicity (% white)	5
Setting	Community
Coexisting conditions/other treatments received during study	Not reported
Targeted behaviour	Drug misuse
Treatment length (weeks)	26
Intervention	Naloxone: average daily dose=757mg

	Naloxone versus placebo
(mean dose; mg/day)	
Comparison	Placebo: average daily dose=1068mg
Note. N= total number of participants; [1] Number randomised	

Table 80: Summary of findings table for the analysis of Naloxone versus placebo for drug misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Naloxone versus placebo (95% CI)
Discontinued medication	97 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 1.53 (0.72 to 3.23)	190 per 1000	101 more per 1000 (from 53 fewer to 425 more)
Number of urine tests positive during treatment	163 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.62 (0.22 to 1.72)	111 per 1000	42 fewer per 1000 (from 87 fewer to 80 more)

[1] unclear risk of bias for detection and performance bias.
[2] optimal information size criterion not met; confidence interval for the effect includes clinically significant benefit and harm

Naltrexone versus active intervention

4 RCTs (N=514) met the eligibility criteria for this review: Cornish 1997(Cornish et al., 1997), Coviello 2010(Coviello et al., 2010), Lee 2016 (Lee et al., 2016; Lee et al., 2015) and Lobmaier 2010(Lobmaier et al., 2010).

An overview of the trials can be found in Table 81. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 82. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

These were all 2-armed trials with Naltrexone treatment in one arm and either a psychosocial intervention (3 trials) or Methadone (1 trial) in the other.

The data for this comparison was of very low quality. No data were available for the outcomes of quality of life and adaptive functioning.

Table 81: Study information table for trials included in the analysis of naltrexone versus active intervention for drug misuse

	Naltrexone versus alternative opioid antagonist	Naltrexone plus psychological intervention versus other active intervention	Naltrexone, probation and counselling versus probation and counselling alone
Total no. of studies (N[1])	1 (44)	2 (419)	1 (51)
Study ID	Lobmaier 2010	(1) Coviello 2010 (2) Lee 2016	Cornish 1997
Study design	RCT	RCT	RCT

	Naltrexone versus alternative opioid antagonist	Naltrexone plus psychological intervention versus other active intervention	Naltrexone, probation and counselling versus probation and counselling alone
Country	Norway	(1,2) USA	USA
Diagnosis	Heroin dependence	(1,2) Opioid dependence	Drug misuse
Age (mean) years	35.1	(1) 33.5 (2) 44.0	39.0
Sex (% female)	6	(1) 18 (2) 15	10
Ethnicity (% white)	Not reported	(1) 47 (2) 23	24
Setting	Initiated in prison, continued in the community	(1,2) Community	Community
Coexisting conditions/other treatments received during study	Not reported	(1) Psychosocial treatment (2) Motivational enhancement counselling	Probation plus brief drug counselling
Treatment length (weeks)	Not reported	(1) 26 (2) 8	26
Intervention (mean dose; mg/day)	Naltrexone	(1) Naltrexone 7975mg/day+70 hours psychosocial contact; (2) 380mg/d	Not reported
Delivery method	Implant (releases drug for 5-6 months)	(1) Oral (2) Intramuscular	Oral
Comparison	Methadone: 30mg/d increasing to 80-130mg/d	(1) Psychosocial treatment (3 hours group therapy, 1 hour individual therapy + 1 hour case management for 6 weeks, then 1 hour individual and 1 hour case management per week for 20 weeks (2) Motivational enhancement	Probation and counselling (3 sessions per week in first 2 weeks)
Note. N= total number of participants; [1] Number randomised			

Table 82: Summary of findings table for the analysis of naltrexone versus active intervention for drug misuse

Outcomes	No of Participa	Quality of the	Relative effect	Anticipated absolute effects

	nts (studies) Follow up	evidence (GRADE)	(95% CI)	Risk with Active intervention	Risk difference with Naltrexone versus active intervention (95% CI)
Retained in treatment	51 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 1.7 (0.76 to 3.82)	294 per 1000	206 more per 1000 (from 71 fewer to 829 more)
Urine test positive for drugs (during treatment) - Alcohol	51 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.5 (0.03 to 7.51)	59 per 1000	29 fewer per 1000 (from 57 fewer to 383 more)
Urine test positive for drugs (during treatment) - Amphetamine	51 (1 study)	⨁⨁⨁⊖ MODERATE[1]	Not estimable[7]	0 per 1000	Not applicable
Urine test positive for drugs (during treatment) - Benzodiazepine	51 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.5 (0.03 to 7.51)	59 per 1000	29 fewer per 1000 (from 57 fewer to 383 more)
Urine test positive for drugs (during treatment) - Cocaine	51 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.69 (0.34 to 1.38)	471 per 1000	146 fewer per 1000 (from 311 fewer to 179 more)
Urine test positive for drugs (during treatment) - Marijuana	51 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.67 (0.17 to 2.65)	176 per 1000	58 fewer per 1000 (from 146 fewer to 291 more)
Urine test positive for drugs (during treatment) - Opiates	51 (1 study)	⨁⨁⊖⊖ LOW[1,2]	RR 0.3 (0.08 to 1.11)	294 per 1000	206 fewer per 1000 (from 271 fewer to 32 more)
Cocaine use (post-treatment)	63 (1 study)	⨁⊖⊖⊖ VERY LOW[2,3]	RR 2.58 (0.54 to 12.33)	62 per 1000	99 more per 1000 (from 29 fewer to 708 more)
Opioid use (post-treatment)	371 (2 studies)	⨁⊖⊖⊖ VERY LOW[2,3,4]	RR 0.67 (0.55 to 0.83)	572 per 1000	189 fewer per 1000 (from 97 fewer to 257 fewer)
Injection drug use (post-treatment)	308 (1 study)	⨁⊖⊖⊖ VERY LOW[2,4]	RR 0.71 (0.28 to 1.81)	65 per 1000	19 fewer per 1000 (from 46 fewer to 52 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Active intervention	Risk difference with Naltrexone versus active intervention (95% CI)
Days of drug use per month (6 month follow-up) - Amphetamine	44 (1 study)	⨁⨁⨀⨀ LOW5,6	-	Mean 8 days (SD 10.45)	MD 2.50 higher (3.86 lower to 8.86 higher)
Days of drug use per month (6 month follow-up) - Benzodiazepine	44 (1 study)	⨁⨁⨀⨀ LOW[5,6]	-	Mean 9.9 days (SD 10.97)	MD 2.0 higher (4.49 lower to 8.49 higher)
Days of drug use per month (6 month follow-up) - Heroin	44 (1 study)	⨁⨁⨀⨀ LOW[5,6]		Mean 20.2 days (SD 12.56)	MD 4.60 lower (12.74 lower to 3.54 higher)
Reincarceration - During treatment	51 (1 study)	⨁⨁⨀⨀ LOW[1,2]	RR 0.5 (0.24 to 1.02)	529 per 1000	265 fewer per 1000 (from 402 fewer to 11 more)
Reincarceration - Post-treatment	308 (1 study)	⨁⨁⨀⨀ LOW[2,4]	RR 0.79 (0.54 to 1.15)	290 per 1000	61 fewer per 1000 (from 134 fewer to 44 more)
Reincarceration - 6 month follow-up	44 (1 study)	⨁⨀⨀⨀ VERY LOW[2,5]	RR 0.91 (0.31 to 2.71)	238 per 1000	21 fewer per 1000 (from 164 fewer to 407 more)
Parole violations (post-treatment)	63 (1 study)	⨁⨀⨀⨀ VERY LOW[2,3]	RR 0.23 (0.05 to 0.98)	281 per 1000	217 fewer per 1000 (from 6 fewer to 267 fewer)
Drug charges (post-treatment)	63 (1 study)	⨁⨀⨀⨀ VERY LOW[2,3]	RR 3.1 (0.34 to 28.19)	31 per 1000	66 more per 1000 (from 21 fewer to 850 more)
Days of criminal activity per month (6 month follow-up)	44 (1 study)	⨁⨀⨀⨀ VERY LOW[5,6]		Mean 14.4 days (SD 13.11)	MD 0.50 higher (7.04 lower to

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Active intervention	Risk difference with Naltrexone versus active intervention (95% CI)
					8.04 higher)
Adverse events (1-year follow-up) - No. of participants experiencing an adverse event	308 (1 study)	⨁⨁⊖⊖ LOW[2,4]	RR 1.34 (1.14 to 1.57)	581 per 1000	197 more per 1000 (from 81 more to 331 more)
Adverse events (1-year follow-up) - Deaths	308 (1 study)	⨁⊖⊖⊖ VERY LOW[2,4]	RR 0.41 (0.08 to 2.06)	32 per 1000	19 fewer per 1000 (from 30 fewer to 34 more)
Adverse events (1-year follow-up) - Non-fatal overdoses	308 (1 study)	⨁⊖⊖⊖ VERY LOW[2,4]	RR 0.11 (0.01 to 2.07)	26 per 1000	23 fewer per 1000 (from 26 fewer to 28 more)

1 Cornish 1997 - unclear randomisation and allocation concealment; unclear blinding; ITT analysis
2 Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome (imprecision) respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
3 Caviello 2010 - Unclear randomisation and allocation concealment; unclear blinding; available case analysis
4 Lee 2016 - Appropriate randomisation and unclear allocation concealment; No blinding to participants; ITT analysis
5 Lobmaier 2010 - appropriate randomisation and allocation concealment; no blinding; ITT analysis
6 Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome (imprecision) respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.
7 No event in either arm of the trial.

Opioid maintenance treatment

Opioid maintenance treatment aims to minimise the harms associated with opioid use, such as blood-borne illnesses associated with needle sharing.

8 RCTs (N=1,565) met the eligibility criteria for this review: Cropsey 2011(Cropsey et al., 2011), Dolan 2003/2005(Dolan et al., 2003; Dolan et al., 2005), Howells 2002(Howells et al., 2002), Magura 2009(Magura et al., 2009), Rich 2015 (Rich et al., 2015), Sheard 2009(Sheard et al., 2009), Shearer 2006(Shearer et al., 2006) and Wright 2011(Wright et al., 2011).

Methadone maintenance versus waiting list control

4 papers from 3 separate RCTs (N=1,047) met the eligibility criteria for this review: Dolan 2003 and Dolan 2005 (Dolan et al., 2003; Dolan et al., 2005), Rich 2015 (Rich et al., 2015) and Shearer 2006 (Shearer et al., 2006).

An overview of the trials can be found in Table 83 . Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 84. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

These were 2-armed trials with service users randomly allocated to either a Methadone treatment arm or waiting list control or forced withdrawal arm.

The data for this comparison were of low to very low quality. No data were available for the outcomes of adaptive functioning and quality of life.

Table 83: Study information table for trials included in the analysis of methadone maintenance versus waiting list control for drug misuse

	Methadone versus waiting list control
Total no. of studies (N[1])	3 (1,047)
Study ID	(1) Dolan 2003/2005 (2) Shearer 2006 (3) Rich 2015
Study design	RCT
Country	(1, 2) Australia (3) USA
Diagnosis	(1, 2,3) Heroin misuse
Age (mean)	(1, 2) 27.0 (3) 34.0
Sex (% female)	(1, 2) 0.0 (3) 22.0
Ethnicity (% white)	(1,2) NR (3) 81.0
Setting	Prison
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	(1) 21 (2,3) Not reported
Intervention (mean dose; mg/day)	Methadone: (1,3) Not reported (2) 61mg/day
Delivery method	(1,2,3) Not reported
Comparison	(1, 2) Wait list control (3) TAU (forced withdrawal)

Note. N= total number of participants; TAU=treatment as usual
[1] Number randomised

Table 84: Summary of findings table for the analysis of methadone versus waiting list control for drug misuse

Outcomes	No of Participants	Quality of the evidence	Relative effect	Anticipated absolute effects Risk with	 Risk difference with

	(studies) Follow up	(GRADE)	**(95% CI)**	**Control**	**Methadone versus control** (95% CI)
Drop-out	382 (1 study)	⨁⨁⊖⊖ LOW[1,2]	RR 1.24 (1.09 to 1.4)	644 per 1000	155 more per 1000 (from 58 more to 258 more)
Positive for opioids - Post-treatment	547 (2 studies)	⨁⊖⊖⊖ VERY LOW[1,2,3,4]	RR 0.86 (0.61 to 1.23)	333 per 1000	47 fewer per 1000 (from 130 fewer to 77 more)
Positive for opioids - 1 month follow-up	197 (1 study)	⨁⊖⊖⊖ VERY LOW[2,3]	RR 0.44 (0.21 to 0.96)	184 per 1000	103 fewer per 1000 (from 7 fewer to 145 fewer)
Positive for opioids - 2 month follow-up	207 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2,3]	RR 0.79 (0.36 to 1.76)	119 per 1000	25 fewer per 1000 (from 76 fewer to 90 more)
Positive for opioids - 3 month follow-up	444 (2 studies)	⨁⊖⊖⊖ VERY LOW[1,2,3]	RR 0.7 (0.5 to 0.99)	242 per 1000	73 fewer per 1000 (from 2 fewer to 121 fewer)
Positive for opioids - 4 month follow-up	538 (2 studies)	⨁⊖⊖⊖ VERY LOW[1,2,3]	RR 0.91 (0.62 to 1.35)	151 per 1000	14 fewer per 1000 (from 57 fewer to 53 more)
Re-incarceration - 1-month follow-up	196 (1 study)	⨁⊖⊖⊖ VERY LOW[2,5]	RR 1.2 (0.51 to 2.8)	92 per 1000	18 more per 1000 (from 45 fewer to 166 more)
Reincarceration - 4-year follow-up	382 (1 study)	⨁⨁⨁⊖ MODERATE[1]	RR 1.04 (0.92 to 1.18)	717 per 1000	29 more per 1000 (from 57 fewer to 129 more)
Adverse events (1 month follow-up) - Deaths	223 (1 study)	⨁⊖⊖⊖ VERY LOW[2,5]	RR 2.87 (0.12 to 69.69)	0 per 1000	-
Adverse events (1 month follow-up) - Non-fatal overdoses	196 (1 study)	⨁⊖⊖⊖ VERY LOW[2,5]	RR 0.39 (0.04 to 4.24)	23 per 1000	14 fewer per 1000 (from 22 fewer to 75 more)

[1] Dolan 2003/2005 - appropriate randomisation and allocation concealment; unclear blinding and available case analysis
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference(MID) for the outcome (imprecision) respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[3] Shearer 2006 - unclear randomisation and allocation concealment; unclear blinding; available case analysis
[4] Evidence was downgraded by one level due to serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of 50%-74.99%) and by two levels due to very serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of >75%).
[5] Rich 2015 - appropriate randomisation and allocation concealment; unclear blinding; ITT analysis

Alpha-adrenergic agonists versus opioid maintenance for substance misuse

1 RCT (N=68) met the eligibility criteria for this review: Howells 2002(Howells et al., 2002).

An overview of the trial can be found in Table 85. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in

Table 86. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

This was a 2-armed trial comparing Lofexidine, an alpha-adrenergic agonist, with methadone. Lofexidine is typically managed in these settings to minimise symptoms of opiate withdrawal. This trial was conducted within a prison setting.

The quality of evidence for this review was low. No evidence was available for the outcomes of offending and reoffending, service utilisation, adaptive functioning or rates of self-injury.

Table 85: Study information table for trials included in the analysis of alpha-adrenergic agonists versus opioid maintenance for substance misuse

	Lofexidine versus methadone
Total no. of studies (N[1])	1 (68)
Study ID	Howells 2002
Study design	RCT
Country	UK
Diagnosis	Opioid dependence
Age (mean) years	30.0
Sex (% female)	0
Ethnicity (% white)	Not reported
Setting	Prison
Coexisting conditions/other treatments received during study	Not reported
Treatment length (days)	10 days
Intervention (mean dose; mg/day)	Oral lofexidine 13mg 2 times per day
Comparison	Oral methadone 175mg 2 times per day
Note. N= total number of participants; [1] Number randomised	

Table 86: Summary of findings table for the comparison of alpha-adrenergic agonists versus opioid maintenance for substance misuse

Outcomes	**No of Participants (studies) Follow up**	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with Opioid maintenance	**Risk difference with alpha-adrenergic (95% CI)**
Total withdrawal symptoms	63 (1 study) 10 days	⨁⨁⨀⨀ LOW[1]	-	The mean total withdrawal symptoms 572.1	MD 24 higher (73.86 lower to 121.86 higher)
1 optimal information size criterion not met; confidence interval of effect includes both appreciable benefit and harm					

Opioid substitution therapy versus active intervention or placebo

4 RCTs (N=450) met the eligibility criteria for this review: Cropsey 2011(Cropsey et al., 2011), Magura 2009(Magura et al., 2009), Sheard 2009(Sheard et al., 2009) and Wright 2011(Wright et al., 2011).

An overview of the trials can be found in Table 87. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in

Table 88. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Each study had 2 arms with buprenorphine in one arm and an alternative opioid substitute or placebo in the other. Three studies were conducted within a prison setting whilst one was conducted in the community.

The data were low to very low quality. No data were available for the outcomes of quality of life or adaptive functioning.

Table 87: Study information table for trials included in the analysis of opioid substitution versus active intervention for substance misuse

	Buprenorphine versus active intervention or placebo
Total no. of studies (N[1])	4 (450)
Study ID	(1) Cropsey 2011 (2) Magura 2009 (3) Sheard 2009 (4) Wright 2011
Study design	RCT
Country	(1, 2) US (3, 4) UK
Diagnosis	Drug (opiate) misuse
Age (mean)	(1) 31.8 (2) 39.5 (3) 29.3 (4) Not reported
Sex (% female)	(1)100 (2, 3) 0 (4) Not reported
Ethnicity (% white)	(1) 88.9 (2, 3) Not reported (4) 92.0
Setting	(1) Community (2, 3, 4) Prison
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	(1) 12 (2) Not reported (3, 4) 3
Intervention (mean dose; mg/day)	Buprenorphine: (1) 2-8 (mean at release=5.8, SD=2.4), (2) 4-38, (3) 96mg, (4) variable
Comparison	(1) Placebo (2, 4) Methadone (3) Dihydrocodeine

	Buprenorphine versus active intervention or placebo
Note. N= total number of participants; [1] Number randomised	

1

2 **Table 88: Summary of findings table for the analysis of opioid substitution versus**
3 **active intervention or placebo for substance misuse**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Opioid substitution therapy versus active intervention or placebo (95% CI)
Drop-out	206 (2 studies)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 0.75 (0.46 to 1.22)	279 per 1000	70 fewer per 1000 (from 151 fewer to 61 more)
Abstinence - Post-treatment	213 (1 study)	⊕⊕⊖⊖ LOW[4,5]	RR 1.06 (0.9 to 1.25)	699 per 1000	42 more per 1000 (from 70 fewer to 175 more)
Abstinence - 1 month follow-up	159 (1 study)	⊕⊖⊖⊖ VERY LOW[4,6]	RR 0.85 (0.68 to 1.06)	736 per 1000	110 fewer per 1000 (from 235 fewer to 44 more)
Abstinence - 3 month follow-up	94 (1 study)	⊕⊖⊖⊖ VERY LOW[4,7]	RR 1.2 (0.87 to 1.65)	562 per 1000	113 more per 1000 (from 73 fewer to 366 more)
Abstinence - 6 month follow-up	150 (2 studies)	⊕⊖⊖⊖ VERY LOW[4,8,9]	RR 1.08 (0.74 to 1.59)	280 per 1000	22 more per 1000 (from 73 fewer to 165 more)
Opioid abuse (3 month follow-up)	116 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,10]	RR 0.81 (0.6 to 1.09)	661 per 1000	126 fewer per 1000 (from 264 fewer to 59 more)
Self-reported injection drug use - Post-treatment	36 (1 study)	⊕⊕⊖⊖ LOW[11]	RR 0.57 (0.27 to 1.2)	583 per 1000	251 fewer per 1000 (from 426 fewer to 117 more)
Self-reported injection drug use - 3 month follow-up	36 (1 study)	⊕⊕⊖⊖ LOW[12]	RR 0.58 (0.25 to 1.35)	500 per 1000	210 fewer per 1000 (from 375 fewer to 175 more)
Number of times rearrested (3 month follow-up)	116 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,10]		Mean 0.71 re-arrests (SD 0.77)	The mean number of times rearrested (3 month follow-up) in the intervention groups was 0.02 standard deviations lower (0.39 lower to 0.34 higher)
Re-arrest for drug crimes (3 month follow-up)	116 (1 study)	⊕⊖⊖⊖ VERY LOW[12,13]	RR 0.57 (0.26 to 1.28)	232 per 1000	100 fewer per 1000 (from 172 fewer to 65 more)
Re-incarceration (post-treatment)	116 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,10]	RR 0.8 (0.53 to 1.2)	500 per 1000	100 fewer per 1000 (from 235 fewer to 100 more)

[1] *high risk performance of bias*

[2] *serious indirectness Maguara 2009 due to population)*
[3] *Optimal information size criterion not met (combined n = 206)*
[4] *high risk performance of bias*
[5] *Optimal information size criterion not met (n = 213)*
[6] *Optimal information size criterion not met (n = 159)*
[7] *Optimal information size criterion not met (n = 94)*
[8] *ROB - Sheared = high performance bias + unclear detection bias + 2 unclear bias.*
[9] *Optimal information size criterion not met (Combined n = 150)*
[10] *Optimal information size criterion not met (n = 116)*
[11] *Optimal information size criterion not met (n = 36)*
[12] *Optimal information size criterion not met (events<100) and CI of effect includes appreciable benefit and harm*

6.4.122 Combined psychological and pharmacological interventions

Antidepressants plus psychological therapy versus psychological therapy alone for substance misuse

1 RCT (N=60) met the eligibility criteria for this review: George 2011(George et al., 2011).

An overview of the trial can be found in Table 89. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 90. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

The trial had 2 arms, with service users being randomly allocated to either receive fluoxetine, a selective serotonin reuptake inhibitor (SSRI), in addition to CBT and motivational therapy or just to receive CBT and motivational therapy. The authors report that fluoxetine was chosen for this study as SSRIs are hypothesised to modulate the processing of environmental stimuli to increase orbital frontal cortex function and accordingly reduce impulsive aggression. This trial was conducted in the community.

The available data for this review was of low quality. No data were available for the outcomes of offending and reoffending, adaptive functioning or rates of self-injury.

Table 89: Study information table for trials included in the analysis of antidepressants plus psychological therapy versus psychological therapy alone for substance misuse

	Fluoxetine plus CBT and motivational therapy versus CBT plus motivational therapy only
Total no. of studies (N[1])	1 (60)
Study ID	George 2011
Study design	RCT
Country	USA
Diagnosis	Alcohol dependence
Age (mean) years	38.9
Sex (% female)	23
Ethnicity (% white)	Not reported
Setting	Community
Coexisting conditions/other treatments received during study	CBT + motivational therapy
Treatment length (weeks)	12 weeks
Intervention	Fluoxetine; 40mg/day plus CBT

	Fluoxetine plus CBT and motivational therapy versus CBT plus motivational therapy only
(mean dose; mg/day)	
Comparison	Placebo plus CBT
Note. N= total number of participants; [1] Number randomised	

Table 90: Summary of findings table for antidepressants plus psychological therapy versus psychological therapy alone for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Psychological therapy only	**Risk difference with Antidepressants + psychological therapy** (95% CI)
No. participants who failed to complete treatment	60 (1 study) 12 weeks	⨁⨁◯◯ LOW[1,2]	RR 1.35 (0.68 to 2.67)	310 per 1000	109 more per 1000 (from 99 fewer to 518 more)
Spielberger state anxiety inventory score (Scale from 20 to 80; lower better)	60 (1 study) 12 weeks	⨁⨁◯◯ LOW[1,2]	-	The mean Spielberger state anxiety inventory score in the control group was 38.2 (SD 13.9)	MD 0.30 lower (6.44 lower to 5.84 higher)
Hamilton rating scale for depression (HRSD) score (Scale from 0 to 52; lower better)	60 (1 study) 12 weeks	⨁⨁◯◯ LOW[1,2]	-	The mean Hamilton rating scale for depression (HRSD) score in the control groups was 11.5 (SD 7.2)	MD 3.10 lower (6.18 to 0.02 lower)
1 unclear selection, detection and attrition bias *2 optimal information size criterion not met*					

6.4.1.251 Support and educational interventions

Psychoeducation versus control or treatment as usual (TAU)

1 RCT (N=60) met the eligibility criteria for this review: Brown 1980(Brown, 1980).

An overview of the trial can be found in Table 91. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 92. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

The RCT had 3 arms with service users being allocated to either psychoeducation, educational drinking (where participants learned to control their drinking behaviour in an experimental bar facility) or treatment as usual. Only the psychoeducation and treatment as usual arms are included here. The psychoeducational intervention consisted of 3-hour sessions comprising a 30 minute talk, a 30 minute film, and then a chaired group discussion. The types of topic covered included drinking and driving, effects of alcohol on physical

health, effects upon family and how to modify drinking habits. Treatment as usual consisted of assigned tasks at the periodic detention centre.

The available data for this review was of very low quality. No data were available for the outcomes of offending or reoffending, adaptive functioning, service utilisation or rates of self-injury.

Table 91: Study information table for trials included in the analysis of psychoeducation versus control or treatment as usual for drug misuse

	Psychoeducation versus control or treatment as usual
Total no. of studies (N[1])	1 (60)
Study ID	Brown 1980
Study design	RCT
Country	New Zealand
Diagnosis	Alcohol misuse
Age (mean)	32.0 years
Sex (% female)	0
Ethnicity (% white)	Not reported
Setting	Community
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	5 weeks
Intervention (mean dose; mg/day)	Psychoeducation; 3 hours per week
Comparison	TAU (The control group did not attend any educational sessions but continued to carry out assigned tasks at the Periodic Detention Centre)
Note. N= total number of participants; [1] Number randomised	

Table 92: Summary of findings table for psychoeducation versus control or treatment as usual for drug misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Psychoeducation versus control/TAU (95% CI)
Number of days with uncontrolled drinking	34 (1 study)	⊕⊖⊖⊖ VERY LOW[1, 2]	-	Mean 26.2 days (SD 1.6)	MD 4.85 days fewer (11.46 fewer to 1.76 more)

1 high risk for performance, detection and selective reporting
2 Optimal information size criterion not met (N<400); 95% CI of effect includes both appreciable benefit and harm

Employment workshop versus control or treatment as usual for substance misuse

2 RCTs (N=555) met the eligibility criteria for this review: Hall 1981(Hall et al., 1981) and Webster 2014(Webster et al., 2014).

An overview of the trials can be found in Table 93. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 94. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Both studies were 2-armed trials conducted in the community. The experimental arm of both trials (employment workshop) consisted of a mixture of individual and group sessions designed to provide information, support and opportunities to practice skills needed to find and maintain employment and seek promotion. The control arm of the Hall 1981 study consisted of a 3-hour sign-posting meeting. The control arm of the Webster 2014 study consisted of treatment as usual.

The evidence for this review was of low to very low quality. No data were available for the outcomes of mental health, offending and reoffending, service utilisation or rates of self-injury.

Table 93: Study information table for trials included in the analysis of employment workshop versus control or treatment as usual for substance misuse

	Employment workshop versus control/TAU
Total no. of studies (N[1])	2 (555)
Study ID	(1) Hall 1981 (2) Webster 2014
Study design	RCT
Country	USA
Diagnosis	Substance misuse
Age (mean) years	(1) 33.9 (2) 30.5
Sex (% female)	(1) 15 (2) 35
Ethnicity (% white)	(1) 34 (2) 62
Setting	Community
Coexisting conditions/other treatments received during study	Not reported
Treatment length (weeks)	(1) 3 days (2) 26 sessions
Intervention (mean dose; mg/day)	Employment workshop: (1) 8 hours total (2) NR
Delivery method	(1, 2) Individual and group
Comparison	(1) 3-hour meeting (2) TAU (not specified)
Note. N= total number of participants; [1] Number randomised	

Table 94: Summary of findings table for employment workshop versus control or treatment as usual for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control/TAU	Risk difference with Employment workshops (95% CI)
No. of participants employed	529 (2 studies) 12-52 weeks	⊕⊖⊖⊖ VERY LOW[1,2,3,4]	RR 1.24 (0.84 to 1.81)	735 per 1000	176 more per 1000 (from 118 fewer to 596 more)
Days in paid employment	477 (1 study)	⊕⊕⊖⊖ LOW[5]	-	The mean days in paid employment in the control	MD 10.20 higher (11.8 lower to 32.2

52 weeks		group was 199.9 days	higher)

[1] high risk of performance, detection and reporting bias, unclear bias on 3 other dimensions
[2] I^2=73%; random effects model used; no reasons for this heterogeneity were identified; study effect estimates were RR=1.58 [1.06, 2.36] for Hall (1961) and RR = 1.06 [0.97, 1.17] for Webster (2014)
[3] Hall 1981-unclear whether the population have a current drug or other mental health problem
[4] Hall 1981, small sample size
[5] high risk of detection and performance bias, unclear risk on 3 other domains

6.4.1.22 Physical interventions

Acupuncture versus active intervention

2 RCTs (N=726) met the eligibility criteria for this review: Berman 2004(Berman et al., 2004) and Konefal 1995(Konefal et al., 1995).

An overview of the trials can be found in Table 95. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 96. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Both studies were 2-armed trials. The Berman 2004 study compared two different forms of acupuncture, the NADA (National Acupuncture Detoxification Association) and the Helix protocols. This study was conducted within a prison setting. In the Konefal 1995 study service users in one arm received acupuncture in addition to frequent urine testing and in the other frequent urine testing only. This study was conducted in the community. Both studies used the NADA protocol for acupuncture in the intervention arm. The NADA protocol consists of 5 points chosen for their ability to assist with detoxification; Shen-Men, sympathetic, kidney, liver, and lung. The Helix protocol involved acupuncture to the ear using five points on the helix of the ear

The evidence for this review was low to very low quality. No data were available for the outcomes of offending or reoffending, adaptive functioning or rates of self-injury.

Table 95: Study information table for trials included in the analysis of acupuncture versus active intervention for substance misuse

	Acupuncture versus active intervention
Total no. of studies (N[1])	2 (726)
Study ID	(1) Berman 2004 (2) Konefal 1995
Study design	RCT
Country	(1) Sweden (2) USA
Diagnosis	(1, 2) Substance misuse
Age (mean) years	(1) 33.5 (2) Not reported
Sex (% female)	(1) 39 (2) 47
Ethnicity (% white)	(Not reported
Setting	(1) Prison (2) Community
Coexisting conditions/other treatments received during study	Not reported

	Acupuncture versus active intervention
Treatment length (weeks)	(1) 4 (2) 16
Intervention (mean dose; mg/day)	Acupuncture; (1) 5 times per week in week 1, then 3 times per week (2) 5 times per week for 2 weeks, 3 times per week until week 12, 2 times per week in weeks 13-16 plus frequent urine testing
Comparison	Acupuncture; (1) Helix protocol (2) Frequent urine testing
Note. N= total number of participants; [1] Number randomised	

Table 96: Summary of findings table for acupuncture versus active intervention for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Acupuncture versus active intervention (95% CI)
Drop-out	158 (1 study)	⨁⨁⊖⊖ LOW[1,2]	RR 1.45 (1.06 to 1.99)	421 per 1000	189 more per 1000 (from 25 more to 417 more)
Urine test positive for drug use post-treatment	108 (2 studies)	⨁⊖⊖⊖ VERY LOW[3,4,5,6]	RR 3.65 (0.33 to 41)	129 per 1000	342 more per 1000 (from 86 fewer to 1000 more)

1 allocation concealment, attrition and selective reporting all high risk of bias
2 Optimal information size criterion not met (N<300 events)
3 Both studies had allocation concealment, attrition and selective reporting all high risk of bias
4 I2 66% - random effects model used; large variation in effect sizes: Berman 16.39, Konefal 1.59, but no explanation for the heterogeneity was identified
5 For one study (Konefal 1995) - only 51% of participants were in contact with CJS
6 Optimal information size criterion not met (N <300 events) and CI of effect includes both appreciable benefit and harm

6.4.1.3 Depression

Three RCTs (N = 206) met the eligibility criteria for this review: Gussak 2009, Johnson 2012 and Wilson 1990 (Gussak, 2008; Johnson & Zlotnick, 2012; Wilson, 1990). Gussak 2009 was arts-based psychotherapy (examples included construction of three-dimensional forms with few supplies); Johnson 2012 used interpersonal psychotherapy (IPT) intervention compared to psychoeducation whereas Wilson 1990 compared group cognitive treatment with individual supportive therapy. Due to the differences in the psychotherapy interventions data were not combined and separate analysis was done and presented for each study. In Johnson 2012, IPT was based on Wilfrey 2000 psychotherapy model while in Wilson 1990, group therapy was based on Hollon and Shaw 1979 cognitive treatment model.

An overview of the trials included in the meta-analysis can be found in Table 97. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 98. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 97 Study information table for trials included in the meta-analysis of psychotherapy for depression

	Interpersonal psychotherapy vs Psychoeducation (PSYCHOED)	Group cognitive treatment vs Individual supportive therapy	Arts-based therapy vs no treatment
Total no. of studies (N[1])	1(38)	1(10)	1(158)
Study ID	Johnson 2012	Wilson 1990	Gussak 2009
Study design	RCT	RCT	RCT
Country	USA	USA	USA
Underlying Mental Health Disorder	Moderate to Severe Depression	Moderate Depression	Axis I diagnosis (Depression)
Diagnosis	DSV-IV criteria	Clinical	Clinical
Criminal justice population	Sentenced volunteers from state prison	Inmates at a large maximum-security prison	Inmates at medium to maximum adult correctional facilities
Age (mean/range) years	35 (median)	33.1	20-51
Gender (% female)	100	Not reported	60.8
Ethnicity (% white)	Not reported	Not reported	64.3
Intervention	Interpersonal psychotherapy	Group cognitive treatment	Arts-based therapy
Comparator	Psychoeducation	Individual supportive therapy	No treatment
Format (number of participants per group)	Individual and group (Not reported)	Group (5/group)	Group (8/group)
Intervention Dose/intensity	3-4hours/session (3 sessions/week)	Not reported	One session/week
Comparator Dose/intensity	1-1.5 hours/session (3 sessions/week)	A total of four 30-min sessions plus weekly check-in visits	Not reported
Intervention setting	Initiated in prison and continued in the community	At subsequent time points in prison	Prison
Treatment length (weeks)	8	52	15
Follow-up length (weeks)	13	39	Not reported
Note. N= total number of participants [1] Number randomised			

Table 98: Summary of findings table of psychological intervention versus active intervention or no treatment for depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with comparator	Risk difference with Psychotherapy versus control/TAU (95% CI)
Depression by HRSD scales at post-treatment (Psychotherapy	38 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 20.6 (SD 10.5)	MD 6.5 lower (12.52 to 0.48 lower)

versus PSYCHOED) (Scale from 0 to 52; lower better)					
Depression by HRSD scales (13 weeks Follow-up) (Psychotherapy versus PSYCHOED) (Scale from 0 to 52; lower better)	38 (1 study)	⨁⊖⊖⊖ VERY LOW[3,2]	-	Control mean 12.0 (SD 12.3)	MD 3.8 higher (3.83 lower to 11.43 higher)
Depression by Beck Depression Inventory (BDI) at post-treatment (Group therapy versus Individual therapy) (Scale from 0 to 63; lower better)	10 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 16.2 (SD 6.76)	MD 3.2 lower (13.56 lower to 7.16 higher)
Depression by Beck Hopelessness Scale (BHS) at post-treatment (Group therapy versus Individual therapy) (Scale from 0 to 20; lower better)	10 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 4.2 (SD 4.14)	MD 2.6 higher (4.98 lower to 10.18 higher)
Depression by MMPI D scale at post-treatment (Group therapy versus Individual therapy)	10 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 57.2 (SD 10.98)	MD 12.6 higher (3.38 lower to 28.58 higher)
Depression by MMPI D scale (39 weeks Follow-up) (Group therapy versus Individual therapy)	10 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 56.4 (SD 14.22)	MD 4.8 higher (9.68 lower to 19.28 higher)
Depression by Multiple affect adjective Check list D scale at post-treatment (Group therapy versus Individual therapy)	10 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 8.2 (SD 3.49)	MD 0.6 higher (4.93 lower to 6.13 higher)
Change in Adult Nowicki-Strickland Locus of Control Scale (ANS) – Total at post-treatment (Arts-based therapy versus TAU) (Scale from 0 to 40; lower better)	122 (1 study)	⨁⊖⊖⊖ VERY LOW[2,4]	-	Control mean 0.56	MD 2.93 lower (4.41 to 1.46 lower)
Change in Adult Nowicki-Strickland Locus of Control Scale (ANS) – Male	62 (1 study)	⨁⊖⊖⊖ VERY LOW[2,4]	-	Control mean 1.04 (SD 3.61)	MD 2.26 lower (4.18 to 0.34 lower)

at post-treatment (Arts-based therapy versus TAU) (Scale from 0 to 40; lower better)					
Change in Adult Nowicki-Strickland Locus of Control Scale (ANS) – Female at post-treatment (Arts-based therapy versus TAU) (Scale from 0 to 40; lower better)	60 (1 study)	⊕⊖⊖⊖ VERY LOW[2,4]	-	Control mean 0.12 (SD 9.8)	MD 6.81 lower (11.97 to 1.65 lower)
Change in Beck Depression Inventory (BDI): Total at post-treatment (Arts-based therapy versus TAU) (Scale from 0 to 63; lower better)	156 (1 study)	⊕⊖⊖⊖ VERY LOW[2,4]	-	Control mean -- 1.844 (SD 8.31)	MD 6.5 lower (9.33 to 3.67 lower)
Change in Beck Depression Inventory (BDI): Total - Male at post-treatment (Arts-based therapy versus TAU) (Scale from 0 to 63; lower better)	60 (1 study)	⊕⊖⊖⊖ VERY LOW[2,4]	-	Control mean 0.12 (SD 9.8)	MD 6.81 lower (11.97 to 1.65 lower)
Change in Beck Depression Inventory (BDI): Total – Female at post-treatment (Arts-based therapy versus TAU) (Scale from 0 to 63; lower better)	96 (1 study)	⊕⊖⊖⊖ VERY LOW[2,4]	-	Control mean -- 4.3 (SD 5.22)	MD 6.37 lower (9.76 to 2.98 lower)
Change in Formal Elements of Arts Therapy Scale rating guide (FEATS):Prominence of colour at post-treatment (Arts-based therapy versus TAU) (Scale from 1 to 5; higher better)	84 (1 study)	⊕⊖⊖⊖ VERY LOW[2,4]	-	Control mean 0.42 (SD 1.44)	MD 0.81 lower (1.51 to 0.11 lower)
Change in Formal Elements of Arts Therapy Scale rating guide	84 (1 study)	⊕⊖⊖⊖ VERY LOW[12,4]	-	Control mean 0.24 (SD	MD 0.45 lower (0.84 to 0.06 lower)

(FEATS):Colour fit at post-treatment (Arts-based therapy versus TAU) (Scale from 1 to 5; higher better)		

[1] Wilson 1990 - Unclear selection bias, No blinding, low attrition rate, low selective outcome reporting, low other risk of bias
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.
[3]Johnson 2012 - Unclear risk of bias, unclear blinding of participants and care administrators, blinding of outcome assessors, low attrition bias, unclear selective outcome bias, low other risk of bias
[4]Gussak 2009 - Unclear randomization and allocation, No blinding of patients and care administrators, Blinding of outcome assessors

6.4.12.4 Individuals with suicidal risk

One RCT (N = 46) met the eligibility criteria for this review: Biggam 2002 (Biggam & Power, 2002). Principal training techniques in social problem-solving group therapy included instruction, active discussion, reflective listening and group exercises to practice the targeted skills. It was delivered in small group format (4-6 individuals/group). The participants in control did not receive principal training techniques.

An overview of the trials included in the meta-analysis can be found in Table 99. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 100. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 99: Study information table for trials included in the meta-analysis of social problem-solving group therapy for vulnerable personality with suicidal risks

	Social problem-solving group therapy vs No treatment control
Total no. of studies (N[1])	1 (46)
Study ID	Biggam 2002
Study design	RCT
Country	UK
Underlying Mental Health Disorder	Victims of bullying who had difficulty adjusting to main circulation in prisons
Diagnosis	Symptoms
Criminal justice population	Vulnerable offenders with suicidal risks or those in formal protection units or those being bullied by another inmate
Age (mean) years	19.3
Gender (% female)	Not reported
Ethnicity (% white)	Not reported
Intervention	Social problem-solving group therapy
Comparator	No treatment control

	Social problem-solving group therapy vs No treatment control
Criminal Justice setting	Prison
Format (number of participants per group)	Group (6/group)
Dose/Intensity	Five 90-minutes sessions (7.5 hours in total duration of intervention)
Treatment length (weeks)	NR
Follow-up length (weeks)	13 weeks
Note. N= total number of participants; NR=Not reported [1] Number randomised	

1

2 **Table 100 Summary of findings table of social problem-solving group therapy versus**
3 **no treatment control for vulnerable personality with suicidal risks**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment control	Risk difference with Social problem solving group versus no treatment control (95% CI)
MH outcomes: Depression by HADS scales at post-treatment (Scale from 0 to 21; lower better)	46 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 9.1 (SD 4.2)	MD 3.6 lower (5.76 to 1.44 lower)
MH outcomes: Anxiety by HADS scales at post-treatment (Scale from 0 to 21; lower better)	46 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 9.6 (SD 3.3)	MD 2.9 lower (4.67 to 1.13 lower)
MH outcomes: Depression by Beck Hopelessness scale at post-treatment (Scale from 0 to 20; lower better)	46 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 6.4 (SD 4.7)	MD 2.5 lower (4.89 to 0.11 lower)
MH outcomes: Decision making ability by SPSI:R scales at post-treatment	46 (1 study)	⨁⨁⊖⊖ LOW[1,2]	-	Control mean 6.8 (SD 4.9)	MD 5.3 higher (2.66 to 7.94 higher)
MH outcomes: Depression by HADS scale (13 weeks Follow-up) (Scale from 0 to 21; lower better)	46 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 8.4 (SD 3.6)	MD 3.3 lower (5.19 to 1.41 lower)
MH outcomes: Anxiety by HADS scales (13 weeks Follow-up) (Scale from 0 to 21; lower better)	46 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 9.6 (SD 3.5)	MD 2.7 lower (4.61 to 0.79 lower)
MH outcomes: Depression by Beck Hopelessness	46 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 7.0 (SD 4.9)	MD 2.8 lower (5.13 to 0.47 lower)

scale (13 weeks Follow-up) (Scale from 0 to 20; lower better)					

[1]Biggam 2002 - Unclear risk of selection bias, No blinding, low attrition bias, unclear selective outcome reporting, low other risk of bias
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries

6.4.12.5 Anxiety Disorder

One RCT (N = 38) met the eligibility criteria for this review: Maunder 2009 (Maunder et al., 2009). The therapy was based on CBT-principles. The intervention group was provided with a booklet with a list of instructions and exercises and completed the time diary and thought about their personal reactions to the booklet. The participants in control group did not receive self-help booklets.

An overview of the trials included in the meta-analysis can be found in Table 101. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 102. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 101 Study information table for trials included in the meta-analysis of self-help materials versus wait-list control for anxiety disorders

	Self-help materials vs Wait-list control
Total no. of studies (N[1])	1(38)
Study ID	Maunder 2009
Study design	RCT
Country	UK
Underlying Mental Health Disorder	Anxiety Disorders
Diagnosis	Hospital Anxiety and Depression Scale (HADS) ≥8
Criminal justice population	Prisoners
Age (mean) years	35.22
Gender (% female)	99.9
Ethnicity (% white)	NR
Intervention	Self-help materials
Comparator	Wait-list control
Format (number of participants per group)	Individual
Dose/intensity (hours)	Not reported
Intervention setting	Prison
Treatment length (weeks)	4
Follow-up length (weeks)	4
Notes. N= total number of participants; NR=Not reported	

	Self-help materials vs Wait-list control
[1] Number randomised	

Table 102 Summary of findings table of self-help materials versus wait-list control for anxiety disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with wait-list control	Risk difference with Self-help materials versus wait-list control (95% CI)
MH outcomes: Anxiety by HADS scale at post-treatment (Scale from 0 to 21; lower better)	33 (1 study)	⊕⊕⊖⊖ LOW[1,2]	-	Control mean 13.67 (SD 3.08)	MD 1.06 lower (3.63 lower to 1.51 higher)
MH outcomes: Anxiety by HADS scale (4 weeks follow-up) (Scale from 0 to 21; lower better)	33 (1 study)	⊕⊕⊖⊖ LOW[1,2]	-	Control mean 13.87 (SD 4.19)	MD 2.98 lower (5.82 to 0.14 lower)

[1] Maunder 2009 - low selection risk of bias, No blinding of participants but blinding of care administrators (+), unclear outcome assessor, unclear attrition risk of bias, unclear other risk of bias (blocked randomization with single blinded trial)

[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries

6.4.1.6 PTSD

Four RCTs (N =290) met the eligibility criteria for this review: Bradley 2003, Ford 2013, Cole 2007 and Valentine 2001(Bradley & Follingstad, 2003; Cole et al., 2007; Ford et al., 2013; Valentine & Smith, 2001). Bradley 2003, Ford 2013 and Cole 2007 studies used group therapy method whereas Valentine 2001 applied traumatic incident reduction psychotherapy model. Bradley 2003 and Cole 2007 compared with no-contact and wait-list control respectively. Thus, outcomes were combined in analysis if they were measured by the same measurement tool. Ford 2013 study compared Trauma affect regulation: Guide for Education and Therapy (TARGET) with small group therapy (SGT). Moreover, Valentine 2001 study evaluated trauma incident reduction compared to wait-list controls. Data from these studies were analysed separately.

An overview of the trials included in the meta-analysis can be found in Table 103. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 104. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 103 Study information table for trials included in the meta-analysis of psychotherapy versus no treatment/wait-list control/active treatment for post-traumatic stress disorders

	Group therapy vs No treatment	Group intervention vs Wait-list control	TARGET vs SGT	TIR vs Wait-list control
Total no. of studies (N[1])	1(49)	1(13)	1(80)	1(148)
Study ID	Bradley 2003	Cole 2007	Ford 2013	Valentine 2001
Study design	RCT	RCT	RCT	RCT
Country	USA	USA	USA	USA
Underlying mental health disorders	PTSD and Depression	PTSD (history of childhood sexual abuse)	PTSD	PTSD
Diagnosis	Symptoms	Symptoms	Symptoms	Diagnosis
Criminal justice population	Inmates in medium-security prisons	Recently-incarcerated women	Inmates in a state prison	Inmates
Age (mean) years	36.7	31	36.3	33.9
Gender (% female)	100	100	100	100
Ethnicity (% white)	38	0.33	60	38.5
Intervention	Group therapy	Time-limited Trauma-focused group intervention	Trauma affect regulation: Guide for Education and Therapy (TARGET)	Trauma incident reduction (TIR)
Comparator	No-contact	Wait-list control	Supportive group therapy	Wait-list control
Format (number of participants per group)	Group (24/group)	Group (7/group)	Group (41/group)	Individual
Dose/intensity (hours)	45	40 (5 hours/week)	13	Not reported
Intervention setting	Prison	Prison	Prison	Prison
Treatment length (weeks)	Not reported	8	Not reported	Not reported
Follow-up (weeks)	Not reported	Not reported	Not reported	13

Notes. N= total number of participants; TIR=Traumatic incident reduction; TARGET=Trauma Affect Regulation: Guide for Education and Therapy; PTSD=Post-traumatic stress disorders

[1] Number randomised

Table 104 Summary of findings table of psychotherapy vs wait-list control/ No treatment/ Active treatment for PTSD

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with wait-list control/no treatment/active treatment	Risk difference with Psychological Therapy wait-list control/no treatment/active treatment (95% CI)
Trauma by TSI at post-treatment (Group	40 (2 studies)	⨁⊖⊖⊖ VERY LOW[1,2,3,4]	-	Control mean	MD 11.67 lower (30.36 lower to 7.02

Therapy vs Wait-list/No-contact Control) (Scale from 0 to 300; lower better)				65.29	higher)
Depression by BDI total at post-treatment (TIR vs Wait-list control) (Scale from 0 to 63; lower better)	123 (1 study)	⨁⊖⊖⊖ VERY LOW[4,6]	-	Control mean 16.4	MD 3.8 lower (7.52 to 0.08 lower)
Depression by BDI total (13 weeks Follow-up) (TIR vs Wait-list control) (Scale from 0 to 63; lower better)	123 (1 study)	⨁⊖⊖⊖ VERY LOW[1,4]	-	Control mean 17.5	MD 7.8 lower (12.64 to 2.96 lower)
PTSD by PSS scales at post-treatment (TIR vs Wait-list control) (Scale from 0 to 51; lower better)	123 (1 study)	⨁⊖⊖⊖ VERY LOW[4,6]	-	Control mean 18.2	MD 4.1 lower (7.96 to 0.24 lower)
PTSD by PSS scales (13 weeks follow-up) (TIR vs Wait-list control) (Scale from 0 to 51; lower better)	123 (1 study)	⨁⊖⊖⊖ VERY LOW[4,6]	-	Control mean 15.8	MD 7.3 lower (11.49 to 3.11 lower)
Generalized Expectancy for Success Scale at post-treatment (TIR vs Wait-list control) (Scale from 30 to 150; higher better)	123 (1 study)	⨁⊖⊖⊖ VERY LOW[4,6]	-	Control mean 106.1	MD 15.9 higher (5.7 to 26.1 higher)
Generalized Expectancy for Success Scale (13 weeks follow-up) (TIR vs Wait-list control) (Scale from 30 to 150; higher better)	123 (1 study)	⨁⊖⊖⊖ VERY LOW[4,6]	-	Control mean 108.6	MD 3.6 higher (2.69 lower to 9.89 higher)
Clinical Anxiety scale at post-treatment (TIR vs Wait-list control) (Scale from 0 to 100; lower better)	123 (1 study)	⨁⊖⊖⊖ VERY LOW[4,6]	-	Control mean 56.0	MD 3.3 lower (8.55 lower to 1.95 higher)
Clinical Anxiety scale (13 weeks follow-up) (TIR vs Wait-list control) (Scale from 0 to 100;	123 (1 study)	⨁⊖⊖⊖ VERY LOW[4,6]	-	Control mean 17.5	MD 7.8 lower (12.64 to 2.96 lower)

lower better)					
PTSD symptoms by CAPS scales at post-treatment (TARGET vs SGT)	72 (1 study)	⊕⊕⊕⊖ MODERATE[5]	-	Control mean 24.3	MD 0.5 lower (11.01 lower to 10.01 higher)
Heartland forgiveness scale at post-treatment (TARGET vs SGT) (Scale from 18 to 126; higher better)	32 (1 study)	⊕⊕⊖⊖ LOW[4,5]	-	Control mean 76.7	MD 4.6 higher (6.73 lower to 15.93 higher)
Symptom checklist-90-R: Global Severity Index at post-treatment (Focused group therapy vs Wait-list control) (Scale from 0 to 4; lower better)	9 (1 study)	⊕⊕⊖⊖ LOW[1]	-	Control mean 76.8	MD 16.3 lower (26.23 to 6.37 lower)
Symptom Checklist-90R: Positive Symptom Distress Index at post-treatment (Focused group therapy vs Wait-list control) (Scale from 0 to 4; lower better)	9 (1 study)	⊕⊖⊖⊖ VERY LOW[1,4]	-	Control mean 75.2	MD 13.9 lower (24.8 to 3 lower)
Symptom Checklist-90R: Positive Symptom Total at post-treatment (Focused group therapy vs Wait-list control) (Scale from 0 to 90; lower better)	9 (1 study)	⊕⊕⊖⊖ LOW[1]	-	Control mean 74.4	MD 16.1 lower (26.67 to 5.53 lower)
IIP-32 scales at post-treatment (Group therapy vs No contact control) (Scale from 0 to 128; lower better)	31 (1 study)	⊕⊖⊖⊖ VERY LOW[2,4-]		Control mean 43.4	MD 10.1 lower (24.43 lower to 4.23 higher)

[1] Cole 2007 - high risks of selection bias, No blinding, Unclear attrition bias, low selective outcome bias and low other risk of bias, *[2]Bradley 2003 - unclear risks of selection bias, No blinding, Unclear attrition, High selective outcomes bias and low other risks of bias*

[3]Evidence was downgraded by one level due to serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of 50%-74.99%) and by two levels due to very serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of >75%). Random effects model used; no explanation for the heterogeneity was identified

[4]The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

[5]Ford 2013- low risk of selection bias, blinding of care administrators and outcome assessors but no blinding of participants, low attrition rate, low selective outcome bias, low other risk of bias

[6]Valentine 2001 - high risk of selection bias, No blinding, unclear attrition bias, low selective outcome bias, low other

risk of bias

6.4.2.11 ADHD

Two studies (N = 84) met the eligibility criteria for this review: Ginsberg 2012 and Konstenius 2013(Ginsberg et al., 2012b; Ginsberg & Lindefors, 2012a; Konstenius et al., 2013). The placebo and the methylphenidate capsules and packaging were identical in appearance. Data were combined by meta-analysis as the population, type of intervention and placebo were the same.

An overview of the trials included in the meta-analysis can be found in Table 105 . Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 106. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 105 Study information table for trials included in the meta-analysis of methylphenidate versus Placebo for Attention Deficit Hyperactivity Disorder (ADHD)

	Methylphenidate versus Placebo
Total no. of studies (N[1])	2 (84)
Study ID	(1) Ginsberg 2012 (2) Konstenius 2013
Study design	RCT
Country	Sweden
Underlying mental health disorders	ADHD
Diagnosis	Symptoms
Criminal justice population	(1) long-term inmates convicted of violent or drug-related crimes (2) medium-security voluntary prisoners
Age (mean) years	(1) 34.4 (2) 41.5
Gender (% female)	(1) 0 (2) Not reported
Ethnicity (% white)	Not reported
Intervention	Methylphenidate
Comparator	Placebo
Format	Per oral
Dose/intensity (mg/day)	(1)18mg/day for first 19 days; 36mg/day increment every 3 days up to maximum of 180mg/day (2)36mg/day for 4 days to 54mg/day for 3 days and then to 72mg/day (titrated individually but not exceeding 1.3mg/kg daily)
Intervention setting	Prison
Treatment length (weeks)	(1) 24weeks (2) 52weeks
Follow-up (weeks)	(1) 156 weeks (2)Not reported
Notes. N= total number of participants; [1] Number randomised	

Table 106 Summary of findings table for Methylphenidate versus Placebo for Attention Deficit Hyperactivity Disorder

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Methylphenidate (MPH) versus placebo (95% CI)
Conners Adult ADHD rating scale - Observer: Screening Version (CAARS-OSV) at post-treatment (Scale from 0 to 90; lower better)	84 (2 studies)	⊕⊖⊖⊖ VERY LOW[1,2,3,4]	-	Control mean 2.98	MD 12.85 lower (22.5 to 3.20 lower)
Conners Adult ADHD rating scale - Observer: Screening Version (CAARS-OSV) – 3-years follow-up (Scale from 0 to 90; lower better)	20 (1 study)	⊕⊕⊖⊖ LOW[1,2,4]	-	Control mean 29.6 (SD 7.7)	MD 16.9 lower (24.5 to 9.3 lower)
Number of participants with drug negative urine at post-treatment	54 (1 study)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 1.5 (0.48 to 4.72)	148 per 1000	74 more per 1000 (from 77 fewer to 551 more)

[1]Ginsberg 2012 - high risk of selection bias, No blinding, low risk of attrition, unclear selective outcome reporting and low risk of other bias
[2]Konstenius 2013- low risk of selection bias, Blinding of participants, care administrators and outcome detectors, unclear attrition bias and unclear selective outcome reporting, low risk of other bias
[3] Evidence was downgraded by one level due to serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of 50%-74.99%) and by two levels due to very serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of >75%). Random effects model used; no explanation for the heterogeneity was identified
[4] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries. If SMD was used, +0.5 and -0.5 on the SMD scale as MID boundaries.'
[5] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference for the outcome (imprecision) respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

6.4.2.52 Antisocial personality disorders

One RCT (N = 12) met the eligibility criteria for this review: Gowin 2012 (Gowin et al., 2012). This study took a total of 6 week, with no treatment in week 1, placebo to both groups in week 2, either tiagabine or placebo in week 3, 4 and 5 and placebo to both groups in week 6. Tiagabine was given orally as an escalating manner; 4,8,12 mg bd over 3 weeks. Corn starch was used to fill the capsules.

An overview of the trials included in the meta-analysis can be found in Table 107. Further information about both included and excluded studies can be found in Appendix L. Summary of findings can be found in Table 108. The full GRADE evidence profiles and associated forest plots can be found Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 107 Study information table for trials included in the meta-analysis of Tiagabine versus placebo for the antisocial personality disorder

	Tiagabine versus Placebo
Total no. of studies (N[1])	1(12)
Study ID	Gowin 2012
Study design	RCT
Country	USA
Underlying mental health disorder	Antisocial personality disorder
Diagnosis	Symptoms
Criminal justice population	Majority were on probation but not limited to
Age (mean) years	28.7
Gender (% female)	0.17
Ethnicity (% white)	Not reported
Intervention	Tiagabine
Comparator	Placebo
Format	Per oral
Dose/intensity (mg/day)	In ascending dose from 4 to 12 mg bd
Intervention setting	Community
Treatment length (weeks)	3
Follow-up (weeks)	Not reported
Notes. N= total number of participants; [1] Number randomised	

Table 108 Summary of findings table for Tiagabine versus placebo for the antisocial personality disorder

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Tiagabine versus placebo (95% CI)
Change in aggressive response at post-treatment	12 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]		Control mean 0.47 (SD 0.45)	MD 1.86 lower (2.70 to 1.02 lower)
Number of reports on adverse effects at post-treatment	222* (1 study)	⊕⊖⊖⊖ VERY LOW[1,3]	RR 0.41 (0.14 to 1.24)	92 per 1000	54 fewer per 1000 (from 79 fewer to 22 more)

[1] Gowin 2012- Unclear risk of selection bias, blinding to participants and care person involved,, low risk of attrition, unclear selective outcome reporting, low risk of other bias.
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.
[3] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome (imprecision) respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

** total number of 'Yes' reports to the side-effects at least once*

6.4.213 Severe mental illness

6.4.2.321 Pharmacological interventions

One RCT (N=450) met the eligibility criteria for the review: Alphs 2015a (Alphs et al., 2015). In this study, Paliperidone was given intramuscularly with 234 mg on day 1 and 158 mg on day 8 with monthly maintenance range of 78-238 mg thereafter from day 38. Patients on oral regime received aripiprazole 33 (15.1%), haloperidol 15(6.9%), olanzapine 36 (16.5%), paliperidone 48(22%), perphenazine 20(9.2%), quetiapine 29 (13.3%) and risperidone 37 (17%).

An overview of the trials included in the meta-analysis can be found in Table 109. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 110. The full GRADE evidence profiles and associated forest plots can be found Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 109 Study information table for trials included in the meta-analysis of IM paliperidone versus oral antipsychotics for schizophrenia

	IM Paliperidone vs Oral Antipsychotics
Total no. of studies (N[1])	1(450)
Study ID	Alphs 2015a
Study design	RCT
Country	USA
Underlying mental health disorder	Schizophrenia
Diagnosis	Diagnosis
Criminal justice population	Offenders on release
Age (mean) years	38.2
Gender (% female)	13.7
Ethnicity (% white)	33.2
Intervention	Paliperidone Palmitate
Comparator	Daily Oral Antipsychotics
Format	Intramuscular (Intervention) vs Oral Antipsychotics (Comparator)
Dose/intensity (mg/day)	Intervention – 156mg once/month Comparator – Not reported
Intervention setting	Clinic
Treatment length (weeks)	60
Follow-up (weeks)	Not reported
Notes. N= total number of participants; [1] Number randomised	

Table 110 Summary of findings table for paliperidone versus daily oral antipsychotics for Schizophrenia

Outcomes	No of Participants (studies)	Quality of the evidence	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with oral antipsychotics	Risk difference with IM Paliperidone versus oral

	Follow up	(GRADE)			antipsychotics (95% CI)
First-time treatment failure at post-treatment	444 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.74 (0.61 to 0.91)	537 per 1000	140 fewer per 1000 (from 48 fewer to 209 fewer)
Incidence of prolactin-related side-effects at post-treatment	445 (1 study)	⊕⊕⊖⊖ LOW[1]	RR 5.71 (2.89 to 11.28)	41 per 1000	194 more per 1000 (from 78 more to 422 more)

[1] *Alphs 2015a- Unclear risk of selection bias, no blinding, low risk of attrition bias, low risk of selective outcome bias, low risk of other bias*
[2] *The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.*

6.4.2.322 Psychological intervention

Two RCTs (N = 204) met the eligibility criteria for this review: Bond 2015 and Clayton 2013 (Bond et al., 2015; Clayton et al., 2013). The Bond 2015 study recruited participants with no competitive job placement in previous three months. The study used Individual Placement and Support (IPS) model, supported by employment specialist and the aim was to help identify and prepare for job search. On the other hand, the Clayton 2013 study included participants with severe mental illness who had a criminal charge in the 2 years prior to enrolment for the Citizenship project. The project consisted of three integrated components: individual peer mentor support (8 hours/week), an 8-week citizenship class and an 8-week valued role component. The separate analysis was performed for different intervention.

An overview of the trials included in the analysis can be found in Table 111. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 112 and Table 113. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of service utilization, and rates of self-injury.

Table 111 Study information table for trials included in the analysis of psychosocial intervention versus treatment as usual for severe mental illness

	The Citizenship Project versus TAU	Individual Placement and Support (IPS) versus Work choice/Peer support
Total no. of studies (N[1])	1(114)	1 (90)
Study ID	Clayton 2013	Bond 2015
Study design	RCT	RCT
Country	USA	USA
Underlying mental health disorder	Severe mental illness	Severe Mental Illness
Diagnosis	Symptoms	Diagnosis
Criminal justice population	Participants with a criminal charge in the 2 years prior to enrolment	Participants with self-disclosed criminal justice history and no competitive employment in the past three months
Age (mean) years	40	43.8
Gender (% female)	32	80
Ethnicity (% white)	0.31	30

	The Citizenship Project versus TAU	Individual Placement and Support (IPS) versus Work choice/Peer support
Intervention	The Citizenship project	IPS
Comparator	Treatment as usual: individual or group treatment medication monitoring, case management, or jail diversion services, as appropriate.	Work choice/Peer support
Format	Individual and Group	Individual
Dose/intensity (mg/day)	8(- 10) hours/week	Not reported
Intervention setting	Outpatients at 2 local mental health centre	Psychiatric agency providing treatment and rehabilitation services
Treatment length (weeks)	52	52
Follow-up (weeks)	Not reported	Not reported
Notes. N= total number of participants; TAU=Treatment as usual [1] Number randomised		

Table 112 Summary of findings table for the Citizenship Project versus TAU for Severe Mental Disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with The Citizenship Project versus TAU (95% CI)
Change in overall quality of life at post-treatment	114 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]		Control mean (-)0.68	MD 0.68 higher (0 to 1.36 higher)
Change in number of all convictions at post-treatment	114 (1 study)	⨁⨁⊖⊖ LOW[1,2]		Control mean (-)0.7	MD 0.05 higher (0.79 lower to 0.89 higher)
Change in Addiction Severity Index (ASI-6): alcohol composite score at post-treatment (Scale from 0 to 9; lower better)	114 (1 study)	⨁⊖⊖⊖ VERY LOW[1,3]		Control mean 0.29	MD 0.29 lower (0.57 to 0.01 lower)
Change in brief psychiatric rating Scale: emotional withdrawal symptoms at post-treatment (Scale from 1 to 7; lower better)	114 (1 study)	⨁⊖⊖⊖ VERY LOW[1,4]		Control mean 0	MD 0.28 higher (0.01 to 0.55 higher)
Change in Addiction Severity Index (ASI-6): drug composite score at post-treatment (Scale from 0 to 9; lower	114 (1 study)	⨁⊖⊖⊖ VERY LOW[1,5]		Control mean 0.04	MD 0.04 lower (0.08 lower to 0 higher)

better)

[1] Clayton 2013 - Unclear selection bias, No blinding, Unclear attrition, low risk of selective outcome reporting, low risk of other bias
[2]The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries

Table 113 Summary of findings table for Individual Placement and Support (IPS) vs Work choice models for severe mental illness

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with peer support	Risk difference with Individual Placement and Support (IPS) versus peer support (95% CI)
Competitive job placement at post-treatment	85 (1 study)	⊕⊕⊖⊖ LOW[1,2]	RR 4.44 (1.36 to 14.46)	70 per 1000	240 more per 1000 (from 25 more to 939 more)
Number of hospitalizations at post-treatment	84 (1 study)	⊕⊕⊖⊖ LOW[1,3]		Control mean 0.7 (SD 1.04)	MD 0.5 higher (0.07 lower to 1.07 higher)
Number of days in hospital at post-treatment	84 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3]		Control mean 4.93 (SD 7.59)	MD 5.51 higher (1.91 lower to 12.93 higher)

[1] Bond 2015 - Appropriate randomization with concealed allocation, no blinding of participants and care administrators, ITT analysis, appropriate outcome report
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[3] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries

6.4.2.44 Interventions for uncategorized mental health disorders

6.4.2.451 Parenting from the inside

One RCT (N=176) met the eligibility criteria for this review: Loper 2011(Loper, 2011). The intervention was based on cognitive behavioural strategy and the sessions were focused on connecting with one's own children emotionally and guiding as a parent, while they were in prison.

An overview of the trials included in the meta-analysis can be found in Table 114. Further information about both included and excluded studies can be found Appendix L.

Table 115. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 114 Study information table for trials included in the analysis of parenting from inside versus wait-list control for mental health disorders

	Parenting from inside vs wait-list control
Total no. of studies (N[1])	1 (176)
Study ID	Loper 2011
Study design	RCT
Country	USA
Underlying mental health disorder	Not reported
Diagnosis	NA
Age (mean) years	33.37
Gender (% female)	100
Ethnicity (% white)	47.5
Intervention	Parenting from inside
Comparator	Wait-list control
Format	Individual and Group
Dose/intensity	Not reported
Intervention setting	prison
Treatment length (weeks)	Not reported; The study ran for approximately 1.5 years.
Intervention (mean dose; mg/day)	Not reported
Notes. N= total number of participants; [1] Number randomised	

Table 115 Summary of findings table for parenting from inside vs Wait-list control for mental health disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with wait-list control	Risk difference with Parenting from the Inside (PFI) versus wait-list control (95% CI)
Parenting Stress Index-Modified at post-treatment (Scale from 27 to 135; lower better)	136 (1 study)	⨁⨁⊖⊖ LOW[1]	-	Control mean 2.14 (SD 0.64)	MD 0.04 higher (0.17 lower to 0.25 higher)
Parenting Alliance Measure at post-treatment (Scale from 20 to 100; higher better)	136 (1 study)	⨁⨁⊖⊖ LOW[1,2]	-	Control mean 80.01 (SD 17.46)	MD 0.31 lower (6.23 lower to 5.61 higher)
Brief Symptom Inventory (BSI): Total at post-treatment (Scale from 0 to 212; lower better)	136 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 0.75 (SD 0.82)	MD 0.2 higher (0.12 lower to 0.52 higher)

[1] *Loper 2011 - Unclear selection bias; No blinding; Unclear attrition bias, low risk of selective outcomes, low risk*

of other bias
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

6.4.2.412 Music Therapy

Two RCTs (N = 215) met the eligibility criteria for this review: Chen 2015 and Hakvoort 2013(Chen et al., 2015; Hakvoort et al., 2013). The two studies used standard care and wait-list control as comparison respectively and the underlying mental health disorders and the reported mental outcome measures were also different. Thus, separate analysis was performed for each study.

An overview of the trials included in the meta-analysis can be found in Table 116. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 117 and Table 118. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending and reoffending, service utilization, adaptive functioning and rates of self-injury.

Table 116 Study information table for trials included in the analysis of music therapy versus standard care or wait-list control for mental health disorders

	Music therapy vs Standard care	**Music therapy vs wait-list control**
Total no. of studies (N[1])	1 (200)	1 (15)
Study ID	Chen 2015	Hakvoort 2013
Study design	RCT	RCT
Country	China	Netherlands
Underlying mental health disorder	Anxiety and depression	Antisocial personality disorder
Diagnosis	Anxiety score≥49 on the State and Trait Anxiety Inventory (STAI:STAI-State or STAI-Trait) or Depression score ≥14 on Beck Depression Inventory (BDI)	60% of participants suffered from Cluster B personality disorders
Criminal justice population	Adult male inmates	Male forensic psychiatric patients enrolled at psychiatric clinics
Age (mean) years	35.5	35.6
Gender (% female)	0%	0%
Ethnicity (% white)	Not reported	Not reported
Intervention	Music therapy	Music therapy
Comparator	Standard care	Wait-list control
Format	Group	Not reported
Dose/intensity	20 (90 minutes) sessions	Not reported
Intervention setting	Prison	Prison
Treatment length (weeks)	Not reported	26
Intervention (mean dose; mg/day)	A total of 30 hours	20 sessions over 6 months

	Music therapy vs Standard care	Music therapy vs wait-list control
Notes. N= total number of participants; [1] Number randomised		

Table 117: Summary of findings table for music therapy vs standard care for depression and anxiety disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Standard care/wait-list control	Risk difference with Music therapy versus standard care/wait-list control (95% CI)
State and Trait Anxiety Inventory – State at post-treatment (Scale from 20 to 80; lower better)	184 (1 study)	⨁⨁⨁⊖ MODERATE[1]		Control mean 48.58	MD 8.05 lower (10.74 to 5.36 lower)
State and Trait Anxiety Inventory – Trait at post-treatment (Scale from 20 to 80; lower better)	184 (1 study)	⨁⨁⨁⊖ MODERATE[1]		Control mean 49.09	MD 8.51 lower (10.91 to 6.11 lower)
Brief Symptom Inventory (BSI): Total at post-treatment (Scale from 0 to 212; lower better)	184 (1 study)	⨁⨁⨁⊖ MODERATE[1]		Control mean 20.32	MD 8.81 lower (11.82 to 5.8 lower)
Rosenberg self-esteem inventory at post-treatment (Scale from 0 to 30; higher better)	184 (1 study)	⨁⨁⨁⊖ MODERATE[1]		Control mean 27.01	MD 2.26 higher (0.98 to 3.54 higher)
Texas social behaviour inventory at post-treatment (Scale from 0 to 128; higher better)	184 (1 study)	⨁⨁⨁⊖ MODERATE[1]		Control mean 96.81	MD 7.54 higher (3.24 to 11.84 higher)

[1] Chen 2015 - Appropriate randomization with proper concealment; blinding of care administrators, but not participants; ITT analysis; appropriate outcome report

Table 118 Summary of findings table for music therapy vs wait-list control for antisocial personality disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Wait-list control	Risk difference with music therapy versus wait-list control(95% CI)
ASP-1: Change in Self-management of psychiatric	13 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]	-	Control mean 0.0 (SD 0.47)	MD 0.44 higher (0.03 lower to 0.91 higher)

symptoms at post-treatment (Scale from 0 to 4; higher better)					
ASP-4: Change in self-management of assaultive symptoms at post-treatment (Scale from 0 to 4; higher better)	13 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	-	Control mean 0.75 (SD 0.35)	MD 0.11 lower (0.67 lower to 0.45 higher)
ASP-9: Change in Interpersonal skills at post-treatment (Scale from 0 to 4; higher better)	13 (1 study)	⊕⊕⊖⊖ LOW[1]	-	Control mean -0.04 (SD 0.08)	MD 0.02 higher (0.06 lower to 0.1 higher)
Change in social dysfunction and aggression scales (SDAS) at post-treatment (Scale from 0 to 44; lower better)	13 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	-	Control mean 0.8 (SD 1.48)	MD 0.8 lower (2.73 lower to 1.13 higher)
Change in forensic psychiatric profiles 40 (FP40): positive coping skills at post-treatment	13 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	-	Control mean 0.02 (SD 0.15)	MD 0.43 higher (0.12 to 0.74 higher)

[1] Hakvoort 2013 - unclear randomisation and concealment;- No blinding; available case analysis; appropriate outcome report
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

6.4.23 Economic evidence

6.4.3.1 Systematic literature review

The systematic search of the literature identified 3 studies that assessed the costs and benefits of interventions for adults with substance misuse disorders who are in contact with the criminal justice system.

Of these:

- 1 study examined the costs and benefits associated with a psychosocial intervention in the US (Daley et al., 2004)
- 2 studies examined the costs and benefits associated pharmacological interventions in Australia (Gisev et al., 2015; Warren et al., 2006)

No studies assessing the cost effectiveness of psychological, social, pharmacological or physical interventions for other disorders recommended in existing NICE guidance, for adults who are in contact with the criminal justice system, were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic review of the economic literature are described in Chapter 3; full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix S. Completed methodology checklists of the studies are provided in Appendix R. Economic evidence profiles of studies considered during guideline development

(that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix T.

6.4.3.2 Psychosocial interventions

6.4.3.2.1 Daley and colleagues (2004)

Daley and colleagues (2004) evaluated the cost effectiveness of a prison-based substance abuse treatment for incarcerated adult offenders with a substance abuse problem in the US, Connecticut. Four tiers of the substance abuse intervention were assessed: 'Tier 1' intervention involved one weekly session of drug/alcohol education for up to 6 group sessions, 'Tier 2' intervention involved 30 outpatient group sessions 3 days a week for 10 weeks, 'Tier 3' intervention involved intensive day treatment programme consisting of 4 sessions a week for 4 months or a total of 64 sessions, and 'Tier 4' intervention comprised of a residential treatment programme consisting of full-time daily treatment for 6 months in a separate housing unit. Different tiers were compared to each other and also to no intervention alternative. The economic analysis was based on an observational cohort study (N=831). Clinical effectiveness data were derived from the observational study participants. The time horizon of the economic analysis was 1 year, and its perspective was the taxpayer. Cost elements comprised intervention costs, including substance abuse and mental health treatment. Cost data were collected for the study participants from interlinked administrative records and databases, accounting data, and, as necessary, were supplemented with authors' assumptions. The primary measure of outcome utilised in the economic analysis was the likelihood of re-arrest. Regression analysis was used to adjust outcomes for baseline differences in service user characteristics including race, age, drug need score, security risk, prior arrests and other programs attended.

The mean cost per participant over 1 year was $0 for no intervention group, $189 for 'Tier 1' group, $672 for 'Tier 2' group, $2,677 for 'Tier 3' group and $5,699 for 'Tier 4' group (in likely 2003 US dollars). The adjusted probability for re-arrest with one year post-release was 45.9% for no intervention, 49.3% for 'Tier 1' group, 37.4% for 'Tier 2' group, 27.2% for 'Tier 3' group and 23.5% for 'Tier 4' group. In terms of cost effectiveness and under a public sector perspective 'Tier 1' intervention was dominated by no intervention group (that is, it was less effective and more costly). The ICER for 'Tier 2' intervention versus no intervention was $7,906 per re-arrest avoided; for 'Tier 3' versus 'Tier 2' it was $19,657 and for 'Tier 4' versus 'Tier 3' it was $81,676.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US and adopted a narrow healthcare payer perspective and has not considered wider social care costs. The measure of outcomes was not expressed in QALYs, which made interpretation of findings difficult. The study was judged by the GC to have potentially serious methodological limitations, including the relatively short time horizon (1 year), the lack of consideration of health outcomes, the study design (observational study), and source of unit cost data was unclear.

6.4.3.3 Pharmacological interventions

6.4.3.3.1 Gisev and colleagues (2015)

Gisev and colleagues (2015) evaluated the cost effectiveness of opioid substitution therapy (OST) upon prison release in New South Wales, Australia. OST treatment was compared with no OST treatment at prison release. The economic analysis was based on a retrospective matched-control study, using records of OST entrants, charges and court appearances, prison episodes, and death notifications. A total of 13,468 individuals were matched (N=6,734 in each group). The time horizon of the economic analysis was 6 months

post-release, and its perspective was the public sector (healthcare and criminal justice system). Cost elements comprised treatment, criminal justice system (court, penalties, prison), and the social costs of crime. It is unclear what social costs of crime are. However, they are likely to include physical injury, psychological trauma, a feeling of vulnerability and a fear of crime. The primary measure of outcome utilised in the economic analysis was the mortality rate.

The mean bootstrapped cost per participant at 6 months was $7,206 for OST group and $14,356 for no treatment group; a difference of -$6,353 (95% CI:–$7,568; -$5,139) (in 2012 AUD). The bootstrapped mortality rate at 6 months was 0.3% for the OST group and 0.7% for the no treatment group; a difference of -0.4% (level of statistical significance not reported). Based on the above findings OST treatment is dominant when compared with no intervention alternative. According to the cost effectiveness acceptability curve, the probability that OST post-release treatment is cost-effective is 96.7% at a willingness to pay of $500 per life saved. The results of the sensitivity analyses highlighted the robustness of the findings to the changes in the assumptions pertaining to the criminal justice system costs (for example, scenario where all 6-month costs were attributed to crime, and excluding prison costs altogether)

The study is only partially applicable to the NICE decision-making context, as it has been conducted in Australia. The measure of outcome was not expressed in QALYs. However, the intervention was found to be dominant. The study was judged by the GC to have potentially serious methodological limitations, including the relatively short time horizon (6 months), the study design (retrospective matched-control study), the lack of consideration of mental health outcomes, and the derivation of unit cost data from a mixture of national and local sources.

6.4.3.3.42 Warren and colleagues (2006)

Warren and colleagues (2006) evaluated the cost effectiveness of a prison methadone programme provided in the context of other prison health services, including counselling and related non-pharmacotherapy treatment services versus SC (no prison-based methadone intervention) in Australia, New South Wales. This was an economic modelling study with effectiveness data obtained from an RCT (N=405). The time horizon of the economic analysis was 1 year, and its perspective was a prison service provider. Cost elements comprised programme provision, including enrolment of prisoners on the programme, provision of daily methadone and associated treatment, and referral of prisoners who exit the programme to other services. Cost data were obtained from and RCT, administrative databases and published sources. The primary measures of outcome utilised in the economic analysis were the days of heroin use, deaths prevented due to substance abuse, and hepatitis C (HCV) cases avoided/delayed.

According to the analysis the intervention resulted in a mean annual cost of $3,234 per participant, in 2003 Australian dollars. SC was assigned the cost of $0 in the analysis. In terms of effectiveness the number of days of heroin use in a year was 15 and 100 in the intervention and SC group, respectively; a difference of -85 days. It was also found that the annual mortality difference was -0.71% between those receiving prison-based methadone treatment and those not receiving methadone and that provision of prison methadone for a year reduced the incidence of HCV by 0.08 cases.

Based on the above findings the ICER associated with the intervention was $38 per additional heroin free day, $458,074 per additional death avoided and $40,428 per HCV case avoided. The authors concluded that in-prison methadone was no more costly than community methadone, and provided benefits in terms of reduced heroin use in prisons, with associated reduction in morbidity and mortality (Warren et al., 2006). The GC could not judge the cost effectiveness of prison-based methadone treatment due to the lack of QALYs.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in Australia and adopted a narrow prison service provider perspective (only

intervention costs were reported). The measure of outcomes was not expressed in QALYs, which made interpretation of the findings difficult. The study was judged by the GC to have potentially serious methodological limitations, including the relatively short time horizon (1 year), the fact that some of the model inputs were based on authors' assumptions (resource use), the lack of consideration of social care and criminal justice sector costs, limited sensitivity analysis, and also the source of unit cost data was unclear.

6.4.74 Clinical evidence statements

6.4.4.81 Substance misuse

6.4.4.1.91 Psychological interventions

Low quality evidence from 1 RCT (N=95) suggested a clinically important difference between CBT alone and CBT plus contingency management in number of days with cannabis positive urine test although there was no difference in number of self-reported days with cannabis use between the two groups.

Very low quality evidence from 1 RCT (N=75) indicated uncertainty about the relative effectiveness of CBT and a 12-step program in terms of number of days with either positive urine test or positive breath analyser test. The same RCT (N=71) indicated a clinically important difference for an increase in number of days abstinent from alcohol with CBT therapy as relative to 12-step program although the effect was non-significant for number of days abstinent from drugs between the two groups.

Very low quality evidence from 1 RCT (N=44) suggested no clinically significant difference between CBT therapy and seeking safety for ASI-6 alcohol or drug composite scores, number of abstinent weeks. Although reincarceration rates were reduced by almost half with CBT, there was considerable uncertainty in the effect estimate.

Very low quality data from 1 RCT (N=27) showed no clinically important difference between CBT and wait-list control for ASI-6 alcohol composite score and uncertainty about their relative effectiveness in terms of drug composite scores and abstinence in the previous 3 months.

Very low to low quality evidence from 1 RCT (N=30) showed a clinically important reduction in ASI-6 alcohol composite score with acceptance and commitment therapy as compared to CBT therapy although the effect was not clinically significant for ASI-6 drug composite scores. The same RCT suggested uncertainty about their relative effectiveness in terms of abstinence from drugs.

Very low quality evidence from 2 RCTs showed uncertainty about the relative effectiveness of acceptance and commitment therapy and wait-list control in terms of ASI-6 alcohol (N=56) and drug (N=52) composite scores. Similarly, one RCT of low quality (N=25) reported uncertainty about the relative effectiveness of the two groups in terms of abstinence from drugs.

Very low quality data from 1 RCT (N=54) showed no clinically significant difference between mindfulness-based relapse prevention (MBRP) and cognitive behavioural therapy (CBT) for number of drug use days and short inventory problems (SIP) scores at follow-up. The same RCT also suggested a clinically important effect of MBRP for reduction of legal composite and medical composite scores of ASI as relative to CBT, the effect was not clinically significant for family-social composite and psychiatric composite scores.

Very low to low quality evidence suggested no clinically important difference between contingency management and counselling for self-reported days (2 RCTs; N=263) or urine test positive (1 RCT; N=136) for cannabis use during treatment.

Similarly, very low quality evidence of 1 RCT (N=65) reported no significant difference in ASI-marijuana scores at post-treatment and follow-up between contingency management and motivational enhancement therapy. Although the same study suggested a clinically important difference for an increase in number of cannabis use days per month at post-treatment with contingency management as relative to motivational enhancement therapy, the effect was uncertain at follow-up.

Very low quality evidence from 1 RCT (N=165) suggested uncertainty about the effectiveness of contingency management compared to psychoeducation in terms of number of participants still in treatment and number of days in treatment at follow-up.

Low quality evidence from 1 RCT (N=20) indicated uncertainty about the effectiveness of contingency compared to treatment as usual in terms of arrests for public drunkenness.

Very low quality data from 1 RCT showed clinically significant effect of motivational enhancement for an increase in percentage of self-reported days abstinent from alcohol or alcohol and drug as relative to psychoeducation at 3-months follow-up (N=238), but the effect was non-significant at 6-months (N=214) and 12-months (N=190) follow-up. Moreover, the same RCT also suggested no important difference between motivational enhancement therapy and psychoeducation for number of drinks per drinking days at 3-months (N=238), 6-months (N=214) and 12-months (N=190) follow-up.

Very low quality evidence from 1 RCT (N=136) suggested a clinically important difference between motivational enhancement therapy and CBT plus contingency management for percentage of cannabis positive urine test use during treatment although the effect was uncertain for percentage of cannabis use days.

Very low quality evidence from 1 RCT (N=27) suggested no clinical effect between motivational enhancement and CBT plus contingency management for self-reported motivation to take steps to change substance misuse scores.

Very low quality evidence from 1 RCT (N=79) suggested uncertainty about the difference in drug positive urine test (during treatment), self-reported drug or alcohol use (at 1-month follow-up) between motivational interviewing therapy and no treatment controls.

Very low quality evidence from 1 RCT (N=114) indicated no clinically significant difference between motivational interviewing and usual planning interviewing for number of self-reported days with drug use and number of days with illegal activity in past 30 days at 10 month follow-up.

Low quality evidence from 1 RCT reported uncertainty about the difference in number of drop-outs from subsequent treatment among either binge drinking group (N=23) or no binge drinking group (N=35) between motivational interviewing and no treatment control. The same RCT also suggested a clinically significant increase in the number of subsequent treatment sessions attended among binge drinking groups as relative to no treatment control although the difference was non-significant among no binge drinking group.

Very low quality evidence from 1 RCT (N=30) suggested uncertainty about the effectiveness of assessment with motivational feedback and assessment without feedback in terms of rates of speciality addiction clinic attendance.

Very low quality evidence from 1 RCT (N=128) suggested uncertainty about the relative effectiveness of group counselling and treatment as usual in terms of re-arrest, number of re-convictions (N=149), re-incarceration and number of incarcerated days (N=149) at 12-months follow-up. The same RCT of very low quality reported that although there was clinical significant difference for reduction in self-reported marijuana use with group counselling compared to treatment as usual, the effect was uncertain for self-reported LSD use, self-reported speed use and self-reported heroin use at 12-months follow-up.

Low quality evidence from 1 RCT (N=183) showed a clinically important difference between self-help journal and no intervention with fewer subsequent bookings at 12-month follow-up in the self-help group.

6.4.4.142 Pharmacological interventions

Very low quality evidence from 1 RCT showed uncertainty about the effectiveness of naloxone compared to placebo for number of participants with discontinued medication (N=97) and number of positive urine tests whilst still engaged with treatment (N=163).

Very low to moderate quality evidence from 1 RCT (N=51) showed uncertainty about the effectiveness of naltrexone compared to placebo for number of participants who retained in treatment as well as number of participants with positive urine test for alcohol, amphetamine, benzodiazepine, cocaine, marijuana or opiates while still engaged with treatment.

Very low quality evidence showed uncertainty about the effectiveness of naltrexone compared to placebo for cocaine use (N=63) and injection drug use (N=308) at post-treatment. However, very low quality evidence from 2 RCTs (N=371) suggested a clinically important reduction in opioid use with naltrexone treatment compared to placebo at post-treatment.

Low quality evidence from 1 RCT (N=44) reported uncertainty about the effectiveness of naltrexone compared to placebo in terms of number of days amphetamine, benzodiazepine or heroin use per month at 6-months follow-up.

Low to very low quality evidence indicated uncertainty about the effectiveness of naltrexone compared to placebo in terms of re-incarceration during treatment (1 RCT; N=51), post-treatment (1 RCT; N=308) and re-incarceration (1 RCT; N=44) in comparison with placebo.

One RCT of very low quality (N=63) suggested a clinically important reduction in parole violations with naltrexone compared to placebo at post-treatment although there was no difference in drug charges between the two groups. One RCT (N=44) of very low quality uncertainty about the effectiveness of naltrexone compared to placebo in terms of number of days of criminal activity per month at 6 month follow-up.

Low quality evidence from 1 RCT (N=308) reported a clinically important increase in the number of participants experiencing an adverse event at 1-year follow-up with naltrexone treatment when compared to placebo. However, there was uncertainty about the relative rates of death and non-fatal overdoses between the two groups at 1-year follow-up.

Low quality evidence from one RCT (N=382) suggested a clinically important increase in drop-outs with methadone compared to the waiting list control group.

Very low quality evidence suggested clinically important difference between methadone and control for opioid positive test at 1-month follow-up (1 RCT; N=197) and 3-months follow-up (2 RCTs; N=444). However, there was uncertainty about the differences at post-treatment (2 RCTs; N=547), 2-months follow-up (1 RCT; N=207) and 4-months follow-up (2 RCTs; N=538).

Very low quality evidence from one RCT (N=196) and moderate quality evidence from another RCT (N=382) suggested no clinically important difference between methadone and control for re-incarceration at 1-month follow-up and 4-years follow-up respectively.

Very low quality evidence from one RCT suggested uncertainty about the difference between methadone and controls for deaths (N=223) and non-fatal overdoses (N=196).

Low quality evidence from 1 RCT (N=63) showed no clinically important difference between alpha-adrenergic agonists and opioid maintenance for total withdrawal symptoms at post-treatment.

Very low to low quality evidence suggested uncertainty about the effectiveness of opioid substitution therapy compared to active intervention or placebo in terms of number of drop-outs (2 RCTs; N=206), abstinence at post-treatment (1 RCT; N=213), at 1-month follow-up (1 RCT; N=159), 3-months follow-up (1 RCT; N=94) and 6-months follow-up (2 RCTs; N=150), opioid abuse at 3-months follow-up (1 RCT; N=116), self-reported injection drug use at post-treatment and 3-month follow-up (1 RCT; N=36) as well as number of times re-arrested at 3-months follow-up, re-arrest for drug crimes at 3-months follow-up and re-incarceration at post-treatment (1 RCT; N=116).

6.4.4.193 Combined psychological and pharmacological interventions

Low quality evidence from 1 RCT (N=60) indicated uncertainty about the effectiveness of fluoxetine plus psychological therapy (CBT and motivational therapy) compared to psychological therapy alone in terms of the number of participants who failed to complete treatment. The same RCT also suggested a clinically important decrease in Hamilton depression rating scores with fluoxetine plus psychological therapy as relative to psychological therapy alone although there was no clinically important difference in Spielberger state anxiety inventory scores between the two groups.

6.4.4.194 Support and education interventions

Very low quality evidence from 1 RCT (N=34) showed that compared with treatment as usual, psychoeducation had uncertain effects on the number of days of uncontrolled drinking.

Low to very low quality evidence indicated uncertainty about the effectiveness of an employment workshop compared to treatment as usual in terms of the number of participants employed (2 RCTs; N=529) and number of days in paid employment (1 RCT; N=477).

6.4.4.195 Physical interventions

One RCT of low quality (N=158) suggested a clinically important increase in drop-out rates with acupuncture as compared to helix control (placebo acupuncture) whereas very low quality evidence from 2 RCTs (N=108) reported uncertainty about the relative effectiveness of acupuncture and other active interventions for substance misuse in terms of drug-positive urine test at post-treatment between.

6.4.4.192 Depression

Very low quality evidence from one RCT (N=38) indicated interpersonal psychotherapy had a clinically important effect on depression by HRSD scales at post-treatment when compared to a psychoeducation intervention, but this effect was uncertain at 13-weeks follow-up.

Very low quality evidence from one RCT (N=10) indicated uncertainty about the effectiveness of group cognitive treatment compared to individual support therapy in terms of depression by BDI scales, Hopeless scales, MMPI D scales and Multiple affect adjective checklist D scales at post-treatment and in depression symptoms by MMPI D scales at 39-weeks follow-up.

Very low quality evidence from one randomized study (N=158) showed that arts-based therapy had a clinically significant effect on Adult Nowicki-Strickland Locus (ANS) of control scale and depression scales by BDI in male, female and combined groups in comparison with no treatment control. Low quality evidence from this trial indicated arts-based therapy had a clinically significant effect on two of the formal elements of the arts therapy scale rating guide (FEATS), prominence of colour and colour fit when compared to no treatment control.

6.4.4.43 Vulnerable inmates with suicidal risk

One randomized study (N=46) provided very low quality evidence of a clinically significant beneficial effect of a social problem solving group on depression symptoms by either HADS

or Beck Hopeless scales at post-treatment and 13-weeks follow-up compared to no treatment control. Similarly, it is clinically significant that the social group therapy decreased anxiety symptoms by HADS scales at post-treatment and 13-weeks follow-up. Low quality evidence from this trial indicated decision making ability as measured by SPSI:R scales was improved by a clinically significant amount in the social problem solving group therapy compared to the no treatment control group.

6.4.4.4 Anxiety Disorder

Very low quality evidence from one randomized study (N=33) indicated uncertainty about the effectiveness of psychological therapy with self-help compared to wait-list control in terms of anxiety symptoms by HADS scales at post-treatment. This trial, however, reported a clinically significant reduction in anxiety with the self-help materials compared to wait-list control after 4-weeks of follow-up.

6.4.4.5 PTSD

Very low quality evidence from two randomized studies (N=40) indicated uncertainty about the effectiveness of psychological therapy (group method) compared to wait-list or no-contact control in terms of trauma symptoms by TSI scales.

Very low quality evidence from one RCT (N=123) reported a clinically significant decrease in depression symptoms by BDI scale and PTSD symptoms by PSS scales at either post-treatment or 13-weeks follow-up, increase in post-treatment generalised expectancy for success scales as well as increase in 13-weeks follow-up clinical anxiety scales with TIR intervention relative to wait-list control. The clinical effects on generalised expectancy for success scales at 13-weeks follow-up and clinical anxiety scales at post-treatment were uncertain and of very low quality.

One RCT provided moderate quality evidence of uncertainty about the effectiveness of TARGET intervention compared to SGT intervention in terms of post-treatment PTSD symptoms by CAPS scales (N=72) and Heartland forgiveness scales (N=32; low quality evidence).

Very low to low quality evidence from one randomized study (N=9) suggested a clinically important beneficial effect of focused group therapy on global severity index, positive symptom distress index and total positive symptom index by symptom checklist-90R compared to wait-list control at post-treatment.

Very low quality evidence from one RCT (N=31) indicated uncertainty about the effectiveness of group therapy compared to no contact control in terms of IIP-32 scales.

6.4.4.6 ADHD

Very low quality evidence from two RCTs (N=84) suggested a clinically important effect of methylphenidate on reduction in ADHD symptoms by CAARS:OSV scales at post-treatment as compared to placebo. Low quality evidence from one RCT (N=20) indicated a clinically significant reduction in ADHD symptoms at 3-years follow-up with methylphenidate compared to placebo.

Very low quality evidence from one RCT (N=54) indicated uncertainty about the effectiveness of methylphenidate compared to placebo in terms of the number of participants with drug negative urine at post-treatment.

6.4.4.7 Antisocial personality disorder

Very low quality evidence from one randomized study (N=12) indicated a clinically significant decrease in aggressive response with tiagabine compared to placebo at post-treatment.

However, there was uncertainty about difference in the number of adverse effects in the two groups.

6.4.4.38 Severe mental illness

6.4.4.8.41 Pharmacological intervention

Very low evidence from one RCT (N=445) suggested a clinically significant reduction in first-time treatment failure rate with IM paliperidone compared to oral antipsychotics. There was low quality evidence of a clinically significant increase in the risk of prolactin-related side-effects, however, with IM paliperidone.

6.4.4.8.92 Psychosocial intervention

Low to very low quality evidence from one RCT (N=114) of uncertainty about the effectiveness of the Citizenship Project compared to TAU in terms of quality of life, number of convictions and addiction severity at post-treatment. Low quality evidence form this trial indicated a clinically significant difference between the groups in terms of alcohol composite score and withdrawal symptoms at post-treatment.

Low quality evidence from one RCT (N=85) indicated that the participants in the IPS intervention group were more than four times more likely to get competitive job placement than work choice group. However, there was uncertainty about the relative effectiveness of IPS and work choice in terms of the number of hospitalizations and number of days in hospital Interventions for uncategorized mental health disorders

6.4.4.8.03 Parenting from the inside

Very low to low quality evidence from one randomized study (N=136) indicated no clinically significant difference between parenting from inside intervention and treatment as usual for parenting stress index, parenting alliance and total brief symptom inventory scales.

6.4.4.8.44 Music therapy

Moderate quality evidence from one randomised study (N=184) suggested that music therapy had a clinically significant effect compared to standard care decreasing anxiety symptoms on state and trait anxiety measures compared to standard care. Music therapy increased self-esteem as measured by the Rosenberg self-esteem inventory and social behaviour as measured by the Texas social behaviour inventory by clinically significant amounts.

Very low to low quality evidence from one randomised study (N=13) indicated uncertainty about the effectiveness of music therapy compared to wait-list control in terms of self-management psychiatric symptoms, self-management assaultive symptoms, interpersonal skills, social dysfunction and aggression. However, very low quality evidence from this trial indicated a clinically important increase in positive coping skills as measured by FP40 scales in the music therapy group.

6.4.75 Economic evidence statements

There was evidence from 1 US study on the cost effectiveness of psychosocial prison-based interventions for people with substance abuse problems. The cost effectiveness analysis was based on an observational cohort study (N=831). It was found that intensive outpatient group treatment (3 days a week for 10 weeks) results in an ICER of $7,906 per re-arrest avoided (when compared with ‘no intervention’ option), and intensive day treatment programme (consisting of 4 sessions a week for 4 months or a total of 64 sessions) results in an ICER of $19,657 (when compared with an outpatient group treatment). This evidence was US-based and is only partially applicable to the NICE decision making-context and is characterised by

potentially serious methodological limitations, including the relatively short time horizon (1 year), the lack of consideration of health outcomes, the study design (observational study), and source of unit cost data was unclear. Due to the lack of QALYs the GC could not judge the cost-effectiveness of psychosocial prison-based interventions in adult offenders with substance abuse problems.

There was mixed evidence from 2 Australian studies on the cost effectiveness of pharmacological treatments for substance abuse problems in incarcerated adult offenders. One cost-effectiveness analysis based on a retrospective observational matched-control study (N=13,468) found opioid substitution therapy (OST) upon prison release to be dominant when compared to no OST treatment. It resulted in cost savings from a public sector perspective and fewer deaths at 6 month follow-up. Another cost-effectiveness analysis based on economic modelling (with effectiveness data from an RCT) found that prison-based methadone programme provided in the context of other prison health services, including counselling and related non-pharmacotherapy treatment services (when compared with no prison-based methadone intervention) resulted in the ICERs of: $38 per additional heroin free day, $458,074 per additional death avoided, and $40,428 per additional hepatitis C case avoided. This evidence is from Australian studies and is only partially applicable to the NICE decision-making context. Outcomes were not reported in the form of QALYs, which made judgements on cost effectiveness difficult, although in one of the studies judgement on cost effectiveness was straightforward since the intervention was found to be dominant. Both studies are characterised by potentially serious limitations, 1 study adopted retrospective matched-control study design, relatively short time horizons (6 months and 1 year), and lack of use of national unit costs.

6.5 Recommendations and link to evidence

Recommendations	**36. Use this guideline with NICE guidelines on any specific mental health problems when available. Take into account:** • **the nature and severity of any mental health problem** • **the presence of a learning disability or any acquired cognitive impairment** • **other communication difficulties (for example, language, literacy, information processing or sensory deficit)** • **the nature of any coexisting mental health problems** • **limitations on prescribing and administering medicine (for example, in-possession medicine) or the timing of the delivery of interventions in certain settings (for example, prison)** • **the development of trust in an environment where health and care staff may be held in suspicion** • **any differences in presentation of mental health problems** • **the treatment setting (the person's home, in the community, primary or secondary care health services, mental health or learning disabilities services, and prison).** **37. Refer to relevant NICE guidance for the psychological treatment of mental health problems for adults in contact with the criminal justice system, taking into account:**

	• **the need to modify the delivery of psychological interventions in the criminal justice system** • **the need to ensure continuity of the psychological intervention (for example, transfer between prison settings or on release from prison).** 38. **Practitioners should consider using contingency management to reduce drug misuse and promote engagement with services for people with substance misuse problems.** 39. **Practitioners delivering contingency management programmes should:** • **agree with the person the behaviour that is the target of change** • **provide incentives in a timely and consistent manner** • **confirm the person understands the relationship between the treatment goal and the incentive schedule** • **make incentives reinforcing and supportive of a healthy and drug-free lifestyle.** 40. **Refer to relevant NICE guidance for pharmacological interventions for mental health problems in adults in contact with the criminal justice system. Take into account:** • **risks associated with in-possession medicines** • **administration times for medication** • **availability of medicines in the first 48 hours of transfer to prison** • **availability of medicines after release from prison.** 41. **Refer to NICE's guidance on attention deficit hyperactivity disorder (ADHD) when prescribing pharmacological interventions for this condition.** 42. **Review all medicines prescribed for sleep problems and the management of chronic pain to:** • **establish the best course of treatment (seek specialist advice if needed)** • **assess the risk of diversion or misuse of medicines.**
Relative values of different outcomes	The GC were mindful that the primary aim of the interventions covered in this review question was to improve substance misuse and mental health outcomes and therefore remission (and relapse - and its prevention - for those had remitted) from the disorder and improvement in symptomatology were seen as critical outcomes. The GC were also mindful of the link between mental health problems and offending (e.g. as may be the case in substance misuse) and so also considered offending as a potentially important outcome. Given the challenge of engaging individuals in contact with the criminal justice system in treatment engagement in treatment was also considered.
Trade-off between	In assessing the trade-off between benefits and harms in the interventions

clinical benefits and harms	covered in this protocol the GC were particular interested in any evidence for an intervention that was specifically developed for use in the criminal justice system that demonstrated a benefit greater than might be expected from the use of an intervention recommended in other NICE mental health guidelines for the disorder or problem that was the target of the interventions. When making this judgement the GC drew on their knowledge of and considered relevant NICE guidance. **Substance misuse –** The GC were aware that some of the interventions reviewed although having some evidence of benefit (e.g. reduced drug misuse) in this review including psychological interventions such as cognitive behavioural therapy and psychoeducation, pharmacological interventions such as naltrexone and the combination of psychological interventions with drugs such as fluoxetine. The GC did not think that there were significant harms associated with the use of psychological interventions but did not that for pharmacological interventions there was a risk of particularly with the pharmacological interventions where the illicit use of prescribed medications are associated with risks of harm including death from accidental overdose. They noted that contingency management is a brief intervention, which is simple to implement and has limited potential to harm. **Individuals with suicidal risk** – One study on social problem-solving intervention compared to no treatment control and found an improvement in mental health outcomes (depression, anxiety and decision making ability). The GC, however, took the view there was no reported evidence of a direct impact on suicidal behaviour therefore decided not to make a recommendation. **Self-help for anxiety disorders** – One study compared the use of self-help materials with a wait-list control group and found a small benefit on anxiety symptoms, the evidence was rated to be of very low quality. The GC noted that there was evidence to support self-help in non-criminal justice population with anxiety disorders but did not think that there was sufficient evidence to recommend the specific intervention under review. **PTSD** – The evidence on effectiveness of focused group psychological therapy (either group or individual) found no clinically important difference and no indication of harm, compared with wait-list/supportive group therapy in studies which were of moderate to very low quality. As such, the GC did not think that there was sufficient evidence to recommend the specific interventions under review. **ADHD** – Two randomised studies found no clinical benefit in the effect of per oral methylphenidate compared to placebo. The GC commented that the evidence was inconclusive and did not make any recommendations that would vary from existing NICE recommended interventions for ADHD. They noted that the prescription of oral methylphenidate in the CJS could result, through onward sale of the drug, in the illicit use of the drug by individuals for whom it was not prescribed leading to the possibility of harm. **Antisocial personality disorder** –There was one trial on the reduction of aggression by tiagabine treatment compared to placebo. The evidence was of very low quality. It is not licensed for use as a mood stabiliser or for impulse control in the UK and may lower the seizure threshold in people without epilepsy and therefore, the GC did not recommend tiagabine for use in the CJS **Severe mental illness** – One community RCT found that paliperidone

	palmitate injection was more effective in reduction of first-time treatment failure (defined by a range of outcome indicators) than oral antipsychotics. The GC noted the potential harms associated with the use of depot medication (e.g. increased social withdrawal) and the specific indications for its use in the NICE Schizophrenia Guideline (CG155). Given the range of other, possibly less costly, depot injections available and that its license was only for people who had previously responded to responsive to paliperidone or risperidone. The GC decided not to make a specific recommendation for paliperidone palmitate. The evidence on psychological intervention (the Citizenship project) and IPS intervention for quality of life, mental health outcomes and substance misuse outcomes was inconclusive. The GC did not think that there was sufficient evidence to recommend the intervention. **Parent training** – There was no clinical difference on mental health outcomes on the effect of the 'parenting from inside' intervention compared to wait-list controls. The GC did not think that there was sufficient evidence to recommend the specific interventions under review. **Anxiety and depressive symptoms** - One RCT which compared the effectiveness of group counselling (with or without video feedback) with treatment as usual among prisoners in minimal community unit. The GC noted that the evidence was of very limited quality to recommend a change. Music therapy can improve anxiety and depression symptoms, however, the evidence was limited to two small studies non-UK setting and the GC identified no comparable recommendations in NICE guidance for depressive or anxiety disorders There was no indication of harm in these studies, Given the low quality of evidence and absence of evidence in other relevant NICE mental health guidelines, the GC did not think that there was sufficient evidence to recommend the specific interventions under review.
Trade-off between net health benefits and resource use	Existing economic evidence on psychological interventions for people who are in contact with the criminal justice system was limited to 1 non-UK study that found that intensive outpatient group treatment (3 days a week for 10 weeks) and intensive day treatment programme (consisting of 4 sessions a week for 4 months or a total of 64 sessions) may potentially be cost effective for the treatment of adult offenders with a substance abuse problem. There was no economic evidence on psychological interventions for the management of other mental health problems in adults in contact with the criminal justice system. Existing economic evidence on pharmacological interventions for people who are in contact with the criminal justice system was limited to non-UK studies and were only for substance abuse treatment in prison setting. Existing evidence indicated that prison-based pharmacological treatments are associated with reduced rates of re-offending, reduced incidence of HCV, improved survival and as a result may potentially be cost effective in people with substance misuse who are in prisons. There was no evidence on pharmacological interventions for the management of other mental health problems in adults in contact with the criminal justice system. The GC considered the economic consequences arising from the presence of mental health problems in people who are in contact with the criminal justice system that is associated with the consumption of extra healthcare resources. The GC also considered the impact of mental health problems on the mortality (increased risk of suicide) and HRQoL and concluded that the provision of effective psychological and pharmacological interventions for the management of mental health problems is likely to improve survival and HRQoL in this population. If untreated, the symptoms are likely to get worse and require the management of mental health problems in more resource-intensive settings, such as secondary care or require expensive

	crisis care. Also, once released back in the community, service users with untreated mental health problems are likely to have repeat interface with the criminal justice system, because their problems are likely to be getting even worse. The GC also considered the impact of potential self-harm and suicide on the HRQoL of family members. All of the above are likely to result in a significant increase in healthcare, social care, and criminal justice sector costs. The GC expressed the opinion that, for safety reasons, people with mental health problems in criminal justice settings receiving pharmacological treatments may benefit from a closer monitoring. The GC concluded that additional monitoring would ensure that service users received adequate and effective treatment. The GC acknowledged that provision of pharmacological interventions to people who are in contact with criminal justice sector may be more resource-intensive compared with provision of pharmacological interventions in the general population, and this may have implications for the cost effectiveness of such interventions, but considered that additional monitoring and support, and further adaptations in the pharmacological treatment of people with mental health problems who are in contact with criminal justice system are essential in order to achieve a positive outcome. There was no economic evidence on contingency management for people with substance misuse. The GC acknowledged that provision of such programmes to people who are in contact with criminal justice sector may require additional resources (that is, costs associated with the voucher/prize incentives, urine testing, etc.,). The GC also considered the economic consequences arising from the presence of mental health problems in people who are in contact with the criminal justice system that is associated with the consumption of extra healthcare resources. The GC expressed the view that the additional costs of the interventions are very likely to be justifiable by the potential improvements in mental health outcomes and potential reduction in reoffending (the link between illicit drug use and crime is well established) as was demonstrated in the NICE Guideline on psychosocial interventions in Drug Misuse. The GC also considered issues relating to equality, and judged that psychological (including contingency management) and pharmacological interventions for the management of mental health problems that have been shown to be cost effective in general population should also be offered to people with mental health problems who are in contact with the criminal justice system, following necessary adaptations and additional monitoring.
Quality of evidence	Most of the evidence reviewed was of very low to low quality. There were a large number of small studies, for example in substance misuse which used a broad range of different interventions and comparators which limited the extent to which data could be pooled. Some of the better quality evidence reviewed was for contingency management. The GC noted that just because the reviewed evidence showed no effect, this does not mean conclusively that an intervention has no benefit, it may reflect the limited nature of the current evidence for the treatment of these disorders in the criminal justice system.
Other considerations	The GC were aware of the need for effective interventions for individuals in contact with the CJS that are in line with existing NICE guidance for the general population. The GC identified nothing in the reviews undertaken for this guideline which would suggest that current available treatment would not be of benefit to this in contact with the CJS. In addition, they identified no significant harms, save for the possible diversion of prescribed drugs (e.g. methylphenidate) which would be of significant concern.

	With this in mind, the GC developed through informal consensus a set of principles of which would guide the use of NICE guidance on mental interventions in the CJS and identifying, where necessary, where specific modifications to the intervention or the manner in which it is delivered were considered. The GC were of the view that when using NICE mental health guideline in the CJS the degree of the degree of learning disabilities or acquired cognitive impairment, communication difficulties, coexisting mental health problems, limits on the limitations on prescribing and administering medicine (e.g. in-possession medicine), the development of trust in an environment where health and care staff may be held in suspicion, differences in presentation of mental health problems and the treatment setting should be borne in mind. Other principles concerned the modification of the delivery of psychological interventions and the need to ensure continuity of psychological care across the pathway. For pharmacological interventions the GC recognised the need to modify drug prescribing to take into account in-possession medication, the administration times for medication, the availability of medicines in the first 48 hours of transfer to prison and the availability of medicines after release from prison. They also wanted to draw attention the importance of proper prescribing for the management of attention deficit hyperactivity disorder, sleep problems and chronic pain management

6.5.11 Research recommendation

4. What is the effectiveness of a structured clinical or case management to improve mental health outcomes using interventions within community rehabilitation centres and national probation services??

Many individuals in contact with the CJs in particularly those managed by community rehabilitation companies have significant personality problems and interpersonal difficulties. Evidence from people with such problems in general mental health services suggest that structure mental health services may be of benefit in improving mental health outcomes. A programme of research which would (a) refine the structured clinical management for use in the CRCs and then (b) test this in a large scale randomised control trial should be undertaken. The comparison should be against standard CRC care. The trail should consider both clinical outcomes (as detailed below) and cost-effectiveness.

Important outcomes could include:

- Offending and re-offending rates
- Mental health outcomes
- Cost-effectiveness
- Health-related quality of life

6.6 Review question: For adults with a paraphilic disorder who are in contact with the criminal justice system, what are the benefits and harms of psychological, social or pharmacological interventions aimed at reducing or preventing the expression of paraphilic behaviour, or preventing or reducing sexual offending or reoffending?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 119. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 119: Clinical review protocol summary for the review on interventions aimed at reducing or preventing the expression of paraphilic behaviour, sexual offending or reoffending in adults with a paraphilic disorder who are in contact with the criminal justice system

Component	Description
Population	Included: • Adults (aged 18 and over) with a paraphilic disorder who are in contact with the criminal justice system Excluded: • people who are cared for in hospital, except for providing guidance on managing transitions between criminal justice system settings and hospital • people in immigration removal centres • children and young people (aged under 18 years) • people who are in contact with the criminal justice system solely as a result of being a witness or victim.
Intervention(s)	Psychological and social interventions: • behavioural interventions (aversion therapy, imaginal desensitisation, covert sensitisation or olfactory conditioning) • cognitive analytic therapy (CAT) • CBT (group or individual) • milieu therapy • motivational interviewing • multisystemic therapy • psychodynamic or psychoanalytic psychotherapy • psychoeducational interventions, including psychologically (CBT or IPT)-informed • psychoeducation (Sex Offender Treatment Programmes [SOTP]) • reintegration programmes (circles of support and accountability) schema therapy • therapeutic communities Pharmacological interventions: • antiandrogen hormone therapy (cyproterone acetate, medroxyprogesterone acetate) • antidepressants (SSRIs) • antipsychotic medication (benperidol) • gonadotropin-releasing hormone agonists (triptorelin)
Comparison	• Treatment as usual • No treatment • Waitlist control • Placebo (including attention control) • Any alternative management strategy
Outcomes	• Critical – Offending and re-offending; • Important – Mental health symptoms; Service utilisation Adaptive functioning (for example, employment status, development of daily living and interpersonal skills, and quality of life); Beliefs and attitudes regarding sexual offending; Acceptability of interventions (e.g. attrition from study arms)
Study design	Systematic reviews of RCTs and RCTs (including

Component	Description
	crossover randomised trials if data from the first phase is available) If the RCT evidence is limited either in terms of numbers of RCTs or numbers of included participants (≤100), the range of included studies was expanded to include non-randomised studies.

6.6.11 Clinical evidence

6.6.12.1 Pharmacological interventions

Three RCTs (N = 84) met the eligibility criteria for this review. These trials were identified in a systematic review (Khan et al., 2015) of seven RCTs, four of which were excluded from this review because they involved psychiatric inpatients or were cross-over trials where the first phase data could not be extracted.

The included trials involved medroxyprogesterone acetate (MPA) a synthetic progesterone proposed to supress sexual desire by countering the libidinal effects of testosterone. Hucker et al. (1988) compared MPA with placebo for paedophilia in the outpatient setting. Langevin et al. (1979) and McConaghy et al. (1988) examined the addition of MPA to outpatient psychological interventions for exhibitionism or varied paraphilic disorders respectively.

An overview of the trials included in the meta-analysis can be found in Table 120. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 121 and Table 122. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 120: Study information table for trials included in the meta-analysis of pharmacological interventions for paraphilia

	MPA vs Placebo	MPA+Psych vs Psych alone
Total no. of studies (N[1])	1 (18)	2 (66)
Study ID	Hucker 1988	(1) Langevin 1979 (2) Mcconaghy 1988
Study design	RCT	RCT
Country	Canada	(1) Canada (2) Australia
Diagnosis	Paedophilia	(1) Exhibitionism (n=35) (2) Exhibitionism (n=12); Paedophilia (n=15); Festishism (n=5); Transvestism (n=4); Voyeurism (n=3)
Age in years (mean)	40.5	(1) Not reported (2) 30
Sex (% female)	0%	0%
Ethnicity (% white)	Not reported	Not reported
Diagnostic status	Symptoms	(1) Symptoms (2) Diagnosis (DSM-III)
CJS setting	Outpatient (sentenced and probation)	(1, 2) Outpatient clinic
Treatment length (weeks)	12	(1) 15 (drug+psych) (2) 52 (drug); 1 (psych)

	MPA vs Placebo	MPA+Psych vs Psych alone
Intervention (mean dose)	200 mg/day	(1) MPA – 150mg/fortnight; Psych- 1hr/week (2) MPA – 150mg/fortnight to 150mg/month; Psych: 5 sessions over 5 days
Comparison	Placebo	(1) Assertion training alone (2) Imaginal desensitization alone
Notes. N= total number of participants; MPA = Medroxyprogesterone acetate; Psych= Psychosocial; mg/day = milligrams per day; CJS = Criminal Justice System [1] *Number randomised.*		

Table 121: Summary of findings table for medroxyprogesterone + psychological intervention compared to psychological intervention only for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with psych intervention only	Risk difference with Medroxyprogesterone + psych intervention
Repetition of anomalous behaviour assessed with: self-report questionnaire and case notes	52 (2 RCTs)	⊕◯◯◯ VERY LOW [1,2]	RR 0.58 (0.04 to 8.30)	222 per 1,000	93 fewer per 1,000 (213 fewer to 1,622 more)
Dropout assessed with: number of participants who did not complete treatment	32 (1 RCT)	⊕⊕◯◯ LOW [4,5]	RR 2.27 (1.00 to 5.14)	294 per 1,000	374 more per 1,000 (0 to 1,000 more)
Reduced anomalous desires assessed with: self-report questionnaire	20 (1 RCT)	⊕◯◯◯ VERY LOW [2,3,4]	RR 0.83 (0.12 to 1.55)	600 per 1,000	102 fewer per 1,000 (528 fewer to 330 more)

1. *Downgraded for inconsistency*
2. Confidence interval of the effect estimate includes appreciable benefit, harm and no effect
3. High risk of selection and performance bias
4. High risk of selection and performance bias.

Table 122: Summary of findings table for medroxyprogesterone compared to placebo for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Medroxyprogesterone
Reduced anomalous desire assessed with: self-report questionnaire	20 (1 RCT)	⊕⊕◯◯ LOW [1,2,3]	RR 0.50 (0.17 to 1.46)	600 per 1,000	300 fewer per 1,000 (498 fewer to 276 more)

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Medroxyprogesterone
follow up: 52					
Reduced anomalous behaviour assessed with: self-report questionnaire follow up: 52	20 (1 RCT)	⨁⨁◯◯ LOW [1,2,3]	RR 0.33 (0.04 to 2.69)	300 per 1,000	201 fewer per 1,000 (288 fewer to 507 more)

1. .High risk of performance and attrition bias.
2. Optimal information size criterion not met (event rate less than 300)
3. Confidence interval for the effect estimate spans both MID thresholds (0.80 to 1.25).

6.6.112 Psychoeducational interventions

Three RCTS (N=779) and 23 controlled non randomised studies (N=12317) met the eligibility criteria for this review. The randomised trials were identified in a systematic review (Dennis et al., 2012): Anderson Varney (1991) and Hopkins (1991) compared CBT based psychoeducational interventions to treatment as usual or waiting list control. Marques et al. (1994a) examined the California Sex Offender Treatment and Evaluation Project. No trials published after this systematic review were found.

The non-randomised controlled trials (Abracen et al., 2011; Aytes et al., 2001; Craissati & McClurg, 1997; Craissati et al., 2009; Di Fazio et al., 2001; Duwe & Goldman, 2009; Friendship et al., 2003; Hanson et al., 2004; Looman et al., 2000; Lowden et al., 2003; Marshall et al., 2008; McGrath et al., 2003; McGrath et al., 1998; McGuire, 2000; O'Reilly et al., 2010; Olver et al., 2013a; Redondo Illescas & Garrido Genoves, 2008; Ruddijs & Timmerman, 2000; Song & Lieb, 1995; Stalans et al., 2001; Turner et al., 2000) involved primarily group CBT based psychoeducation (including SOTP). Content of the psychoeducation included: offence disclosure, accepting responsibility, cognitive distortions/cognitive restructuring, victim empathy, offending cycle, individual risk factors and recognition cues, relapse prevention and social skills. Methods included group discussion, exposure to video or audio accounts presented by victims, positive modelling, role-play, skills practice and decision matrices. The control groups received either no treatment, treatment as usual (which was not specified) or were waitlist controls.

An overview of the trials included can be found in Table 123. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 124 and Table 125. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Significant heterogeneity was noted in some of the outcomes and subgroup analysis was done according to country (which reduced heterogeneity) and reported in the summary of findings tables. No other sources of heterogeneity were identified: the non-randomized controlled trials were typically not adjusted for confounders.

Table 123: Study information table for trials included in the analysis of psychoeducational interventions for paraphilia

	Psychoeducational interventions versus treatment as usual, no treatment or waiting list
Total no. of studies (N)	25 (13096)

	Psychoeducational interventions versus treatment as usual, no treatment or waiting list
Study ID	(1) Abracen 2011 (2) Anderson-Varney 1991 (extracted from Dennis 2012) (3) Aytes 2001 (4) Craissati 1997 (5) Craissati 2009 (6) Di Fazio 2001 (7) Duwe 2009 (8) Friendship 2003 (9) Hanson 2004 (10) Hopkins 1991 (extracted from Dennis 2012) (11) Illescas 2008 (12) Looman 2000 (13) Lowden 2003 (14) Marques 1994a/1994b/2005/Miner 1990 (15) Marshall 2008 (16) McGrath 1998 (17) McGrath 2003 (18) McGuire 2000 (19) O'Reilly 2010 (20) Procter 1996 (21) Ruddijs 2000 (22) Scalora 2003 (23) Song 1995 (24) Stalans 2001 (25) Turner 2000 (26) Olver 2013
Study design	(2,10,14) RCT, (all others) non-randomised controlled trials
Country	Canada (1, 6, 9, 12, 15, 26), Ireland (19), Netherlands (21), Spain (11), UK (4, 5, 8, 10, 20), US (2, 3, 7, 13, 14, 16, 17, 18, 22, 23, 24, 25)
Diagnosis	Paraphilic disorder. Proportion sex offenders against children: <50% (6, 12, 16, 17), 73% (5), 73% (15), 76% (21), 79% (23), 90% (19), 90% (20), 100% (2, 4, 22, 24), not reported (1, 3, 7, 8, 9, 10, 11, 13, 14, 18, 25, 26)
Age in years (mean)	27.9 (12), 32.4 (24), 34 (21), 34.3 (25), 34.4 (23), 34.9 (7), 35.2 (6), 35.3 (3), 35.9 (16), 36.2 (19), 36.2 (22), 36.5 (13), 37.4 (9), 38 (20), 38.2 (17), 38.7 (2), 41.5 (26) 49.1 (15), not reported (1, 11, 4, 5, 8, 10, 14, 18)
Sex (% female)	0% (2, 4, 5, 6, 8, 9, 10, 12, 13, 15, 16, 17, 19, 20, 21, 22, 23, 25) Not reported (1, 3, 7, 11, 14, 18, 24, 26)
Ethnicity (% white)	24% (24), 51% (13), 65% (7), 72% (22), 82% (2), 87% (23), 96% (3), 99% (16), 99% (17), not reported (1, 4, 5, 6, 8, 9, 10, 11, 12, 14, 15, 18, 19, 20, 21, 25)
Diagnostic status	Unclear
CJS setting	In the community (3, 4, 5, 9, 16, 20, 21, 24, 25), inpatient (1, 6, 12, 14, 22), prison (2, 7, 8, 10, 11, 13, 15, 17, 18, 19, 23, 26)
Treatment length (weeks)	2 (20), 6 (10), 7 (15), 8 (2), 20 (18), 24 (6), 43 (19), 44 (25), 59 (5), 123 (22), 130 (3), 166 (17), 177 (16), 26 (13), 39-52 (4), 43-52 (11), 52-208 (23), 61-130 (14), Not reported (1, 7, 8, 9, 12, 21, 24, 26)
Intervention	CBT informed psychoeducation (2, 3, 4, 5, 10, 11, 16, 17, 20, 21, 22), sex offender treatment programmes (1, 6, 8, 9, 12, 13, 14, 18, 19, 23, 24, 25, 26) preparatory programme for SOTP (15)
Delivery method	Face to face (2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26), not reported (1)

	Psychoeducational interventions versus treatment as usual, no treatment or waiting list
Comparison	No treatment (1, 3, 5, 7, 11, 12, 13, 14, 19, 20, 21, 22, 23, 25, 26), treatment as usual (not specified: 2, 4, 6, 8, 9, 15, 16, 17, 18, 24), waiting list control (10)

CBT, cognitive behavioural therapy; SOTP, sex offender treatment programme; RCT, randomised controlled trial

Table 124 Summary of findings table (RCTs) for psychoeducational interventions, principally CBT-informed psychoeducation (including SOTP) versus treatment as usual, no treatment or waitlist control for paraphilic disorders.

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual, no treatment or waitlist control	Risk difference with psychoeducational intervention
Cognitive distortions (Abel and Becker Cognition Scale [ABCS]) (Scale from 26 to 130; higher better)	60 (1 RCT)	⊕⊕⊕◯ MODERATE [1]	-	The mean cognitive distortions (Abel and Becker Cognition Scale [ABCS]) - RCT was 134.53	MD 13.43 lower (20.05 lower to 6.81 lower)
Acceptance of accountability (Multiphasic Sex Inventory [MSI]: Justifications)	60 (1 RCT)	⊕⊕◯◯ LOW [1,2]	-	-	MD 0.8 lower (6.13 lower to 4.53 higher)
Sexual anxiety (Multiphasic Sex Inventory [MSI]: Sexual inadequacies)	60 (1 RCT)	⊕⊕◯◯ LOW [1,3]	-	The mean sexual anxiety (Multiphasic Sex Inventory [MSI]: Sexual inadequacies) - RCT was 48.33	MD 6.2 lower (13.43 lower to 1.06 higher)
Anxiety (Social Avoidance and Distress Scale, SADS) (Scale from 0 to 28; lower better)	75 (2 RCTs)	⊕⊕◯◯ LOW [1,4,5]	-	The mean anxiety (Social Avoidance and Distress Scale, SADS) - RCT was 10.5	MD 2.19 lower (7.31 lower to 2.92 higher)
Violent reconviction	233 (1 RCT)	⊕⊕◯◯ LOW [6,7]	RR 1.16 (0.80 to 1.69)	102 per 1,000	62 fewer per 1,000 (90 fewer to 30 more)
Sexual reconviction	480 (1 RCT)	⊕⊕◯◯ LOW [6,7]	RR 0.39 (0.12 to 1.29)	189 per 1,000	30 more per 1,000 (38 fewer to 130 more)

1. Anderson-Varney 1991 - unclear risk of selection bias; no blinding; low risk of attrition bias; low risk of selective outcome bias; low risk of other bias
2. The MID calculated from SD of control was +/-5.41.
3. The MID calculated from SD of control was +/-6.01.
4. Hopkins 1991 - Unclear selection bias; No blinding; low risk of attrition bias; low risk of selective outcome bias; low risk of other bias.

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual, no treatment or waitlist control	Risk difference with psychoeducational intervention
5.Hopkins 1991 - Participants involved roughly equal numbers of incarcerated paedophile and rapists *6.Indirectness – inpatient setting* *7. Imprecision – the CI for the effect spans no effect and both MID thresholds*					

Table 125: Summary of findings table (observational studies) for psychoeducational interventions, principally CBT-informed psychoeducation (including SOTP) versus treatment as usual, no treatment or waitlist control for paraphilic disorders.

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual, no treatment or waitlist control	Risk difference with psychoeducational intervention
Reconviction (Any)	2796 (9 observational studies)	⊕◯◯◯ VERY LOW [1,2,3,4,5,6,7,8,9,10,11,12]	RR 0.49 (0.30 to 0.82)	382 per 1,000	195 fewer per 1,000 (267 fewer to 69 fewer)
Reconviction (Any) - UK studies	338 (1 observational study)	⊕◯◯◯ VERY LOW [2,11]	RR 0.21 (0.15 to 0.31)	1,000 per 1,000	790 fewer per 1,000 (850 fewer to 690 fewer)
Sexual reconviction	5261 (11 observational studies)	⊕◯◯◯ VERY LOW [2,3,4,5,6,8,9,10,11,12,13,14,15,16]	RR 0.66 (0.47 to 0.93)	70 per 1,000	24 fewer per 1,000 (37 fewer to 5 fewer)
Sexual reconviction - UK studies	2885 (3 observational studies)	⊕◯◯◯ VERY LOW [2,12,13,16]	RR 0.96 (0.64 to 1.44)	36 per 1,000	1 fewer per 1,000 (13 fewer to 16 more)
Violent reconviction	2181 (6 observational studies)	⊕◯◯◯ VERY LOW [12,16]	RR 0.62 (0.40 to 0.96)	261 per 1,000	99 fewer per 1,000 (157 fewer to 10 fewer)
Violent reconviction - UK studies	240 (1 observational study)	⊕◯◯◯ VERY LOW [8,9,12]	RR 0.70 (0.36 to 1.36)	166 per 1,000	50 fewer per 1,000 (106 fewer to 60 more)
Revocation	2186 (5 observational studies)	⊕◯◯◯ VERY LOW [1,8,11,12,16,17,18,19]	RR 0.66 (0.35 to 1.23)	410 per 1,000	140 fewer per 1,000 (267 fewer to 94 more)
Revocation - UK studies	240 (1 observational study)	⊕◯◯◯ VERY LOW [16]	RR 0.31 (0.14 to 0.66)	241 per 1,000	167 fewer per 1,000 (208 fewer to 82 fewer)
1. Stalans 2001 - Controlled Non-RCT; significant group differences at baseline in current offence and on prior criminal history; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias					

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual, no treatment or waitlist control	Risk difference with psychoeducational intervention

2. Friendship 2003 - Controlled Non-RCT; confounders controlled in analysis; no blinding; unclear risk of attrition bias; high risk of selective outcome bias; low risk of other bias
3. Ruddijs 2000 - Controlled Non-RCT; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
4. Marshall 2008 - Controlled Non-RCT; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
5. Illescas 2008 - Controlled Non-RCT; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
6. Hanson 2004 - Controlled Non-RCT; higher proportion of prior sexual offences in intervention group compared with control group; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
7. Aytes 2001 - Controlled Non-RCT; significant group differences at baseline in prior incarceration and prior felony conviction; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
8. McGrath 1998 - Controlled Non-RCT; significant group differences at baseline in prior convictions; average time incarcerated and type of sexual offence committed; no blinding; low risk of attrition bias; low risk of selective outcome bias; low risk of other bias
9. McGrath 2003 - Controlled Non-RCT; significant group differences at baseline on prior convictions and time at risk in the community; no blinding; low risk of attrition bias; low risk of selective outcome bias; low risk of other bias
10. 50%<I2<80%
11. Unclear proportion of paraphilia participants
12. The 95% CI considered for imprecision was 0.80 to 1.25.
13. Procter 1996 - Controlled Non-RCT; no blinding; low risk of attrition bias; low risk of selective outcome bias; low risk of other bias
14. Turner 2000 - McGrath 1998 - Controlled Non-RCT; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
15. Olver 2013a - Controlled Non-RCT; low risk of selection bias (confounders properly controlled); no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
16. Craissati 2009 - Controlled Non-RCT; high risk of selection bias; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
17. Lowden 2003 - Controlled Non-RCT; significant group differences at baseline on age, marital status and criminal history; high risk of selection bias; no blinding; low risk of attrition bias; high risk of selective outcome bias; low risk of other bias
18. McGuire 2000 - Controlled Non-RCT; no blinding; unclear risk of attrition bias; low risk of selective outcome bias; low risk of other bias
19. I2>80%

6.6.113 Good Lives Model (GLM) versus Relapse Prevention (RP)

2 controlled non randomised studies (N=1278) met the eligibility criteria for this review. Barnett et al. (2014) and Harkins et al. (2012)compared the Good Lives Model or the revised Better Lives Model with standard relapse prevention in UK sexual offenders against children.

An overview of the trials included can be found in Table 126. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 127. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 126: Study information table for trials included in the analysis of Good Lives Model versus Relapse Prevention for paraphilia

	Good Lives Model versus Relapse Prevention
Total no. of studies (N[1])	2 (1278)
Study ID	(1) Barnet 2014

	Good Lives Model versus Relapse Prevention
	(2) Harkins 2012
Study design	Non-randomised controlled trial
Country	UK (1, 2)
Diagnosis	Paraphilia. 89% offenders against children (2), not otherwise reported (1)
Age in years (mean)	Not reported (1, 2)
Sex (% female)	Not reported (1, 2)
Ethnicity (% white)	96% (1), not reported (2)
Diagnostic status	Unclear (1, 2)
Setting	In the community (1, 2)
Treatment length (weeks)	Not reported (1, 2)
Intervention	Good lives model (GLM) (1, 2)
Delivery method	Face to face (1, 2)
Comparison	Relapse prevention (1, 2)

Table 127: Summary of findings table for Good Lives Model (GLM) versus Relapse Prevention (RP) for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Relapse Prevention	Risk difference with Good Lives Model (GLM)
Cognitive distortions (Children and Sex Questionnaire) (Scale from 0 to 75; lower better)	501 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	-	Mean 13.86 (SD 13.58)	MD 7.15 lower (9.06 lower to 5.25 lower)
Emotional congruence with children (Children and Sex Questionnaire) (Scale from 0 to 75; lower better)	501 (1 observational study)	⊕◯◯◯ VERY LOW [1,3]	-	Mean 18.17 (SD 15.90)	MD 7.72 lower (10.13 lower to 5.3 lower)
Victim empathy distortions (Victim Empathy Distortions scale) (Scale from 0 to 120; lower better)	501 (1 observational study)	⊕◯◯◯ VERY LOW [1,4]	-	Mean 15.69 (SD 16.96)	MD 0.44 higher (2.56 lower to 3.44 higher)
Treatment response for pro-offending attitudes	587 (1 RCT)	⊕⊕◯◯ LOW [5]	RR 0.98 (0.82 to 1.16)	704 per 1,000	14 fewer per 1,000 (127 fewer to 113 more)
Drop-out (any cause)	269 (1 observational study)	⊕◯◯◯ VERY LOW [5,6]	RR 2.09 (0.30 to 14.60)	11 per 1,000	12 more per 1,000 (8 fewer to 149 more)

1. Barnett 2014 - Controlled Non-RCT; no blinding; data on drop-out was not available for some outcomes; low risk of other bias.
2. The MID calculated from SD of control was +/-6.79.
3. The MID calculated from SD of control was +/-7.95.
4. The MID calculated from SD of control was +/-8.48.
5. Harkins 2012 - Controlled Non-RCT; No blinding; data for individual scales were not reported; low other

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Relapse Prevention	Risk difference with Good Lives Model (GLM)

risk of bias.

6. *The 95% CI considered for imprecision was 0.80 to 1.25.*

6.6.114 Reintegration programmes

One RCT (Duwe, 2013)(N=62) and three controlled non randomised studies (Bates et al., 2014; Wilson et al., 2009; Wilson et al., 2007b) (N=350) met the eligibility criteria for this review. All of the studies involved the Circles of Support and Accountability (COSA) intervention. The COSA inner circle consists of the core member (the sex offender) and up to six volunteers from the community. The COSA outer circle consists of, supervision agents, law enforcement personnel and treatment professionals. Volunteers are recruited and trained in preparation for their role, with topics covered such as typology, manipulation, personal boundaries, and managing risk. The goal for each circle is to provide the core member with support during their reintegration into the community.

An overview of the trials included can be found in Table 128. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 129 and Table 130. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 128: Study information table for trials included in the analysis of reintegration programmes versus treatment as usual for paraphilia

	Reintegration programmes versus Treatment as usual
Total no. of studies (N[1])	4 (N=412)
Study ID	(1) Bates 2014 (2) Duwe 2013 (3) Wilson 2007 (4) Wilson 2009
Study design	(2) RCT (1,3,4) Non-randomised controlled trial
Country	Canada (3,4), UK (1), US (2)
Diagnosis	Paraphilic disorder: 86% offenders against children (1), >50% offenders against children (3), sex offenders – but unclear what proportion had paraphilic disorder (2, 4)
Age in years (mean)	47.8 (1), 37.5 (2), 45.5 (3), 42.8 (4)
Sex (% female)	0% (2, 3), not reported (1, 4)
Ethnicity (% white)	16% (2), not reported (1, 3, 4)
Diagnostic status	Unclear (1, 2, 3, 4)
Setting	In the community (1, 2, 3, 4)
Treatment length (weeks)	69 (1), 52 (2), not reported (3, 4)
Intervention	Reintegration programme – circles of support and accountability (1, 2, 3, 4)
Delivery method	Face to face (1, 2, 3, 4)
Comparison	Treatment as usual (not specified: 1, 2, 3, 4)

Table 129: Summary of findings table (RCTs) for reintegration programmes versus treatment as usual for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Reintegration programme
Rearrest at 2-year follow-up (CJS database)	62 (1 RCT)	⊕◯◯◯ VERY LOW [1,2,3]	RR 0.60 (0.36 to 1.00)	645 per 1,000	258 fewer per 1,000 (413 fewer to 0 fewer)
Sex offence rearrest at 2-year follow-up (CJS database)	62 (1 RCT)	⊕◯◯◯ VERY LOW [1,2,3]	RR 0.33 (0.01 to 7.88)	32 per 1,000	22 fewer per 1,000 (32 fewer to 222 more)
Reconviction at 2- to 4-year follow-up (CJS database)	62 (1 RCT)	⊕◯◯◯ VERY LOW [1,2,3]	RR 0.57 (0.28 to 1.16)	452 per 1,000	194 fewer per 1,000 (325 fewer to 72 more)
Resentence at 2-year follow-up (CJS database)	62 (1 RCT)	⊕◯◯◯ VERY LOW [1,2,3]	RR 0.38 (0.11 to 1.28)	258 per 1,000	160 fewer per 1,000 (230 fewer to 72 more)
Any reincarceration at 2-year follow-up (CJS database)	62 (1 RCT)	⊕◯◯◯ VERY LOW [1,2,3]	RR 0.79 (0.50 to 1.25)	613 per 1,000	129 fewer per 1,000 (306 fewer to 153 more)

1. *Duwe 2013 - high risk of selection bias (Prior sex crime conviction was 32% in intervention group compared with 10% in control group); No blinding; low attrition risks; low selective outcome bias; low risk of other bias.*
2. *'Sex offender' - unclear proportion of participants with a paraphilic disorder*
3. *The 95% CI considered for imprecision was 0.80 to 1.25.*

Table 130: Summary of findings table (observational studies) for reintegration programmes versus treatment as usual for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Reintegration programme
Reconviction at 2- to 4-year follow-up (CJS database)	350 (3 observational studies)	⊕◯◯◯ VERY LOW [1,2,3]	RR 0.52 (0.33 to 0.81)	326 per 1,000	156 fewer per 1,000 (218 fewer to 62 fewer)
Reconviction (Any) - UK studies	142 (1 observational study)	⊕◯◯◯ VERY LOW [3,4]	RR 0.50 (0.21 to 1.16)	197 per 1,000	99 fewer per 1,000 (156 fewer to 32 more)
Reconviction (sexual)	350 (3 observational studies)	⊕◯◯◯ VERY LOW [1,2,3]	RR 0.41 (0.18 to 0.94)	120 per 1,000	71 fewer per 1,000 (98 fewer to 7 fewer)
Reconviction (sexual) - UK studies	142 (1 observational study)	⊕◯◯◯ VERY LOW [3,4]	RR 0.80 (0.22 to 2.86)	70 per 1,000	14 fewer per 1,000 (55 fewer to 131 more)
Reconviction (violent)	350 (3	⊕◯◯◯	RR 0.34 (0.19 to	246 per 1,000	162 fewer per 1,000

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Reintegration programme
	observational studies)	VERY LOW [1,2]	0.61)		(199 fewer to 96 fewer)
Reconviction (violent) – UK studies	142 (1 observational study)	⊕◯◯◯ VERY LOW [3,4]	RR 0.07 (0.00 to 1.15)	99 per 1,000	92 fewer per 1,000 (99 fewer to 15 more)

1. *Bates 2014 - Controlled Non-RCT; high risk of selection bias; no blinding; unclear attrition risk of bias; low risk of selective outcome bias; low risk of other bias; Wilson 2007, Wilson 2009 - Controlled Non-RCT; high risk of selection bias; significant differences in baseline risk factors between groups; no blinding; unclear attrition risk of bias; low risk of selective outcome bias; low risk of other bias*
2. *Proportion of participants with paraphilia was unclear (Wilson 2009); only over half (Wilson 2007); majority (86%) of sample (Bates 2014).*
3. *The 95% CI considered for imprecision was 0.80 to 1.25.*
4. *Bates 2014 - Controlled Non-RCT; high risk of selection bias; no blinding; unclear attrition risk of bias; low risk of selective outcome bias; low risk of other bias;*

6.6.115 Therapeutic communities

One controlled non randomised study (N=1217) met the eligibility criteria for this review. Lowden et al. (2003) involved the Sex Offender Treatment and Monitoring Programme (SOTMP) phase 1 and 2, a modified sex offender therapeutic community housing inmates together in a therapeutic milieu where individuals work and live with others who are working on similar treatment issues.

An overview of the trial can be found in Table 131. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 132. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 131: Study information table for trials included in the analysis of therapeutic communities versus no treatment for paraphilia

	Therapeutic communities versus No treatment
Total no. of studies (N[1])	1 (N=1217)
Study ID	Lowden 2003
Study design	Non-randomised controlled trial
Country	US
Diagnosis	Paraphilic disorder: sex offenders – unclear proportion or participants with a paraphilic disorder.
Age in years (mean)	36.5
Sex (% female)	0%
Ethnicity (% white)	51%
Diagnostic status	Unclear
Setting	Prison
Treatment length (weeks)	Not reported
Intervention	Therapeutic communities: sex offender treatment and monitoring programme (SOTMP) phase 1 and 2
Delivery method	Face to face

	Therapeutic communities versus No treatment
Comparison	No treatment

Table 132: Summary of findings table for Therapeutic communities versus no treatment for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment	Risk difference with Therapeutic communities
Rearrest (CJS database)	1217 (1 observational study)	⊕◯◯◯ VERY LOW [1]	RR 0.62 (0.48 to 0.80)	553 per 1,000	210 fewer per 1,000 (287 fewer to 111 fewer)
Sex offence rearrest (CJS database)	1217 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.91 (0.45 to 1.84)	74 per 1,000	7 fewer per 1,000 (41 fewer to 62 more)
Violent rearrest (CJS database)	1217 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.83 (0.58 to 1.19)	262 per 1,000	45 fewer per 1,000 (110 fewer to 50 more)
Incarceration (CJS database)	1217 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.49 (0.28 to 0.84)	208 per 1,000	106 fewer per 1,000 (150 fewer to 33 fewer)
Incarceration for sexual offence (CJS database)	1217 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 1.32 (0.57 to 3.04)	38 per 1,000	12 more per 1,000 (16 fewer to 78 more)
Incarceration for violent offence (CJS database)	1217 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.37 (0.12 to 1.17)	67 per 1,000	42 fewer per 1,000 (59 fewer to 11 more)
Revocation (CJS database)	1425 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.33 (0.21 to 0.50)	477 per 1,000	320 fewer per 1,000 (377 fewer to 239 fewer)

1. *Lowden 2003 - Controlled Non-RCT; significant group differences at baseline on age, marital status and criminal history; high risk of selection bias; no blinding; low risk of attrition bias; high risk of selective outcome bias; low risk of other bias*
2. *The 95% C.I. considered for imprecision was 0.80 to 1.25*

6.6.1.46 Cognitive behavioural therapy

One controlled non randomised study (N=61) met the eligibility criteria for this review. Marshall et al. (1991) involved a treatment programme that conceptualized exhibitionism in cognitive and social terms, rather than simply sexual motivation. Treatment was aimed at teaching skills to deal with all sources of stress, and intervention content included: assertiveness training; stress management; cognitive restructuring; training in relationship skills.

An overview of the trial can be found in Table 133. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 134. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 133: Study information table for trials included in the analysis of cognitive behavioural therapy versus treatment as usual for paraphilia

	Cognitive behavioural therapy versus treatment as usual
Total no. of studies (N[1])	1 (61)
Study ID	Marshall (1991)
Study design	Non-randomised controlled trial
Country	Canada
Diagnosis	Paraphilic disorder: exhibitionists
Age in years (mean)	29
Sex (% female)	0%
Ethnicity (% white)	Not reported
Diagnostic status	Unclear
Setting	In the community
Treatment length (weeks)	Not reported
Intervention	Cognitive behavioural therapy: modified treatment programme for exhibitionists
Delivery method	Face to face
Comparison	Treatment as usual (not specified)

Table 134: Summary of findings table for cognitive behavioural therapy (CBT) versus treatment as usual for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with CBT
Sexual reconviction (CJS database; non-randomised controlled trials; longest follow-up available) - 4-year follow-up (exhibitionists)	38 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.41 (0.16 to 1.05)	571 per 1,000	337 fewer per 1,000 (480 fewer to 29 more)

1. *Marshall 1988a/b/1991 - Controlled Non-RCT with 4 and 9-year follow-up; No baseline risk differences; No blinding; unclear attrition risk of bias; low risk of selective outcome bias; low risk of other bias*
2. *The 95% CI considered for imprecision was 0.80 to 1.25.*

6.6.177 Behavioural therapy

Two randomised trials (McConaghy et al., 1985; McConaghy et al., 1988) and two controlled non randomised studies (Marshall & Barbaree, 1988b; Marshall et al., 1991) (N=187) met the eligibility criteria for this review.

McConaghy et al. (1985) compared imaginal desensitization to covert sensitization for varied paraphilic disorders. McConaghy et al. (1988) examined the addition of imaginal desensitization to MPA for varied paraphilic disorders. Marshall & Barbaree (1988b) and Marshall et al., (1991) compared behavioural treatment programs to treatment as usual in paedophiles and exhibitionists respectively

An overview of the trials can be found in Table 135. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 136, Table 137 and Table 138. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 135: Study information table for trials included in the analysis of behavioural therapy for paraphilia

	Behavioural therapy versus treatment as usual	Behavioural therapy versus other active intervention
Total no. of studies (N[1])	2 (187)	2 (40)
Study ID	(1) Marshall 1988a, 1988b (2) Marshall 1991	(3) McConaghy 1985 (4) McConaghy 1988
Study design	Non-randomised controlled trial	RCT
Country	Canada (1,2)	Australia (3, 4)
Diagnosis	Paraphilic disorder: men who had sexually molested children (1), exhibitionists (2)	Paraphilic disorder: men seeking treatment for anomalous sexual urges or behaviours (3,4)
Age in years (mean)	34.6 (1), 29.0 (2)	36 (3), 30 (4)
Sex (% female)	0% (1, 2)	0% (3, 4)
Ethnicity (% white)	Not reported (1, 2)	Not reported (3, 4)
Diagnostic status	Unclear (1, 2)	Unclear (3), clinical diagnosis (4)
Setting	In the community (1, 2)	Inpatient (3), in the community and inpatient (4)
Treatment length (weeks)	Not reported (1), 26 (2)	1 (3), 26 (4)
Intervention	Behavioural treatment programme for child molesters (1), behavioural treatment programme for exhibitionists (2)	Imaginal desensitization (3), imaginal desensitization plus MPA (4)
Delivery method	Face to face (1, 2)	Face to face (3, 4)
Comparison	Treatment as usual (not specified: 1, 2)	Covert sensitization (3), MPA (4)

Notes. N= total number of participants; NR=Not reported; MPA = medroxyprogesterone
[1] *Number randomised*

Table 136: Summary of findings table for behavioural therapies versus treatment as usual for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with behavioural therapies versus TAU
Sexual reconviction (CJS database) – 4 year follow-up (sex offenders against children)	44 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.42 (0.19 to 0.91)	600 per 1,000	348 fewer per 1,000 (486 fewer to 54 fewer)
Sexual reconviction (CJS database) – 9 year follow-up (exhibitionists)	44 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 0.68 (0.36 to 1.29)	571 per 1,000	183 fewer per 1,000 (366 fewer to 166 more)

1. *Marshall 1988a/b/1991 - Controlled Non-RCT with 4 and 9-year follow-up; No baseline risk differences; No blinding; unclear attrition risk of bias; low risk of selective outcome bias; low risk of other bias*

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with behavioural therapies versus TAU
2. The 95% CI considered for imprecision was 0.80 to 1.25.					

Table 137: Summary of findings table behavioural intervention versus other active intervention

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with active intervention	Risk difference with behavioural therapy versus active intervention
Reduction in anomalous behaviours (26 weeks follow-up)	20 (1 RCT)	⨁◯◯◯ VERY LOW [1,2,3]	RR 1.12 (0.78 to 1.63)	800 per 1,000	96 more per 1,000 (176 fewer to 504 more)
Reduction in anomalous desires (26 weeks follow-up)	20 (1 RCT)	⨁◯◯◯ VERY LOW [1,2,3]	RR 1.67 (0.54 to 5.17)	300 per 1,000	201 more per 1,000 (138 fewer to 1,251 more)
1. McConaghy 1988 - unclear risk of selection bias, no blinding, low risk of attrition bias, high risk of selective outcome bias, low risk of other bias.					
2. Unclear what percentage are currently in contact with the criminal justice system					
3. The 95% CI considered for imprecision was 0.80 to 1.25.					

Table 138: Summary of findings table for imaginal desensitization versus covert sensitization for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with covert sensitization only	Risk difference with Imaginal desensitization
Reduction in anomalous behaviours	20 (1 RCT)	⨁◯◯◯ VERY LOW [1,2,3]	RR 1.75 (0.74 to 4.14)	400 per 1,000	300 more per 1,000 (104 fewer to 1,256 more)
Reduction in anomalous desires	20 (1 RCT)	⨁◯◯◯ VERY LOW [1,2,3]	RR 0.60 (0.19 to 1.86)	500 per 1,000	200 fewer per 1,000 (405 fewer to 430 more)
1. McConaghy 1985 - unclear selection bias, no blinding, high risk of attrition bias, high risk of selective outcome bias, low other risk of bias,					
2. 13/20 had previously received convictions but unclear what percentage of the sample were currently in contact with the criminal justice system. Also 5 individuals requested treatment due to being homosexual, which would no longer be considered a paraphilia.					
3. The 95% CI considered for imprecision was 0.80 to 1.25.					

6.6.15.8 Aversive conditioning training and milieu therapy

One controlled non randomised study (N=197) met the eligibility criteria for this review. Hanson et al. (1993) examined a specialized treatment programme provided in a separate minimum security setting. The programme aimed to increase the social competence of offenders through individual and group counselling and by creating a therapeutic milieu that encouraged the men to recognise and correct social and sexual adjustment problems. The offenders also received aversive conditioning training, involving pairing shocks to stimulus sets tailored for each participant on the basis of their offence history.

An overview of the trial can be found in Table 139. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 140. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 139: Study information table for trials included in the analysis of aversive conditioning training and milieu therapy for paraphilia

	Aversive conditioning training and milieu therapy vs treatment as usual
Total no. of studies (N[1])	1 (197)
Study ID	Hanson 1993
Study design	Non-randomised controlled trial
Country	Canada
Diagnosis	Paraphilic disorder: male child molesters released from maximum security prison
Age in years (mean)	33.1
Sex (% female)	0%
Ethnicity (% white)	Not reported
Diagnostic status	Unclear
Setting	Prison
Treatment length (weeks)	22
Intervention	Aversive conditioning and milieu therapy
Delivery method	Face to face
Comparison	Treatment as usual (not specified)
Notes. N= total number of participants; NR=Not reported [1] Number randomised	

Table 140: Summary of findings table for aversive conditioning and milieu therapy versus treatment as usual for paraphilic disorders

Outcomes	**№ of participants (studies) Follow-up**	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with treatment as usual	**Risk difference with Aversive conditioning training and milieu therapy**
Sexual or violent reconvictions at 21-year follow-up (CJS database)	197 (1 observational study)	⊕◯◯◯ VERY LOW [1,2]	RR 1.15 (0.82 to 1.61)	385 per 1,000	58 more per 1,000 (69 fewer to 235 more)
1. Hanson 1993 - Controlled Non-RCT; significant baseline risk differences (+); no blinding; unclear attrition risk of bias; low risk of selective outcome bias; low risk of other bias. *2. The 95% CI considered for imprecision was 0.80 to 1.25.*					

6.6.109 Individual and Group Psychotherapy

Two controlled non randomised studies (Craissati 2009, Peters 1968; N=440) met the eligibility criteria for this review. Peters et al. (1968) involved a group psychotherapy programme for sex offenders. Craissati et al. (2009) involved individual supportive psychotherapy for sex offenders deemed inappropriate for more structured offence-focused treatment due to disruptive or antagonistic behaviour or denial of the offence

An overview of the trials can be found in Table 141. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 142. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 141: Study information table for trials included in the analysis of psychotherapy versus no treatment or treatment as usual for paraphilia

	Psychotherapy vs No treatment/Treatment as usual
Total no. of studies (N[1])	2 (335)
Study ID	(1) Craissati 2009 (2) Peters 1968
Study design	Non-randomised controlled trial
Country	UK (1), US (2)
Diagnosis	Paraphilic disorder: sex offenders 73% against children (1), sex offenders unclear proportion with paraphilic disorder (2)
Age in years (mean)	Not reported (1,2)
Sex (% female)	0% (1, 2)
Ethnicity (% white)	Not reported (1), 50% (2)
Diagnostic status	Unclear (1, 2)
Setting	In the community (1,2)
Treatment length (weeks)	Not reported (1), 26 (2)
Intervention	Individual supportive psychotherapy (1), group psychotherapy programme for sex offenders (2)
Delivery method	Face to face (1,2)
Comparison	No treatment (1), treatment as usual (not specified: 2)
Notes. N= total number of participants; NR=Not reported [1] Number randomised	

Table 142: Summary of findings table for psychotherapy versus no treatment or treatment as usual for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment or treatment as usual	Risk difference with Psychotherapy
Rearrest (2-year follow-up)	167 (1 observational study)	⨁◯◯◯ VERY LOW [1,2]	RR 0.12 (0.04 to 0.40)	267 per 1,000	235 fewer per 1,000 (256 fewer to 160 fewer)
Sex offence rearrest (2-year follow-up)	167 (1 observational study)	⨁◯◯◯ VERY LOW [1,2,3]	RR 0.14 (0.02 to 1.10)	80 per 1,000	69 fewer per 1,000 (78 fewer to 8 more)
Sexual reconviction (CJS database)	168 (1 observational study)	⨁◯◯◯ VERY LOW [3,4]	RR 1.85 (0.76 to 4.54)	117 per 1,000	100 more per 1,000 (28 fewer to 415 more)

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment or treatment as usual	Risk difference with Psychotherapy
Violent reconviction (CJS database)	168 (1 observational study)	⨁◯◯◯ VERY LOW [3,4]	RR 0.79 (0.26 to 2.41)	166 per 1,000	35 fewer per 1,000 (122 fewer to 233 more)
Breaches of the Sex Offender Register (CJS database)	168 (1 observational study)	⨁◯◯◯ VERY LOW [3,4]	RR 1.44 (0.77 to 2.70)	241 per 1,000	106 more per 1,000 (56 fewer to 410 more)

1. *Peters 1968 - Controlled Non-RCT; group differences at baseline; no blinding; unclear attrition risk of bias; low risk of selective outcome bias; low risk of other bias.*
2. *'Sex offender' - unclear proportion of participants with a paraphilic disorder; also an unknown proportion of participants in the intervention group had treatment delivered in a psychiatric inpatient unit*
3. *The 95% CI considered for imprecision was 0.80 to 1.25.*
4. *Craissati 2009 - Controlled Non-RCT; there might have selection bias issues such as unequal baseline risks between 2 groups and the individual psychoeducation group was also offered to those who had already attempted group work; No blinding; only participants with available follow-up data were included; low risk of selective outcome bias; low risk of other bias*

6.6.1.10 Polygraph testing

One controlled non randomised study (N=208) met the eligibility criteria for this review. McGrath et al. (2007) examined the effectiveness of periodic polygraph compliance exams as a condition of probation or parole in a group of primarily sex offenders against children.

An overview of the trial can be found in Table 143. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 144. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 143: Study information table for trials included in the analysis of polygraph testing versus treatment as usual for paraphilia

	Polygraph testing vs Treatment as usual
Total no. of studies (N[1])	1 (208)
Study ID	McGrath 2007
Study design	Non-randomised controlled trial
Country	US
Diagnosis	Paraphilic disorder: sex offenders 61% against children
Age in years (mean)	35.6
Sex (% female)	0%
Ethnicity (% white)	98%
Diagnostic status	Unclear
Setting	In the community
Treatment length (weeks)	212
Intervention	Periodic polygraph compliance exams
Delivery method	Face to face
Comparison	Treatment as usual (the control group did not undergo any polygraph

	Polygraph testing vs Treatment as usual
	exams)
Notes. N= total number of participants; NR=Not reported [1] Number randomised	

Table 144: Summary of findings table for polygraph testing versus treatment as usual for paraphilic disorders

Outcomes	№ of participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with polygraph testing
Reconviction (CJS database) - 5-year follow-up	208 (1 observational study)	⨁◯◯◯ VERY LOW [1,2]	RR 1.14 (0.80 to 1.63)	346 per 1,000	48 more per 1,000 (69 fewer to 218 more)
Sexual reconviction (CJS database) - 5-year follow-up	208 (1 observational study)	⨁◯◯◯ VERY LOW [1,2]	RR 0.86 (0.30 to 2.46)	67 per 1,000	9 fewer per 1,000 (47 fewer to 98 more)
Violent reconviction (CJS database) - 5-year follow-up	208 (1 observational study)	⨁◯◯◯ VERY LOW [1,2]	RR 0.25 (0.07 to 0.86)	115 per 1,000	87 fewer per 1,000 (107 fewer to 16 fewer)
Incarceration (CJS database) - 5-year follow-up	208 (1 observational study)	⨁◯◯◯ VERY LOW [1,2]	RR 1.23 (0.89 to 1.68)	385 per 1,000	88 more per 1,000 (42 fewer to 262 more)
Violation of supervision conditions (CJS database) - 5-year follow-up	208 (1 observational study)	⨁◯◯◯ ⨁◯◯◯ VERY LOW [1,2]	RR 1.15 (0.87 to 1.52)	452 per 1,000	68 more per 1,000 (59 fewer to 235 more)

1. McGrath 2007 - Controlled Non-RCT; baseline characteristics were similar between the groups; no blinding; low risk of detection bias; low attrition bias; low selective outcome bias; low risk of other bias

2. The 95% CI considered for imprecision was 0.80 to 1.25.

6.642 Economic evidence

The systematic search of the literature identified 1 Australian study in two publications that assessed the cost-benefit of psychological therapy for adults with a paraphilic disorder who are in contact with the criminal justice system Donato 2001(Donato & Shanahan, 2001); Shanahan 2001(Shanahan & Donato, 2001).

No studies assessing the cost effectiveness of pharmacological interventions for adults with a paraphilic disorder who are in contact with the criminal justice system were identified by the systematic search of the economic literature undertaken for this guideline.

Details on the methods used for the systematic review of the economic literature are described in Chapter 3; full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix S. Completed methodology checklists of the studies are provided in Appendix R. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix T.

Donato & Shanahan (2001) conducted a cost-benefit analysis of intensive prison-based paedophile treatment (CBT) in Australia. This was a modelling study with clinical effectiveness data based on published sources and authors' assumptions. The time horizon of the economic analysis was lifetime, and its perspective was public sector (healthcare, social care, and out of pocket expenses). Cost elements comprised CBT provision, the judiciary (court), police, family services (counselling, mediation, child contact services, domestic violence prevention programmes), child-focused health services, medicines, medical services (psychiatrists, general practitioners), out-of-pocket expenses by victims and their families, incarceration and other victim and offender related expenses. Cost data were obtained from various international, federal and state level sources, and authors' assumptions. The analysis utilised the net benefit (NB) framework. The NB was defined as the sum of tangible benefits (resource savings) and intangible benefits (value of health consequences such as avoiding pain and suffering) less the programme provision costs. Intangible benefits were valued using both revealed preferences and contingent valuation methods. When using revealed preferences approach intangible benefits were approximated using a US study that reported the amounts compensated in child sex abuse cases. When using the contingent valuation method intangible benefits were approximated by linking road traffic injuries and associated costs with injuries associated with sexual abuse.

The analysis demonstrated that the total programme provision cost was $10,000 per treated prisoner, the tangible benefits of preventing re-offense were approximately $157,290, and the intangible benefits of preventing re-offense varied from $0 to $198,900 depending on the monetary valuation placed upon intangible benefits (in 1998 AUS dollars). Based on the above the economic benefits associated with intensive prison-based CBT ranged from an expected net loss of $6,850 to an expected NB of $39,870 per treated prisoner (depending on the monetary valuation placed upon intangible benefits and the efficacy of the treatment programme). For example, when intangibles were valued at zero and prison-based CBT was assumed to reduce recidivism by 2%, the intervention resulted in a net loss of $6,850 per treated prisoner. However, when intangibles were valued at ten times the value of tangible benefits, and intervention was assumed to reduce recidivism rate by 14%, the economic benefits were expected to reach $39,870 per treated prisoner. The deterministic sensitivity analysis indicated that if there were two victims per re-offender the economic benefits of a treatment programme would range from an expected net loss of $6,850 to an expected net benefit of $76,710 per treated prisoner (again depending on the monetary valuation placed upon intangible benefits and the efficacy of the treatment programme). Based on these results, the authors concluded that 'based on a reasonable set of parameter estimates, prison-based CBT for paedophiles is likely to be of net benefit to society' (Donato & Shanahan, 2001).

This study is only partially applicable to the NICE decision-making context. It was conducted in Australia and the measure of outcome was not expressed in QALYs. The study was judged by the GC to have potentially serious limitations. Clinical effectiveness (recidivism rate) was based on authors' assumptions. The valuation of intangible costs was approximated using compensation rates for road traffic accident victims when using revealed preferences approach the values. Resource use and unit cost data were based on a mixture of national and local sources and as necessary were supplemented with information from published studies.

6.6.3 Clinical evidence statements

6.6.3.1 Pharmacological interventions

Very low quality evidence from two randomised controlled trials (N=66) indicated uncertainty about the benefit of adding MPA to psychosocial interventions for paraphilia in terms of anomalous desires or behaviour.

Low quality evidence from one randomised controlled trial (N = 32) indicated that adding MPA to a psychosocial intervention for paraphilia increased the risk of study dropout by a clinically important amount compared to the psychological intervention alone.

Very low quality evidence from one randomised controlled trial (N=20) indicated uncertainty about whether MPA alone was more effective than imaginal desensitization alone in terms of anomalous desires or behaviour.

6.6.372 Psychoeducational interventions

Moderate quality evidence from one randomised controlled trial (N=60) indicated that a psychoeducational CBT intervention reduced cognitive distortions by a clinically important amount when compared to no treatment. This trial provided low quality evidence of uncertainty about the effects of psychoeducation on acceptance of accountability, sexual anxiety and anxiety.

The only evidence about reconviction from randomised studies was from a single trial in the inpatient setting (N =480) which provided low quality evidence of uncertainty about the benefit of psychoeducation in terms of sexual or violent reconviction rates.

Very low quality evidence from nine non-randomised controlled trials (N=2796) indicated that psychoeducational interventions led to a clinically important reduction in reconviction rates when compared to treatment as usual, no treatment or waitlist control. This was also the case when restricting the analysis to UK studies (1 study; N=338).

There was very low quality evidence from 11 non-randomised controlled trials (N=5261) that psychoeducational interventions were associated with a clinically important reduction in reconviction rates for sexual offenses when compared to treatment as usual, no treatment or waitlist control. There was very low quality evidence of uncertainty about the effect of psychoeducation on reconviction for sexual offenses when restricting the analysis to UK studies (three studies; N=2885).

There was very low quality evidence from 6 non-randomised controlled trials (N=2181) that psychoeducational interventions were associated with a clinically important reduction in reconviction rates for violent offenses when compared to treatment as usual, no treatment or waitlist control. There was very low quality evidence of uncertainty about the effect of psychoeducation on reconviction for violent offenses when restricting the analysis to UK studies (1 study; N=240).

There was very low quality evidence from 6 non-randomised controlled trials (N=2181) of uncertainty about the effect of psychoeducational interventions on revocation rates when compared to treatment as usual, no treatment or waitlist control. There was very low quality evidence of a clinically important reduction in revocation rates with psychoeducation when restricting the analysis to UK studies (1 study; N=240).

6.6.373 Good Lives Model (GLM) versus Relapse Prevention (RP)

Very low quality evidence from one observational study (N=501) suggested that the Good Lives Model reduced cognitive distortions and emotional congruence with children by a clinically important amount when compared with relapse prevention.

Very low quality evidence from one observational study (N=501) suggested no clinically important difference in the effectiveness of the Good Lives Model and relapse prevention in terms of victim empathy distortions.

Very low quality evidence from one observational study (N=2698) suggested uncertainty about the relative treatment dropout rates associated with the Good Lives Model and relapse prevention.

Low quality evidence from one RCT (N=587) indicated no clinically important difference in the effectiveness of Good Lives Model and relapse prevention for reducing pro-offending attitudes.

There was no evidence about the relative effectiveness of the Good Lives Model and relapse prevention in terms of offending or reoffending.

6.6.364 Reintegration programmes

One randomised trial (N = 62) provided very low quality evidence that a support group, Circles of Support and Accountability, reduced rates of re-arrest at two years of follow-up by a clinically important amount compared to treatment as usual. From this trial there was very low quality evidence of uncertainty about the relative effectiveness of the support group compared to treatment as usual in terms of: sex offence re-arrest, reconviction, resentence and reincarceration at two years follow up.

Very low quality evidence from three non-randomised controlled trials (N=350) indicated reintegration programmes were associated with clinically important reductions in reconviction rates (including for sexual offenses) when compared to treatment as usual. Restricting the analysis to UK only studies there was uncertainty whether reintegration programmes were more effective than treatment as usual.

6.6.385 Therapeutic communities

Very low quality evidence from one non-randomised controlled trials (N=1217) indicated therapeutic communities reduced rates of re-arrest, incarceration and revocation by a clinically important amount compared with no treatment. There was however uncertainty about the effectiveness of the therapeutic community intervention in terms of re-arrest or incarceration when looking at specific sexual or violent offenses. .

6.6.406 Cognitive behavioural therapy

Very low quality evidence from one observational study (N=38) indicated uncertainty about whether CBT reduces the rate of sexual reconviction when compared with treatment as usual.

6.6.427 Behavioural therapy

Very low quality evidence from a small non-randomised study (N=44) suggested a behavioural treatment programme had a clinically important effect on sexual reconviction rates at 4 years of follow up in sex offenders against children but there was uncertainty about its effectiveness for sexual reconviction rates at 9 years in exhibitionists.

Very low quality evidence from a randomised trial (N=20) indicated uncertainty about the relative effectiveness of imaginal desensitization plus MPA versus MPA alone in terms of anomalous desires and behaviours.

Very low quality evidence from a randomised trial (N=20) indicates uncertainty about the relative effectiveness of imaginal desensitization versus covert sensitization alone in terms of anomalous desires and behaviours.

6.6.448 Aversive conditioning training and milieu therapy

Very low quality evidence from one observational study (N=197) indicated uncertainty about whether aversive conditioning training and milieu therapy is more or less effective than treatment as usual in terms of sexual or violent reconvictions.

6.6.3.9 Individual and Group Psychotherapy

Low quality evidence from one non-randomised controlled trials (N=167) indicated a clinically important reduction in rearrest rates following psychotherapy for paraphilic disorders when compared to treatment as usual.

Low quality evidence from two non-randomised controlled trials (N=335) indicated uncertainty about the effectiveness of psychotherapy for paraphilic disorders when compared to no treatment or treatment as usual in terms of sex-offence rearrest or reconviction, violent reconviction and breaches of the sex offender register.

6.6.3.10 Polygraph testing

Very low quality evidence from a non-randomised controlled trials (N=208) indicated uncertainty about the effectiveness of periodic polygraph compliance exams when compared to treatment as usual in terms of sexual reconviction, incarceration or violation of supervision conditions. Violent reconviction, however, was reduced by a clinically important amount in the polygraph testing group.

6.6.4 Economic evidence statements

No evidence on the cost effectiveness of pharmacological interventions for adults with a paraphilic disorder who are in contact with the criminal justice system is available.

There was evidence from 1 Australian study on the cost-benefit of psychosocial intervention for adults with a paraphilic disorder who are in contact with the criminal justice system. The analysis was based on modelling suggesting that prison-based, cognitive behavioural therapy treatment programme for paedophiles may be of net benefit to society. The economic benefits associated with intensive prison-based CBT ranged from an expected net loss of $6,850 to an expected NB of $39,870 per treated prisoner (depending on the monetary valuation placed upon intangible benefits and the efficacy of the treatment programme). This evidence is partially applicable to the NICE decision-making context since it was Australian study, and is characterised by potentially serious limitations, including clinical effectiveness (recidivism rate) being based on authors' assumptions; the valuation of intangible costs (pain and suffering) being undertaken using both contingent valuation and revealed preferences methods. However, when using revealed preferences approach the values were approximated using compensation rates for road traffic accident victims. Resource use and unit cost data were based on a mixture of national and local sources supplemented with information from the published studies. The GC could not draw any conclusions based on this evidence.

6.7 Recommendations and link to evidence

Recommendations	43. Consider psychological interventions for paraphilias only when delivered as part of a research programme.
Relative values of different outcomes	The guideline committee considered offending and reoffending (i.e. paraphilic activity) to be the critical outcomes for this question. Some studies reported cognitive distortions (measure of attitudes/beliefs to paraphilic activity) but the GC did not consider this to a good surrogate offending behaviour. There was no evidence for service utilisation, adaptive functioning or rates of self-injury.
Trade-off between clinical benefits and	Psychological, including psychoeducational interventions, and pharmacological interventions for paraphilia's aim to reduce

harms	the rate of sexual offending with potential benefits for the victims of such offences and their families and communities which may be substantial and long-lasting given that such offences in particularly against children, may be associated with lifelong harm. They also aim to reduce the distress experienced by the offender and improve their mental health and attitudes toward sexual offending. A large number of psychological interventions were reviewed although the majority were not randomised. The clearest indication of a benefit came from the studies of psychoeducational interventions which were of low quality and the estimate of the outcomes (typically reduction in offending) was uncertain. Evidence from a range of observational studies or small randomised trial for a range of other psychological interventions including relapse prevention, reintegration programmes, therapeutic communities, cognitive and behavioural therapies, aversion therapy or individual and group psychotherapy produce low quality evidence which the GC did not think could support any recommendation. Polygraph testing in a small single low quality study suggested some benefit on violent reconviction, but the GC did not this think that they could support a recommendation for such testing. Very low quality evidence from three randomised controlled trials did not show clear benefit for medroxyprogesterone acetate alone or in combination with a psychological interventions or alone. Medroxyprogesterone acetate was also associated with high attrition from treatment and the GC did not think that they could make a treatment recommendation for its use. The primary harms are associated with the use of ant-adrenergic drugs which are associated with significant side effects, including breast development in men. These side effects are also associated with a high drop-out from treatment and poor compliance with treatment regimens. In addition, many programmes, including many psychological interventions, are delivered in custodial environments where attitude change may be a proxy indicator of benefit but which in the absence of opportunity to asses a reduction in offending behaviour may lead to under-estimation of the risk of re-offending after completion of treatment.
Trade-off between net health benefits and resource use	The costs of treatment are limited as they consist of time limited psychological interventions which can be delivered in community or residential settings including prisons. However, the majority of psychological or pharmacological (which require the prescription and monitoring of patent drugs) do not need to be delivered in residential settings. Long-term monitoring of pharmacological interventions or follow up of psychological interventions will often take place in in a context where long-term monitoring of the risk of re-offending is undertaken. Effective treatment will not only reduce the use of resources associated with the care of the individual with paraphilia but also likely contribute to a use of health care resource by individuals who would have become victims of sexual offences if the problem was not successfully treated. One Australian cost-effectiveness study of limited applicability suggested that psychological interventions in a prison setting may be cost effective. The GC thought that this study had significant

	limitations including the parameters included in the economic model, the populations included in the study and the assumptions made by the authors of the study about the effectiveness of the interventions.
Quality of evidence	The quality of the evidence from a small number of RCTs and a larger number of non-randomised studies ranged from moderate to very low. The randomised trials typically had small sample sizes leading to wide confidence intervals for effect estimates and although generally showing evidence of benefit on re-offending and attitudes across a range of interventions (the majority of which consisted of specially developed CBT based psycho-educational interventions), there was uncertainty about the harms and benefits of the interventions. Only one of the randomised trials used adequate blinding. Many of the studies which included sex offenders did not report the proportion and type of paraphilic disorders (the particular focus of this review). The guideline committee were concerned that populations in these studies might not be applicable to those with paraphilic disorders seen in the UK criminal justice system (the UK system has a high proportion of paedophiles and the proportion of those in non-UK studies was less than in the UK) and for this reason paid close attention to the sub-group analysis of UK studies. Although the evidence as a whole suggested that psychological interventions, including psychoeducational interventions, may be effective in reducing re-conviction rates, the evidence was not as clear as non-UK based studies to support their effectiveness in reducing re-conviction for sexual offences in the UK and in particular in the populations likely to managed in the UK CJS.
Other considerations	The National Offender Management Service (NOMS) Psychology and Interventions Teams provide an accredited Sex Offender Treatment Programme for those in custody in England and Wales. The GC thought that because this programme is currently the standard intervention used across the CJS, is well delivered and has good outcome monitoring in place, evidence about its effectiveness was essential to inform any recommendations in this guideline about treatment of paraphilic disorders within the UK criminal justice system. Although an expert witness from NOMS provided testimony on co-commissioning mental health services for offenders, NOMS did not agree to a GC request to release relevant reports or data from the outcome of their Sex Offender Treatment Programme. Given the absence of evidence from the NOMS programmes and the uncertainty about the evidence reviewed, in particular the UK evidence, the guideline committee decided to made no treatment recommendations about interventions for people with paraphilic disorders. Instead they recommended that psychological interventions for paraphilias only be delivered as part of a research programme. Given the high drop-rate and poor compliance identified in the review of pharmacological interventions and lack of direct evidence on sexual offending the GC did not make a recommendation about drug treatments. Given the importance of this topic the GC also made a further research recommendation to determine if either pharmacological or psychological intervention s are effective in reducing re-offending in paraphilic disorders. This should address the use of these interventions in a range of settings in the UK criminal

	justice system.

6.7.11 Research recommendations

5. What is the clinical, cost-effectiveness and safety of psychological and pharmacological interventions both in and out of the prison among people with paraphilic disorders?

The limited evidence for pharmacological interventions (for example, medroxyprogesterone acetate) provides no clear evidence of benefit in people with paraphilias. A randomised trial with adequate sample size is required to examine the effectiveness of medroxyprogesterone acetate in these populations.

There is insufficient evidence on the use of psychological interventions for people with paraphilias in the criminal justice system. Individual patient data analysis of paedophiles who have been treated should be conducted to inform treatment and future research. Psychological interventions paraphilias (such as sex offender treatment programme) should be tested in large randomised controlled trials in criminal justice populations. This research could have a significant impact upon updates of this guideline..

Important outcomes could include:

- Offending and re-offending rates
- Effect on mental health problems
- Cost-effectiveness
- Health-related quality of life

While designing the trials, consideration should be given to timing, intensity and duration of interventions in the context of the criminal justice system.

6.8 Review question: For adults with acquired cognitive impairment who are in contact with the criminal justice system, what are the benefits and harms of psychological, social or pharmacological interventions aimed at rehabilitation?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 145. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 145: Clinical review protocol summary for the review of psychological, social or pharmacological interventions aimed at rehabilitation of adults with acquired cognitive impairment (ACI) in contact with the criminal justice system

Component	Description
Population	Adults with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Intervention(s)	• Psychological and social interventions • Pharmacological interventions • Combined psychological or social and pharmacological

Component	Description
	interventions • Support and education interventions
Comparison	• Treatment as usual • No treatment • Waitlist control • Placebo • Any alternative management strategy
Outcomes	• Critical – Improvement in cognitive functioning; Improvement in adaptive functioning; Offending and re-offending outcomes • Important – mental health outcomes (symptomatology; self-harm and suicide)
Study design	Systematic reviews and RCTs

6.8.11 Clinical evidence

No directly relevant RCTs or systematic reviews were found to address this review question and when agreeing the review protocol GC decided it would be inappropriate to descend the evidence hierarchy as they were aware, on the basis of their existing knowledge of the literature, that it was unlikely to be fruitful and was therefore not considered a good use of time and resource.

In the absence of direct evidence, indirect evidence from populations outside of the criminal justice system was considered. The GC decided that extrapolation from non-criminal justice populations was potentially useful because acquired cognitive impairment is a common sequela of acquired brain injury regardless of population.

Seven systematic reviews of various interventions designed to remediate difficulties associated with ACI were identified. These are summarised narratively below. Summary study characteristics can be found within Table 146. Full details of these reviews can be found in Appendix N. As this was a narrative overview of these systematic reviews GRADE analysis was not conducted, because the evidence was not yet summarised by comparisons and outcomes at this stage. After considering the overview it was decided that further analysis according to study design (RCT versus observational study), intervention or outcome would be unlikely to alter the committee's conclusions, given that in general no clinically significant improvements were observed. The evidence was considered low quality because it was not from non-criminal justice populations.

The systematic reviews identified spanned a range of different disorders associated with acquired cognitive impairment, some progressive and some static; mild cognitive impairment as a precursor to dementia (Cooper 2013), epilepsy (Farina 2015), various neurological conditions (Krasny-Pacini 2013), multiple sclerosis (O'Brien 2008), stroke (Whyte 2011) and stroke as well as other acquired, non-progressive brain injuries (Chung 2013) and (Coleman 2015). The terms ABI (acquired brain injury) and TBI (traumatic brain injury) are used throughout this section as they have been by review authors. ABI is used to describe non-degenerative acquired brain injuries including stroke and impact-related injuries. TBI is specifically used to describe brain injury resulting from head trauma, such as that acquired in a car crash or whilst playing sports.

Whyte 2011 and Chung 2013 both conducted reviews of rehabilitative interventions for stroke and other acquired, non-progressive brain injuries. Whyte 2011 was a narrative review identifying two broad targets for intervention; adaptation and remediation. They note that the difficulties associated with ABI can make engagement with therapeutic interventions more challenging, and that there is little evidence for remediation of deficits at present, but that theoretically high frequency repetition (i.e. intense neuro-rehabilitation) may be beneficial.

The Chung 2013 (N=770) paper was a Cochrane review focusing on cognitive rehabilitation for executive dysfunction, which is commonly impaired in people with ACI. They included randomised studies looking at restorative or adaptive interventions and compensatory strategies for TBI, stroke or 'other acquired brain injury'. All included studies compared the intervention of interest with no treatment, placebo or another active intervention. 3 included studies compared cognitive rehabilitation with sensorimotor therapy, 6 compared cognitive rehabilitation with no treatment or placebo, 10 compared different rehabilitative approaches. Only 2 studies (N=82) reported data on a primary outcome (global executive function measured with the Behavioural Assessment of Dysexecutive Syndrome [BADS], Chung 2007 and Spike 2010), demonstrating no clinically significant effect. Krasny-Pacini 2013 (N=n/r) was a narrative review of a mixture of RCTs and case-reports focussed on a specific rehabilitative technique called 'Goal Management Training' (GMT). 4 'proof-of-principle' studies and 8 experimental studies concerned with implementing the technique in practice were included. The authors concluded that GMT may have some benefits in terms of adaptive functioning, but that if used it would be more efficacious as part of a comprehensive rehabilitative package.

Coleman 2015 (N=388) also conducted a systematic review of assessment (8 studies) and intervention (2 studies) delivered via tele-practice for acquired, non-degenerative brain injuries including TBI and stroke. The 2 studies investigating rehabilitative interventions both compared different forms of problem solving training, 1 with the same intervention delivered instead in person and 1 with a control group. The authors found that there was no positive effect on cognitive skills following participation in these interventions.

Cooper 2013 (N=7,896) systematically reviewed RCTs looking at any intervention intended for mild cognitive impairment on cognitive, neuropsychiatric or functional outcomes, quality of life and the onset of dementia. The focus was on preventing further decline, rather than rehabilitation. This review included 41 placebo-controlled papers, 20 of which included primary outcomes, 9 of which investigated psychological interventions, 5 of which investigated exercise interventions and 22 of which investigated pharmacological or dietary interventions. The authors concluded that there was no replicated evidence that any intervention was effective.

Farina 2015 and O'Brien 2008 both reviewed rehabilitative interventions in neurological conditions including multiple-sclerosis (MS) and epilepsy. Farina 2015 (N=640) narratively reviewed 18 studies relating to issues in cognitive rehabilitation for epilepsy. 9 of these were experimental papers testing out various rehabilitative strategies, 2 randomised and 7 observational. The strategies used included psychoeducation, imagery, compensatory strategies and cognitive training programmes. The authors concluded that a holistic rehabilitative approach was more useful than selective interventions for cognitive impairments in this group. O'Brien 2008 (N=787) conducted a review of cognitive rehabilitative interventions in people with multiple sclerosis. Included studies ranged from RCTs through to uncontrolled studies or case reports. The examined interventions were designed to target attention, memory or executive functioning and included computer-based interventions, memory techniques, repetition strategies, psychoeducation and psychological therapy. The authors concluded that although some memory strategies appeared promising, that further research was needed to inform practice recommendations.

Table 146: Study characteristics for the narrative review of rehabilitative interventions for acquired cognitive impairment in the criminal justice system

Study ID	Chung 2013	Coleman 2015	Cooper 2013	Farina 2015	Krasny-Pacini 2013	O'Brien 2008	Whyte 2011
Type of review	Systematic	Systematic	Systematic	Narrative	Narrative	Narrative	Narrative
Total number of studies (N[1])	13 (770)	10 (388)	41 (7,896)	18 (640)	12 (NR)	16 (787)	NA
Types of study	RCTs	RCTs (k=9), non-randomised crossover study (k=1); of these assessment studies (k=8), intervention studies (k=2)	RCTs	RCTs, observational studies; of these intervention studies (k=9)	RCTs, case-reports	RCTs, observational studies, case reports	NA
Diagnosis	Stroke or non-progressive ABI	Stroke or non-progressive ABI	Mild cognitive impairment	Epilepsy	ABI	MS	ABI
Interventions	Restorative (including neurorehabilitation, goal management, self-awareness and working memory training), compensatory (including neurorehabilitation, and video-feedback) or adaptive interventions	APSST and MOPS; both delivered via tele-practice	Pharmacological and dietary (including Donepezil for dementia and fish oils), computer-assisted cognitive training, group psychological interventions (including psychoeducation and memory training), family interventions, exercise	Cognitive training, computer-assisted memory training, compensatory memory strategies, psychotherapy, psychoeducation, meta-cognitive therapy, mental imagery, occupational training,	Goal management training	Computer-assisted programmes, memory aids, metacognitive therapy, communication skills, psychoeducation	Adaptive and remediative interventions
Treatment length	2 weeks – 1 year	20 x 45 min sessions and	3 weeks – 2 years	NR	NR	4 weeks – 6 months	NA

Study ID	Chung 2013	Coleman 2015	Cooper 2013	Farina 2015	Krasny-Pacini 2013	O'Brien 2008	Whyte 2011
		1 hr/wk for 6 wks					
Comparator	Active intervention, no intervention or placebo	APPST in person Control	Variable	NR	NR	Variable	NA

[1]N=total number of included participants
k=number of studies
ABI=acquired brain injury
APPST=analogical problem solving skills training
MOPS=military online problem-solving video-phone intervention
MS=Multiple Sclerosis
RCT=randomised controlled trial
NR=Not reported
NA=Not applicable

6.8.12 Expert testimony

Professor Huw Williams, Associate Professor of Clinical Neuropsychology and Co-Director of the Centre for Clinical Neuropsychology Research (CCNR) at Exeter University, provided expert testimony on the relationship between traumatic brain injury (TBI) and mental health problems in young offenders. This is described in greater detail in his own words within Appendix W. The guideline committee sought this expert testimony due to the lack of direct evidence about the rehabilitation of adults with acquired cognitive impairment in contact with the criminal justice system.

Professor Williams highlighted to the GC the high prevalence of TBI in individuals in contact with the criminal justice system, and the strong correlations between TBI and mental health problems, in particular substance misuse, self-harm and suicide. He also described the economic and social cost of this link. Professor Williams provided theoretical reasoning and pre-clinical evidence for this association and areas of potential focus for intervention. Professor Williams argued that identification of individuals with a history of TBI is key, and that more research is required to identify ways of supporting this group.

6.8.13 Economic evidence

No studies assessing the cost effectiveness of psychological, social or pharmacological interventions for adults with acquired cognitive impairment who are in contact with the criminal justice system were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.8.14 Clinical evidence statements

No direct evidence was found about the effect of rehabilitative interventions on cognitive or adaptive functioning and offending outcomes in people with cognitive impairment in contact with the criminal justice system.

Low quality, indirect evidence from 7 systematic reviews (N>10,481) of studies conducted in non-criminal justice populations indicated no clinically significant improvement in cognitive or adaptive functioning from a range of interventions including psychological, pharmacological and adaptive interventions that could be considered for the remediation of deficits associated with ACI.

6.8.15 Economic evidence statements

No evidence on the cost effectiveness of psychological, social or pharmacological interventions for adults with acquired cognitive impairment who are in contact with the criminal justice system is available.

6.9 Recommendations and link to evidence

Recommendations	No recommendation made
Relative values of different outcomes	The GC agreed that given the high prevalence of acquired cognitive impairment (ACI) in the criminal justice population, that identification was very important, even if no appropriate rehabilitative interventions are currently available because knowledge of the presence of ACI could impact on an understanding of a person's problems and contribute to the development of any care of management plan.
Trade-off between clinical benefits and	There was no evidence that directly related to the use of interventions to manage ACI in the criminal justice system or provide direct evidence on any

harms	harms. The GC agreed that there was a potentially significant clinical benefit from identifying service users who had experienced ACI, as this may assist with clinical decision making, development of management plans and the assessment of risk. This may in time contribute to overall better care and management in the CJS and the NHS and possibly to a reduction in criminal activity. The GC did not identify any harms associated with this, other than the possible harms associated with a false positive arising from inaccurate identification. In developing recommendations in this area the GC drew on expert testimony and used informal consensus to develop their recommendations
Trade-off between net health benefits and resource use	There was no evidence on the cost-effectiveness of interventions for people with acquired cognitive impairment who are in contact with the criminal justice system. The GC expressed the view that any additional costs associated with the identification, assessment and provision of appropriate care are likely to be offset by the negative consequence associated with lack of knowledge of the presence of the acquired cognitive impairment and inadequately developed care plans. The GC considered the increased rate of ACI in this population and the potential life-long physical and mental problems (many psychological conditions are more prevalent in this population) it can cause, and associated high health care costs. The GC also considered the link between ACI and greater convictions and violence, and the associated increase in the costs to the criminal justice sector.
Quality of evidence	No direct evidence was found for interventions to remediate difficulties associated with ACI in adults within the criminal justice system. In the absence of direct evidence on interventions for people with ACI, indirect low quality evidence on cognitive rehabilitation of ACI with multiple, different causes from 7 systematic reviews (no one of which focused exclusively on ACI) was considered, as well as expert testimony. They showed limited evidence of some benefit when particular cognitive functions were targeted by remediative interventions (e.g. short-term memory function, attention, executive function) which was not directly related to ACI or in the view of the GC could not be applied to ACI. The absence of populations drawn from the CJS, the laboratory based and experimental nature of a number of the interventions and, limited testing in routine health care settings in these reviews also contributed to the GC feeling unable to make any recommendations for any specific interventions for ACI.
Other considerations	The GC agreed that it was important to make recommendations relating to the identification of ACI in this group even though there was no high quality evidence showing that interventions can remediate the deficits associated with ACI. This was because these service users have a higher risk of self-harm and an awareness of the presence of ACI could help a person better adapt to their difficulties and that this information might also inform the general care and management of a person. On this basis the GC decided that a question should be added to the first stage of reception screening in prison to facilitate identification of ACI. They also decided that a recommendation should be made for staff to receive training on the impact of ACI in service users within the criminal justice system. The GC also agreed that given the lack of quality evidence in a condition with a high prevalence in the criminal justice population and with potentially significant clinical implications, they would make a research recommendation to look at the effectiveness of remedial interventions in the CJS for ACI.

6.9.11 Research recommendations

6. What interventions are clinically and cost-effective for the remediation of difficulties associated with acquired brain injuries (including TBI) in adults with mental health problems within the criminal justice system?

Acquired brain injuries are common in adults in contact with the criminal justice system and are associated with an increased prevalence of mental health problems including increased suicidal risk and an increased risk of re-offending. Recognition of ACI is poor and there is currently no effective intervention used in the CJS to address the problems presented by ACI. This leads to poor management in the criminal justice system and poor longer term outcomes in terms of mental health and offending. There is limited evidence on effective models to remediate the consequences of ACI in the general population but no evidence for remediative interventions in the adult criminal justice population. A programme of research and development is required which will (a) develop novel interventions for remediation specially to address the type of ACI commonly seen in the adult CJS population (b) test these interventions in small pilot studies and (c) if the pilot studies show promise test the interventions in large scale randomised clinical trials in the criminal justice system

Important outcomes could include:

- Improved adaptive functioning
- Improved cognitive performance
- Improved mental health
- Reductions in offending

6.10 Review question: For adults with a personality disorder (other than antisocial or borderline personality disorder) who are in contact with the criminal justice system, what are the benefits and harms of psychological, social or pharmacological interventions aimed at reducing personality disorder symptomatology, or preventing or reducing offending or reoffending?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 147. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 147: Clinical review protocol summary for the review of interventions to reduce symptomatology, offending and reoffending in adults with a personality disorder other than antisocial or borderline personality disorder

Component	Description
Population	Adults with, a personality disorder (other than antisocial or borderline personality disorder) who are in contact with the criminal justice system
Intervention(s)	• Psychological, pharmacological and social interventions
Comparison	• Treatment as usual • No treatment • Waitlist control

Component	Description
	• Placebo (including attention control) • Any alternative management strategy
Outcomes	• Critical – Improvement in symptom severity (e.g. borderline personality disorders); Offending and re-offending outcomes; Rates of self-harm; • Important - Adaptive functioning (for example, employment status, development of daily living and interpersonal skills and quality of life)
Study design	Systematic reviews and RCTs

6.10.1 Clinical evidence for the most appropriate assessment procedures and interventions for individuals with a personality disorder within the criminal justice system

No RCT evidence was identified for this question. The GC decided it would be inappropriate to descend the evidence hierarchy as they were aware, on the basis of their existing knowledge of the literature, that it was unlikely to be fruitful and was therefore not considered a good use of time and resource and given the very high prevalence of personality disorders among people in contact with the criminal justice system any recommendations about assessments or interventions could have significant cost impact and should not be based on low quality evidence from non-randomised studies. They decided that extrapolation from non-criminal justice populations would not be appropriate as the criminal justice system is very different from other settings. The GC therefore decided to develop a set of principles to inform assessment and intervention for personality disorders within this population using a modified form of the nominal group technique. The method used for the nominal group technique is described in full within the methods section in Chapter 3.

Key issues related to assessment and intervention within this population were identified through a range of sources and from discussions within the GC meetings. These issues were used to generate nominal statements covering a range of areas that had been identified as important by the GC. These included an understanding of how a personality disorder diagnosis may impact upon psychological wellbeing and interpersonal skills, about common co-occurring difficulties within this group, and the ways that interventions should be delivered to best support service users. These statements were grouped together in the form of a questionnaire and distributed to the GC to be rated. An example of statement that was rated highly by the committee is ‘People with personality disorders should not be excluded from any health or social care service because of their diagnosis’.

The questionnaire was completed by 12 of the 19 GC members. Some members were unable to attend the relevant committee meeting, however they had the opportunity to discuss the statements from the nominal group process and contributed to the subsequent recommendations. Percentage consensus values were calculated, and comments collated, for each statement. The rankings and comments were then presented to the GC members, and used to inform a structured discussion within the GC meeting. Agreement within the GC was high enough that a second round of ratings was not deemed necessary. This discussion led to the development of recommendations in this area. A brief summary of the outcome of this process is depicted in Table 148 below. The full list of statements and ratings can be found in Appendix V and blank copies of the questionnaires used can be found in Appendix U.

Table 148: Summary of the nominal group technique process followed for the development of recommendations for the care, assessment and interventions for people with a personality disorder within the criminal justice system

Round 1		Round 2		No. of recommendations generated
Level of agreement	Statements N (total=24)	Level of agreement	Statements N (total=0)	5 recommendations
High	23	High	n/a	
Moderate	1	Moderate	n/a	
Low	0	Low	n/a	

6.10.2 Economic evidence

No studies assessing the cost effectiveness of psychological, social or pharmacological interventions for adults with a personality disorder (other than antisocial or borderline personality disorder) who are in contact with the criminal justice system were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.10.3 Clinical evidence statements based upon formal consensus ratings

The GC endorsed statements relating to principles of care stating that:

- a personality disorder diagnosis should not result in preventing service users accessing services
- staff should be aware that this population may have longstanding impairments in a range of areas of functioning including interpersonal difficulties, that structure and clear expectations are helpful for this group of service users and that a personality disorder diagnosis may complicate treatment of co-occurring disorders
- it is important to be both validating and judiciously challenging when interacting with these service users.

Regarding assessment, the GC endorsed statements stating that staff should:

- be able to identify and appropriately adjust for common features of a personality disorder
- be aware that these service users may struggle to interpret and manage emotions, have difficulties with impulse control, feel as though they have a lack of autonomy and have an unstable sense of self or struggle with social functioning
- establish which other services are involved in the care of the person with a personality disorder, and clarify the roles and responsibilities of each service.

Regarding interventions, the GC endorsed statements stating that:

- if complex interventions are required these should be delivered in a multi-disciplinary setting
- staff should ensure that adequate case management and advocacy are in place for the service user
- interventions should be supportive, facilitate learning and encourage the development of new behaviours, and that the service user should be offered interventions for any comorbid disorders in line with relevant NICE guidelines

- staff should work alongside the service user to develop a crisis plan and assist them to feel responsible for their care
- when changing treatments or services that a structured and phased approach should be taken
- when developing care plans the following components should be considered; problem-solving, articulation and management of emotions, managing interpersonal relationships, impulse control, self-harm and medication management.

The GC expressed moderate agreement for increasing the duration or intensity of psychological interventions.

6.10.4 Economic evidence statements

No evidence on the cost effectiveness of psychological, social or pharmacological interventions for adults with a personality disorder who are in contact with the criminal justice system is available.

6.11 Recommendations and link to evidence

Recommendations	**44. Be aware that many people in contact with the criminal justice system, (including people with a diagnosis of personality disorder) may have difficulties with:** • **accurately interpreting and controlling emotions** • **impulse control (for example, difficulty planning, seeking high levels of stimulation, ambivalent about consequences of their negative actions)** • **experiencing themselves as having a lack of autonomy (for example, seeing their actions as pointless, having difficulties in setting and achieving goals)** • **having an unstable sense of self that varies depending on context or is influenced by the people they interact with** • **social functioning (for example, relating to, cooperating with, and forming relationships with others, difficulties understanding their own and others' needs)** • **occupational functioning.** **45. Providers of services should ensure staff are able to identify common features and behaviours associated with personality disorders and use these to inform the development of programmes of care.** **46. Practitioners should ensure interventions for people with a diagnosis of personality disorder or associated problems are supportive, facilitate learning and develop new behaviours and coping strategies in the following areas:** • **problem solving** • **emotion regulation and impulse control** • **managing interpersonal relationships** • **self-harm** • **medicine management (including reducing polypharmacy).** **47. Practitioners should be aware when delivering interventions for people with mental health problems that having a personality disorder or an associated problem may reduce the effectiveness of interventions. Think about:** • **providing additional support.** • **adjusting the duration and intensity of psychological interventions if standard protocols have not worked** • **delivering complex interventions in a multidisciplinary context.** **48. Practitioners should not exclude people with personality disorders from any health or social care service, or intervention for comorbid disorders, as a direct result of their diagnosis.**
Relative values of	The GC discussed issues specific to work with individuals with personality

different outcomes	disorders in the criminal justice system including particular difficulties with establishing and maintaining a therapeutic relationship, the need for more complex therapeutic interventions, the greater levels of risk and difficult social relationships. They noted that personality disorders are often poorly understood, and that people with a diagnosis of personality disorder are sometimes denied access to services as a result of this diagnosis. Despite these problems people with personality disorder are over represented in the CJS and criminal justice population and in groups of people who make high use of emergency health care services. Effective access to services followed by prompt treatment may therefore have implications not only for improved mental health and well-being of people with personality disorder but also reduce service utilisation.
Trade-off between clinical benefits and harms	The GC discussed the fact that greater awareness on the part of staff about the nature of personality disorders and information about how best to approach this group therapeutically could have a significant positive clinical impact. They agreed that adapting interventions so that they are more structured, treatment sessions are more frequent or of longer duration, or working alongside other professionals and also collaboratively with the individual were likely to result in improved therapeutic engagement, better clinical outcomes and less use of services in the future. They agreed that clinicians feeling confident enough to maintain structure and boundaries in therapeutic relationships is key to working with this group. The GC agreed that, given the proper adaptation of effective interventions there would likely be no harms associated with psychosocial interventions for people with a personality disorder and that it may lead to a reduction in the extent of the self-harm often seen in people with personality disorder.
Trade-off between net health benefits and resource use	Effective treatment for people with personality disorders is likely to lead to increase use of health service resources in the short-term arising from the need for more intensive, structured treatments of longer duration. However given that such individuals are higher uses of emergency health services and over represented in the prison system that the effective treatment could lead to significant costs savings in the long term.
Quality of evidence	No RCT evidence was identified that was relevant to this review. The GC used a nominal group technique to generate evidence statements to support the development of the recommendations. This evidence was of low quality and was used make general 'in principle' recommendations for this population of service users on the basis of their expert knowledge. These statements focused on interpreting and controlling emotions, impulse control, lack of autonomy, having an unstable sense of self, and social and occupational functioning. The GC felt that these problems could be addressed by improved by interventions focused on problem solving, emotion regulation and impulse control, managing interpersonal relationships, reducing self-harm and better medicine management, along with adjustments to the delivery of psychological interventions.
Other considerations	The GC were aware of the need to produce recommendations that supported the provision of effective interventions that are in line with existing NICE guidance including that on personality disorder. The GC was particularly concerned with the engagement of individuals so that they could access effective NICE recommended interventions. With this in mind, and because of the limited quality of evidence in criminal justice populations, the group focused the development of the general principle recommendations which would guide general treatments of people with personality disorders and the use of NICE guidance. Given that much evidence on personality disorder focused on borderline and antisocial personality disorder the GC decided to make a research recommendation for research into psychosocial interventions for people with other types of personality disorder.

6.11.1 Research recommendations

7. What psychosocial interventions are clinically and cost-effective for people with a personality disorder (other than ASPD or PBD) within the criminal justice system?

Personality disorders are common in adults in contact with the criminal justice system and are associated with an increased risk of re-offending, increased self-harm and suicidality and increased drug and alcohol misuse. Personality disorder may also contribute to significant management problems in the criminal justice system, these management problems may in part arise because the disorders are not recognised and potentially effective interventions are not made available. There are effective treatments for antisocial and borderline personality disorders and, in particular, for antisocial personality disorder are available in the criminal justice system. However, although other types of personality disorder are also present in the criminal justice population there is very limited evidence to guide effective treatment for these problems. A programme of research and development is required which will (a) develop interventions for personality disorder (other than ASPD or PBD) within the criminal justice system specially for use in the adult CJS population (b) test these interventions in a series of pilot studies and (c) if the pilot studies show promise, test the interventions in large scale randomised clinical trials in the criminal justice system

Important outcomes could include:

- Remission of the disorder
- Improved interpersonal performance
- Improved mental health
- Reductions in offending

7 Service Delivery

7.1 Introduction

People with mental health problems and in contact with the criminal justice system receive interventions from a wide range of mental services provided by statutory (including health and social care), and the criminal justice service and voluntary sector. Services may be provided by a number of these organisations simultaneously and the coordination of activity across these various agencies remains a major challenge; a challenge that is compounded by the very different cultures of the organisations, differing rules regarding confidentiality and the incompatibility of many information systems. In addition to these organisational challenges, the knowledge and skills of the staff to meet these challenges varies considerably and many staff may lack basic knowledge and understanding of organisations and agencies other than their own.

Despite the difficulties arising from the challenges above, there have been a number of developments which seeks to address this difficulties; they include the development of Street triage models of care (Reveruzzi, 2016) which promote better from working between criminal justice and health care staff, the development of court diversion and liaison schemes to better support mentally ill people who enter the court system (Sainsbury Centre for Mental Health, 2009) , the development of specialist metal health or drug courts (Winstone,2010), the development of psychologically informed prison environments to prompt mental health and well-being in the prisons system (Turley). A more substantial change in the prison system has been the provision of specialist mental health teams commissioned and provided by health services with the prison system.

Another significant problem that remains, is the effective engagement of services users into health care services at all levels of the criminal justice system; engagement is typically poor and there is a need for clear pathways and case management systems in place which support engagement and ensure effective transitions between services. The evidence considered in the following sections should be seen not in isolation as evidence about service delivery but also in the context of a changing political landscape about prison organisation and community supervision of offenders; these factors may affect the relevance of emerging findings and recommendations.

7.2 Review question: What are the most effective care plans and pathways, and organisation and structure of services, for the assessment, intervention and management of mental health problems in people in contact with the criminal justice system to promote:

- appropriate access to services?
- positive experience of services?
- positive mental health outcomes?
- integrated multi-agency care?
- successful transition between services?
- successful discharge from services?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 149: Clinical review protocol summary for the review of the most effective care plans and pathways, and organisation and structure of services, for the assessment, intervention and management of mental health problems in people in contact with the criminal justice system

A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 149: Clinical review protocol summary for the review of the most effective care plans and pathways, and organisation and structure of services, for the assessment, intervention and management of mental health problems in people in contact with the criminal justice system

Component	Description
Population	Adults (aged 18 and over) with, or at risk of developing, a mental health problem who are in contact with the criminal justice system
Intervention(s)	Any service delivery model, including: • Assertive Community Treatment (ACT) • Case management (including intensive case management) • CARAT (Counselling, Assessment, Referral, Advice and Throughcare) • Collaborative care • Dangerous and Severe Personality Disorder (DSPD) programme • Drug Arrest Referral Schemes (DARS) • Drug Interventions Programme (DIP) • Drug Rehabilitation Requirements (DRRs) • Drug Treatment and Testing Orders (DTTO) • Integrated Drug Treatment System (IDTS) • Mental health courts • Prison/court liaison and diversion programmes • Psychologically Informed Planned Environments (PIPEs) • Re-entry programmes • Street triage
Comparison	• Treatment as usual • No treatment • Waitlist control • Placebo (including attention control) Any alternative service delivery model
Outcomes	• Critical – Service /utilization outcomes (e.g. hospital admissions,s136 detentions); Mental health outcomes • Important – Offending and re-offending; Adaptive functioning (for example, employment status, development of daily living and interpersonal skills and quality of life);
	• RCTs, systematic reviews • Non-randomised controlled trials, Prospective or Retrospective cohort studies, Before and After (B & A) studies were included if they were done in UK

7.2.1 Clinical evidence

7.2.1.1 Street Triage

Three before and after observational cohort studies (N=13303) met the eligibility criteria for this review: Hywel Dda 2015, Powys 2015 and Reveruzzi 2016(Dyfed Powys Police and Powys Teaching Health Board., 2015; Morgan, 2015; Reveruzzi & Pilling, 2016). Street Triage is a joint police/health care which works with people who present in a mental health crisis in public places and might typically be taken to a place of safety under s136 of the

Mental Health Act. Aims of Street triage include reducing the number of s136 and seeking alternative routes into care.

An overview of the included studies can be found in Table 150. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 151: The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of mental health, offending and reoffending outcomes, adaptive functioning and rate of self-injury.

Table 150: Study information table for the analysis of before and after street triage scheme

	Street triage scheme
Total no. of studies (N[1])	3 (200464)
Study ID	(1) Hywel Dda 2015 (2) Powys 2015 (3) Reveruzzi 2016
Study design	Before and after study
Country	(1, 2) Wales, UK (3) England, UK
Underlying Mental Health Disorders	Any
Diagnosis	Clinical
Age (mean/range) years	(1)18 to 59 (84%) Male; 18-59 (90%) Female (2, 3) Not reported
Gender (% female)	(1)46% (2, 3) Not reported
Ethnicity (% white)	(1, 2) Not reported (3)70%
Criminal justice setting	(1, 2, 3) Community
Treatment length (weeks)	(1, 2)52 (3)82[2]
Follow-up length (weeks)	Not applicable
Intervention (mean dose; mg/day)	After Street triage scheme
Comparison	Before Street triage scheme
Notes. N= total number of participants; NR=Not reported [1]= number of participants under section 136 detention after street triage; [2] = longest duration	

Table 151: Summary of findings table for before versus after street triage schemes for mental health disorders

Outcomes	**No of Participants (studies) Follow up**	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with no street triage	**Risk difference with street triage (95% CI)**
Total s136 detentions per 100,000	200000* (1 B & A study)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.83 (0.63 to 1.1)	1 per 1000	18.2 fewer per 100,000 (from 39.6 fewer to 10.7 more)
Number of s136 detentions in	49914	⊕⊕⊖⊖	RR 0.68	361 per 1000	115 fewer per 1000

custody per total number of s136 detentions	(2 B & A studies)	LOW[1,3,4]	(0.67 to 0.7)		(from 108 fewer to 119 fewer)
Number of s136 detentions in hospital per total number of s136 detentions	49953 (3 B & A studies)	⊕⊖⊖⊖ VERY LOW[1,5]	RR 1.18 (1.16 to 1.19)	639 per 1000	115 more per 1000 (from 102 more to 121 more)

[1] *Reveruzzi 2016 - before and after study; low risk of selection bias as the groups were formed by before and after implementation of street triage; high risk of performance bias as there was no blinding involved; high rate of missing data and complete case analysis*
[2] *The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.*
[3] *Hywel Dda 2015 - before and after study; low risk of selection bias as the groups were formed by before and after implementation of street triage; high risk of performance bias as there was no blinding involved; high rate of missing data and complete case analysis.*
[4] *Evidence was upgraded if the effect estimate was considered to be large (i.e. 95% CI of RR <0.75 or RR>1.25).*
[5] *Powys 2015 - before and after study; low risk of selection bias as the groups were formed by before and after implementation of street triage; high risk of performance bias as there was no blinding involved; high rate of missing data and complete case analysis*

*The total population being looked at was not provided and the data was calculated per 100,000.

7.2.122 Diversion Services

Four observational studies (N=1002) met the inclusion criteria for this review: Chambers 1999, Exworthy 1997, Weaver 1997 and James 2002(Chambers & Rix, 1999; Exworthy & Parrott, 1997; James et al., 2002; Weaver et al., 1997).

Chambers 1999 study was a controlled cohort study where prisoners were assessed by a doctor or a nurse prior to appearing before the magistrates and compared with no assessment. Exworthy 1997 and Weaver 1997 studies were before and after studies of the court diversion. In Exworthy 1997 study, a psychiatrist attended the court once a week whereas in Weaver 1997 study, offenders were referred to Bentham unit (a remand and assessment service for mentally disorder patients) based in a hospital. At the same time, James 2002 study, which was a controlled cohort study, compared between community and court diversion services.

An overview of the studies included in the analysis can be found in Table 152:. Further information about both included and excluded studies can be found in Appendix L.
Summary of findings can be found in

Table 153 and Table 155. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of quality of life or service user and carer satisfaction.

Table 152: Study information table for trials included in the analysis of diversion services

	Before and After Court Diversion (Same setting)	Assessment at Court vs No assessment	Court vs Community Diversion
Total no. of studies (N)	2 (653)	1(284)	1(428)
Study ID	(1) Exworthy 1997 (2) Weaver 1997	Chambers 1999	James 2002
Study design	Before and after study	Retrospective cohort study	Retrospective cohort study

	Before and After Court Diversion (Same setting)	Assessment at Court vs No assessment	Court vs Community Diversion
Country	(1, 2) UK	UK	UK
Underlying Mental Health Disorders	(1, 2) Mental illness	Not reported	Mental illness
Diagnosis	(1, 2) Clinical	Not reported	Clinical
Age (mean/range) years	(1)30.8 (2) NR	28	35.1
Gender (% female)	(1, 2) Not reported	Not reported	15
Ethnicity (% white)	(1, 2) Not reported	93[1]	59
Criminal justice setting	Prisoners on remand	Prisoners on remand	NA
Treatment length (weeks)	(1)78 (2)22	26	Not reported
Intervention (mean dose; mg/day)	After - (1) Custody Scheme (2) Bentham unit	Court diversion after a doctor or a nurse's assessment	Court Diversion
Comparison	Before (1) Custody Scheme (2) Bentham unit	No assessment	Community Diversion
Notes. N=total number of participants; 1 Doctor group only; NA=Not applicable			

Table 153: Summary of findings table for trials included in the analysis of diversion services (before and after)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with before diversion services Control mean±SD	Risk difference with after diversion services versus before diversion services (95% CI)
Duration between remand and assessment (days)	611 (2 B & A studies)	⨁⊖⊖⊖ VERY LOW[1,2]		47.1±78.1	MD 21.64 lower (29.87 to 13.41 lower)
Days of total time on remand	565 (1 B & A study)	⨁⊖⊖⊖ VERY LOW[1]		67.1±71.3	MD 17.6 lower (28.64 to 6.56 lower)

[1] Exworthy 1997- before and after study with no confounder being controlled; no blinding; unclear drop out and available case analysis
[2] Weaver 1997 – before and after study with no confounder being controlled; no blinding; unclear dropout with available case analysis

Table 154 Summary of findings table for assessment by a doctor or a nurse versus no assessment at court

Outcomes	No of	Quality of	Relative	Anticipated absolute effects

	Participants (studies) Follow up	the evidence (GRADE)	effect (95% CI)	Risk with no assessment	Risk difference with assessment versus no assessment at court (95% CI)
Proportions of prisoners on bail (doctor' or nurse's assessment vs no assessment)	220 (1 retrospective cohort study)	⊕⊖⊖⊖ VERY LOW[3,4]	RR 1.25 (0.76 to 2.04)	204 per 1000	51 more per 1000 (from 49 fewer to 212 more)
Attendance at alcohol and drug treatment programmes (doctor' or nurse's assessment vs no assessment)	70 (1 retrospective cohort study)	⊕⊖⊖⊖ VERY LOW[3,4]	RR 1.02 (0.51 to 2.07)	310 per 1000	6 more per 1000 (from 152 fewer to 332 more)
OPD attendance rate for those release on bail (doctor' or nurse's assessment vs no assessment)	36 (1 retrospective cohort study)	⊕⊖⊖⊖ VERY LOW[3,4]	RR 0.89 (0.46 to 1.72)	538 per 1000	59 fewer per 1000 (from 291 fewer to 388 more)
Registration of care programmes and supervision registration (doctor' or nurse's assessment vs no assessment)	220 (1 retrospective cohort study)	⊕⊖⊖⊖ VERY LOW[3,4]	RR 2.01 (0.65 to 6.21)	41 per 1000	41 more per 1000 (from 14 fewer to 213 more)

[3] Chambers 1999 – retrospective cohort study with no confounder being controlled; no blinding; unclear drop out and available case analysis
[4] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

Table 155: Summary of findings table for court diversion vs community diversion

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with community diversion	Risk difference with court diversion versus community diversion(95% CI)
Rate of re-incarceration in two years after index discharge	428 (1 retrospective cohort study)	⊕⊕⊖⊖ LOW[1,2]	RR 5.45 (2.95 to 10.08)	51 per 1000	229 more per 1000 (from 100 more to 467 more)
100% attendance rate of appointments	428 (1 retrospective cohort study)	⊕⊖⊖⊖ VERY LOW[1,3]	RR 0.59 (0.44 to 0.81)	369 per 1000	151 fewer per 1000 (from 70 fewer to 207 fewer)
Number of days in hospital	428 (1 retrospective cohort study)	⊕⊖⊖⊖ VERY LOW[1]		Control mean 129	MD 17 lower (64.44 lower to 30.44 higher)
Number of diverted participants with no mental health disorders	428 (1 retrospective cohort study)	⊕⊖⊖⊖ VERY LOW[1,3]	RR 13 (0.74 to 229.33)		

[1] James 2002 - retrospective cohort study; No blinding; Few missing cases and available case data analysis
[2] The effect size is considered large if 95% of RR<0.8 or RR>1.25.
[3] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

7.2.1521 Patient Navigation Intervention

One RCT (N= 40) met the eligibility criteria for this review: Binswanger 2015(Binswanger et al., 2015). This study assessed the patient navigation programme where participants were directed to Colorado Indigent Care Program (CICP) themselves and it was compared with facilitated enrolment where participants were referred to CICP by enrolment specialist for facilitated enrolment. CICP was a programme funding to clinics and hospitals so that medical services could be provided at a discount.

An overview of the trials included in the analysis can be found in Table 156. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 157. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of mental health and service utilization rate.

Table 156 Study information table for trials included in the analysis of patient navigation intervention versus facilitated enrolment for substance misuse disorders

	Patient navigation intervention
Total no. of studies (N[1])	1 (40)
Study ID	Biswanger 2015
Study design	RCT
Country	USA
Underlying Mental Health Disorders	Substance misuse disorders
Diagnosis	Unclear
Age (mean)years	42.4
Gender (% female)	18
Ethnicity (% white)	Not reported
Criminal justice setting	In the community
Treatment length (weeks)	13
Follow-up length (weeks)	26
Intervention (mean dose; mg/day)	Patient navigation plus care discount programme
Comparison	Facilitated enrolment into indigent care programme
Notes. N=total number of participants; [1]Number being randomised	

Table 157 Summary of findings table for patient navigation intervention versus facilitated enrolment at 26 weeks follow-up for substance misuse disorders

Outcomes	**No of Participants (studies) Follow up**	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with facilitated enrolment	**Risk difference with patient navigation intervention versus**

					facilitated enrolment (at 26 weeks follow-up) (95% CI)
Number of participants who used drugs	18 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.62 (0.07 to 5.72)	200 per 1000	76 fewer per 1000 (from 186 fewer to 944 more)
Number of participants who used alcohol to intoxication	18 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.42 (0.05 to 3.28)	300 per 1000	174 fewer per 1000 (from 285 fewer to 684 more)
Average days when mental health was not good in the last 30 days	18 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,3]		Control mean 8.6	MD 1.1 lower (9.74 lower to 7.54 higher)

[1] Binswanger 2015 – RCT; unclear randomization with appropriate allocation concealment, no blinding and appropriate attrition rate
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[3] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

7.2.172 Neighbourhood outreach

One before and after study (N=213) met the eligibility criteria for this review: Earl 2015(Earl et al., 2015). The service delivery model applied was Cornwall Criminal Justice Liaison and Diversion Services (custody-based and neighbourhood outreach services) which operated from Monday to Friday 9am to 5pm staffed by Community Psychiatric Nurses. This service assessed people with apparent vulnerability and/or mental ill health coming to the attention of public services without meeting thresholds for criminal intervention or imminent mental health crisis. The outcomes were then compared with data before the service implementation.

An overview of the trials included in the analysis can be found in Table 158. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 159. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcome of mental health.

Table 158 Study information table for trials included in the analysis of neighbourhood outreach (Before and After)

	Before and After Neighbourhood outreach
Total no. of studies (N)	1 (213)
Study ID	Earl 2015
Study design	Before and after study
Country	UK
Underlying Mental Health Disorders	Population with apparent vulnerability and/or mental ill health coming to attention of public services without meeting thresholds for criminal intervention or imminent mental health crisis.

	Before and After Neighbourhood outreach
Diagnosis	Sub-thresholds symptoms
Age (median)years	34.5
Gender (% female)	28.1
Ethnicity (% white)	92.9
Criminal justice setting	In the community
Treatment length (weeks)	Not reported
Follow-up length (weeks)	26
Intervention (mean dose; mg/day)	After Cornwall Criminal Justice Liaison and Diversion Services (CJLDS) (custody-based and neighbourhood outreach services) schemes
Comparison	Before Cornwall CJLDS scheme
Notes. N=total number of participants	

Table 159 Summary of findings table for before versus after neighbourhood outreach for mental health disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with before neighbourhood outreach	Risk difference with after neighbourhood outreach and before neighbourhood outreach (95% CI)
Proportion of crime contacts with policing team escalated to court	506 (1 B & A study)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.68 (0.54 to 0.85)	484 per 1000	155 fewer per 1000 (from 73 fewer to 223 fewer)

[1] Earl 2015 – before and after study; available case analysis; high risk of selective outcome report
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

7.2.1752 Drug Rehabilitation Program

One controlled cohort study (N=73) was included for this review: Naeem 2007(Naeem et al., 2007). In Naeem 2007 study, Drug Rehabilitation Requirement (DRR) [formerly Drug Testing and Treatment Order (DTTO)] was compared with clients in mainstream services. This service level intervention had three main requirements: a treatment requirement, a testing requirement and a court review requirement.

An overview of the trials included in the analysis can be found in Table 160 Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 162. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of offending rate and service utilization rate.

Table 160 Study information table for trials included in the analysis of Drug Rehabilitation Requirement versus TAU

	Drug Rehabilitation Requirement (Previously Drug Testing and Treatment Order) vs TAU
Total no. of studies (N)	1(73)
Study ID	Naeem 2007
Study design	Non-randomised controlled trial
Country	UK
Underlying Mental Health Disorders	Offenders with drug misuse
Diagnosis	Unclear
Diagnosis	31.6
Age (median)years	15
Gender (% female)	NR
Ethnicity (% white)	Not reported
Criminal justice setting	In the community
Treatment length (weeks)	Not reported
Follow-up length (weeks)	52
Intervention (mean dose; mg/day)	DRR (DTTO)
Comparison	Treatment as usual: mainstream services
Notes. n=total number of participants; DRR=Drug Rehabilitation Requirement; DTTO= Drug Testing and Treatment Order	

Table 161 Summary of findings table for DRR vs TAU/mainstream services for mental health disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with mainstream services	Risk difference with DRR versus mainstream services (95% CI)
MAP total scores	52 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]		Control mean 151.8	20.2 lower (52 lower to 11.6 higher)
HoNOS total scores (Scale from 0 to 28; lower better)	52 (1 study)	⨁⊖⊖⊖ VERY LOW[1,2]		Control mean 9.9	0.2 lower (2.44 lower to 2.04 higher)
Overall satisfaction scores (Scale from 0 to 7; higher better)	52 (1 study)	⨁⊖⊖⊖ VERY LOW[1]		Control mean 3.2	2.1 higher (1.16 to 3.04 higher)

[1] *Naeem 2007 –non-randomised controlled trial; missing data imputed by regression*
[2] *The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.*

7.2.1713 Case Management

Thirteen RCTs met the eligibility criteria for this review: Cosden 2003/2005, Cusack 2010, Friedmann2012/Johnson2011, Guydish 2011, Hanlon 1999, Jarrett 2012, Martin 1993, Needels 2005, Prendergast 2011, Rossman 1999, Scott 2012, Solomon 1994 and Wang 2012(Cosden et al., 2005; Cosden et al., 2003; Cusack et al., 2010; Friedmann et al., 2012; Guydish et al., 2011; Hanlon et al., 1999; Jarrett et al., 2012; Johnson et al., 2011; Martin et al., 2013; Needels et al., 2005; Prendergast et al., 2011; Rossman et al., 1999; Scott & Dennis, 2012; Solomon et al., 1994; Wang et al., 2012). Seven trials (Friedmann2012/Johnson2011, Guydish 2011, Hanlon 1999, Needels 2005, Prendergast 2011, Rossman 1999 and Scott 2012) studied case management whereas one trial (Martin 1993) looked at assertive community treatment (ACT) among participants with substance misuse disorders. The other five trials examined different case management among severe mental illness subjects (Cosden 2003/2005, Cusack 2010, Jarrett 2012), schizophrenia (Solomon 1994) and uncategorized mental health disorder (Wang 2012).

An overview of the trials included in the meta-analysis can be found in Table 162 (for substance misuse disorders) and

Table **166** (for mental health disorders other than substance misuse). Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 163, Table 164,

Table 165 and Table 167. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

Table 162: Study information table for trials included in the meta-analysis of case management of substance misuse disorders

	Case management vs Active intervention/Treatment as usual	ACT vs TAU
Total no. of studies (N[1])	7 (3645)	1(400)
Study ID	(1) Fridemann2012/Johnson2011 (2) Guydish2011 (3) Hanlon1999* (4) Needels2005 (5) Prendergast2011 (6) Rossman1999 (7) Scott2012	Martin 1993
Study design	RCT	RCT
Country	(1 to 7) USA	USA
Underlying Mental Health Disorders	(1, 3, 4) Drug misuse (2, 5, 6, 7) Substance (alcohol and/or drug) misuse disorders	Drug misuse
Diagnosis	(1, 5) Symptoms (2, 3, 4, 6, 7) Unclear	Unclear
Age (mean range) years	(1 to 7) 31 to 37	29
Gender (% female)	(1, 3, 5, 6)15 to 24 (2, 4, 7) 100	37
Ethnicity (% white)	(1, 4) Not reported (2, 3, 5 to 7) 8 to 47.2	Not reported
Criminal justice setting	(1, 2, 3, 5, 6) in the community (4, 7) initiated in the prison and continued in the community	In the community
Treatment length (weeks)	(1, 7) 13	Not reported

	Case management vs Active intervention/Treatment as usual	ACT vs TAU
	(2) Not reported (3, 4, 6**) 52 (5) 22 to 35	
Follow-up length (weeks)	(1, 4) 65 (2, 3, 6**) 52 (5) 48 (7) 13	26
Intervention (mean dose; mg/day)	(1) Collaborative behavioural management (once a week) (2) Case management (3) Case management and urine monitoring (4) Case management and intensive discharge planning (5) Transitional case management (once a month pre-release; weekly for 3 months after release and monthly for a further 3 months, as required) (6) Opportunity to succeed aftercare program (7) Recovery management check-up	ACT (Not reported)
Comparison	(1) Standard parole (2) Standard probation (3) TAU (routine parole) or Urine monitoring only (4) discharge planning (5) TAU (standard parole) (6) TAU (routine supervision) (7) TAU (not specified)	TAU
*3-armed study; **52 to 104 weeks; TAU – Treatment as usual;		

1

2

3

Table 163 Summary of findings table for case management versus TAU for substance misuse disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Case management versus TAU (95% CI)
Rearrest – Post-treatment	504 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,4,5]	RR 0.9 (0.7 to 1.14)	415 per 1000	41 fewer per 1000 (from 124 fewer to 58 more)
Rearrest - 3 month follow-up	462 (1 RCT)	⨁⨁⊖⊖ LOW[2,4,5]	RR 1.24 (0.88 to 1.74)	202 per 1000	48 more per 1000 (from 24 fewer to 149 more)
Reconviction – Post-treatment	504 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,4,5]	RR 0.76 (0.51 to 1.14)	207 per 1000	50 fewer per 1000 (from 102 fewer to 29 more)
Reincarceration – Post-treatment	504 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,4,5]	RR 0.82 (0.61 to 1.11)	326 per 1000	59 fewer per 1000 (from 127 fewer to 36 more)
Reincarceration - 3 month follow-up	462 (1 RCT)	⨁⊖⊖⊖ VERY LOW[2,5]	RR 1.04 (0.75 to 1.45)	231 per 1000	9 more per 1000 (from 58 fewer to 104 more)
Reincarceration – 12 month follow-up: Total	862 (2 RCTs)	⨁⨁⊖⊖ LOW[5,6]	RR 0.91 (0.76 to 1.10)	346 per 1000	31 fewer per 1000 (from 83 fewer to 35 more)
Reincarceration - 12 month follow-up: female sample	154 (1 RCT)	⨁⊖⊖⊖ VERY LOW[6]	RR 0.73 (0.41 to 1.27)	286 per 1000	77 fewer per 1000 (from 169 fewer to 77 more)
Reincarceration - 12 month follow-up: male sample	708 (1 RCT)	⨁⨁⊖⊖ LOW[5,6]	RR 0.94 (0.77 to 1.16)	359 per 1000	22 fewer per 1000 (from 83 fewer to 57 more)
Number of days jailed in past 6 months (12 month follow-up)	411 (1 RCT)	⨁⨁⨁⊖ MODERATE[6]		Control mean 14.8	MD 0.47 higher (6.65 lower to 7.59 higher)
Drug related crimes in past 6 months (12 month follow-up)	411 (1 RCT)	⨁⨁⊖⊖ LOW[5,6]		Control mean 804.2	MD 25.6 lower (235.88 lower to 184.68 higher)
Drug related criminal activity during treatment (12 months follow-up)	284 (1 RCT)	⨁⊖⊖⊖ VERY LOW[5,7]	RR 0.9 (0.59 to 1.39)	241 per 1000	24 fewer per 1000 (from 99 fewer to 94 more)
Self-reported alcohol use - During treatment	288 (1 RCT)	⨁⨁⊖⊖ LOW[5,7]	RR 0.83 (0.69 to 0.99)	679 per 1000	115 fewer per 1000 (from 7 fewer to 210 fewer)
Self-reported alcohol use - Post-treatment	680 (1 RCT)	⨁⨁⊖⊖ LOW[5,8]	RR 1.09 (0.86 to 1.39)	269 per 1000	24 more per 1000 (from 38 fewer to 105 more)
Self-reported alcohol use – 12 month follow-up: Total	862 (1 RCT)	⨁⊖⊖⊖ VERY LOW[5,6]	RR 0.42 (0.09 to 1.92)	436 per 1000	253 fewer per 1000 (from 397 fewer to 401 more)
Self-reported alcohol	154	⨁⨁⨁⊖	RR 0.18	286 per	234 fewer per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Case management versus TAU (95% CI)
use - 12 month follow-up: female sample	(1 RCT)	MODERATE[6]	(0.07 to 0.5)	1000	(from 143 fewer to 266 fewer)
Self-reported alcohol use - 12 month follow-up: male sample	708 (1 RCT)	⊕⊕⊖⊖ LOW[5,6]	RR 0.83 (0.7 to 0.99)	469 per 1000	80 fewer per 1000 (from 5 fewer to 141 fewer)
Self-reported drug use - During treatment (marijuana)	288 (1 RCT)	⊕⊕⊖⊖ LOW[5,7]	RR 0.81 (0.58 to 1.14)	358 per 1000	68 fewer per 1000 (from 150 fewer to 50 more)
Self-reported drug use - During treatment (hard drugs)	288 (1 RCT)	⊕⊖⊖⊖ VERY LOW[5,7]	RR 1 (0.79 to 1.26)	504 per 1000	0 fewer per 1000 (from 106 fewer to 131 more)
Self-reported drug use - Post-treatment	680 (1 RCT)	⊕⊕⊖⊖ LOW[5,8]	RR 1.07 (0.84 to 1.37)	269 per 1000	19 more per 1000 (from 43 fewer to 100 more)
Self-reported drug use – 12 month follow-up: Total	862 (1 RCT)	⊕⊕⊖⊖ LOW[5,6]	RR 0.76 (0.59 to 0.98)	251 per 1000	60 fewer per 1000 (from 5 fewer to 103 fewer)
Self-reported drug use - 12 month follow-up: female sample	154 (1 RCT)	⊕⊖⊖⊖ VERY LOW[5,6]	RR 0.62 (0.27 to 1.4)	169 per 1000	64 fewer per 1000 (from 123 fewer to 68 more)
Self-reported drug use - 12 month follow-up: male sample	708 (1 RCT)	⊕⊕⊖⊖ LOW[5,6]	RR 0.78 (0.6 to 1.02)	268 per 1000	59 fewer per 1000 (from 107 fewer to 5 more)
Injection drug use (post-treatment)	462 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 0.8 (0.34 to 1.85)	50 per 1000	10 fewer per 1000 (from 33 fewer to 43 more)
Abstinence - During treatment (at 12 months)	283 (1 RCT)	⊕⊕⊖⊖ LOW[5,7]	RR 1.3 (0.86 to 1.94)	221 per 1000	66 more per 1000 (from 31 fewer to 207 more)
Abstinence - Post-treatment	462 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 1.04 (0.75 to 1.45)	231 per 1000	9 more per 1000 (from 58 fewer to 104 more)

[1] Hanlon 1999 - Unclear randomisation; No blinding; Unclear attrition
[2] Scott 2012 - appropriate randomisation with concealment; No blinding; Unclear attrition bias; No selective outcomes report
[3] Evidence was downgraded by one level due to serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of 50%-74.99%) and by two levels due to very serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of >75%). Random Effect Model was used if I-squared inconsistency statistic was more than or equal to 50%.
[4] Evidence was downgraded by one level because study population of one study (Hanlon 1999) differed from the review question in that the study included unclear proportion of ex-heroin/cocaine users.
[5] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[6] Johnson 2011/Friedmann 2012 - Unclear randomisation with unclear allocation concealment; No blinding; ITT analysis; Appropriate outcome report
[7] Rossman 1999 - Appropriate randomisation with allocation concealment; No blinding; Unclear drop-out; Appropriate selective outcome report
[8] Prendergast 2011 - Unclear randomisation with unclear allocation concealment; No blinding; Unclear attrition risk; high risk of selective outcome report

Table 164 Summary of findings table for case management versus active intervention for substance misuse disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with active intervention	Risk difference with case management (95% CI)
Remained in treatment for 6 months	369 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 1.75 (1.31 to 2.33)	343 per 1000	258 more per 1000 (from 106 more to 457 more)
Rearrest - Post-treatment	369 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,5]	RR 0.78 (0.59 to 1.02)	444 per 1000	98 fewer per 1000 (from 182 fewer to 9 more)
Rearrest - 3 month follow-up	511 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,5]	RR 1.1 (0.88 to 1.38)	352 per 1000	35 more per 1000 (from 42 fewer to 134 more)
Rearrest for drug crime (3 month follow-up)	511 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,4]	RR 1.05 (0.73 to 1.5)	186 per 1000	9 more per 1000 (from 50 fewer to 93 more)
Reconviction - Post-treatment	369 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,5]	RR 0.65 (0.4 to 1.05)	212 per 1000	74 fewer per 1000 (from 127 fewer to 11 more)
Reconviction - 3 month follow-up	511 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,5]	RR 1.33 (0.97 to 1.81)	205 per 1000	68 more per 1000 (from 6 fewer to 166 more)
Re-incarceration - Post-treatment	369 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,5]	RR 0.93 (0.64 to 1.35)	283 per 1000	20 fewer per 1000 (from 102 fewer to 99 more)
Re-incarceration - 3 month follow-up	511 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,5]	RR 1.09 (0.86 to 1.39)	326 per 1000	29 more per 1000 (from 46 fewer to 127 more)
Any self-reported drug use (3 month follow-up)	511 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,5]	RR 1.07 (0.86 to 1.33)	379 per 1000	27 more per 1000 (from 53 fewer to 125 more)
Positive hair test (3 month follow-up) - Crack/Cocaine	511 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,6]	RR 1.05 (0.84 to 1.3)	375 per 1000	19 more per 1000 (from 60 fewer to 112 more)
Positive hair test (3 month follow-up) - Marijuana	511 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,6]	RR 0.75 (0.55 to 1.03)	269 per 1000	67 fewer per 1000 (from 121 fewer to 8 more)

[1] Hanlon 1999 - Unclear randomisation; No blinding; Unclear attrition
[2] Evidence was downgraded by one level because study population of one study (Hanlon 1999) differed from the review question in that the study included unclear proportion of ex-heroin/cocaine users.
[3] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[4] Needels 2005 - Unclear randomisation and allocation concealment; No blinding; Available case analysis with unclear drop-out; appropriate outcome report
[5] Kinlock 2007/Kinlock 2009/ Gordon 2008 - Permuted block randomisation with unclear allocation concealment; No blinding; ITT analysis with differing drop-out rates
[6] Evidence was downgraded by one level due to serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of 50%-74.99%) and by two levels due to very serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of >75%). Random Effect Model was used if I-squared inconsistency statistic was more than or equal to 50%.

Table 165 Summary of findings table for assertive community treatment versus TAU

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Assertive Community Treatment versus TAU (95% CI)
Urine test positive for drug use during treatment	90 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 2.33 (0.98 to 5.53)	133 per 1000	177 more per 1000 (from 3 fewer to 604 more)
Self-reported injection drug use during treatment	119 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.8 (0.39 to 1.66)	222 per 1000	44 fewer per 1000 (from 136 fewer to 147 more)
Self-reported drug use during treatment	119 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 1.13 (0.88 to 1.44)	635 per 1000	83 more per 1000 (from 76 fewer to 279 more)
Re-incarcerated during treatment	119 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.91 (0.63 to 1.33)	508 per 1000	46 fewer per 1000 (from 188 fewer to 168 more)

[1] Martin 1993 - Unclear randomisation and allocation concealment; no blinding; Available case analysis with unclear drop-out; appropriate outcome report
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

Table 166: Study information table for trials included in the meta-analysis of case management for mental health disorders other than substance misuse

	Case Management versus TAU/Active intervention
Total no. of studies (N[1])	5 (829)
Study ID	(1) Cosden 2003/2005 (2) Cusack 2010 (3) Jarrett 2012 (4) Solomon 1994* (5) Wang 2012
Study design	RCT
Country	(1, 2, 4, 5) USA (3) UK
Underlying Mental Health Disorders	(1 to 3) Severe Mental Illness (4) Schizophrenia (5) Uncategorized
Diagnosis	(1 to 5) Clinical
Age (mean range) years	(1) NR (2 to 5) 35 to 43
Gender (% female)	(1, 2, 4, 5) 7 to 58.5 (3) NR
Ethnicity (% white)	(1) 83 (2) 63 (3, 4, 5) 7 to 19
Criminal justice setting	(1) Initiated in prison and continued in the community (2) in the community
Treatment length (weeks)	(1, 2, 4, 5) NR (3) 6

	Case Management versus TAU/Active intervention
Follow-up length (weeks)	(1, 2) 104 (3) NR (4) 26 (5) 2
Intervention (mean dose; mg/day)	(1) MHTC with ACT (Non-adversarial court proceedings) (2) FACT (3) Case management with CTI manager (4) ACT or Individual case management (5) Transition clinics – primary care-based complex management program
Comparison	(1) TAU (Adversarial court proceedings) (2) TAU (County-operated public behaviour health system) (3) TAU (Care from prison in-reach team) (4) TAU (Referral CMHC) (5) TAU (Expedited primary care)
*3-armed study; NR-Not reported; MHTC – Mental Health Treatment Court; ACT – Assertive Community Treatment; FACT – Forensic assertive community treatment; CTI – Critical Time Intervention; CMHC – Community Mental Health Centre	

Table 167 Summary of findings table for case management versus treatment as usual for mental health disorders other than substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Case management versus TAU (95% CI)
Service utilization	223 (2 RCTs)	⨁⊖⊖⊖ VERY LOW[1,2,3,4]	RR 0.98 (0.56 to 1.72)	473 per 1000	9 fewer per 1000 (from 208 fewer to 340 more)
Rate of re-offending	432 (3 RCTs)	⨁⊖⊖⊖ VERY LOW[2,4,5,6]	RR 1.04 (0.87 to 1.26)	505 per 1000	15 more per 1000 (from 81 fewer to 136 more)
Number of days in jail (up to 24 months follow-up)	369 (2 RCTs)	⨁⊖⊖⊖ VERY LOW[4,5,6]		Control mean 34.0	MD 12.24 lower (21.87 lower to 2.61 lower)
Quality of life	92 (1 RCT)	⨁⊖⊖⊖ VERY LOW[4,5]		Control mean 4.08	MD 0.09 higher (0.51 lower to 0.69 higher)

[1] *Jarrett 2012 – Unclear randomisation and allocation concealment; No blinding; Available case analysis*
[2] *Wang 2012 – Appropriate randomisation and allocation concealment; Unclear blinding; ITT analysis*
[3] *Evidence was downgraded by one level due to serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of 50%-74.99%) and by two levels due to very serious heterogeneity (chi-squared p<0.1, I-squared inconsistency statistic of >75%).* Random Effect Model was used if I-squared inconsistency statistic was more than or equal to 50%.
[4] *Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.*
[5] *Cosden 2003 – Unclear randomisation and allocation concealment; Unclear blinding; Available case analysis*
[6] *Solomon 1994 – Unclear randomisation and allocation concealment; No blinding; Unclear risk of attrition bias*

[7] *Cusack 2010 – Unclear randomisation and allocation concealment; ITT analysis*

7.2.2011 Drug courts

Four RCTs (N=607) met the eligibility criteria for this review: Dakof2010(Dakof et al., 2010), Gottfredson2005(Gottfredson et al., 2005), Jones2013(Jones, 2013) and Messina2012(Messina et al., 2012).

An overview of the trials included in the analysis can be found in Table 168. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 169 and Table 170**Error! Reference source not found.**. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for mental health outcomes.

Table 168: Study information table for trials included in the analysis of drug court for substance misuse disorders

	Drug court vs Active intervention/Treatment as usual
Total no. of studies (N[1])	4 (607)
Study ID	(1) Dakof 2010 (2) Gottfredson 2005 (3) Jones 2013 (4) Messina 2012
Study design	RCT
Country	(1, 2, 4) USA (3) Australia
Underlying Mental Health Disorders	(1, 3) Substance (alcohol and/or drug) misuse disorders (2, 4) Drug misuse disorders
Diagnosis	(1 to 4) Unclear
Age (mean/range) years	(1) 30.2 (2) 34.8 (3) 32.4 (4) 35.9
Gender (% female)	(1, 4) 100 (2) 26 (3) 16
Ethnicity (% white)	(1) 23 (2, 3) Not reported (4) 58
Criminal justice setting	(1, 2) in court custody (3, 4) in the community
Treatment length (weeks)	(1)52-65 (2) Not reported (3) Mean – 33 (4) approximately 78
Follow-up length (weeks)	(1)78 (2)156 (3) Not reported (4) Mean-96
Intervention	(1) Engaging Moms Drug Court

	Drug court vs Active intervention/Treatment as usual
(mean dose; mg/day)	(2) Baltimore City Drug Court (3) Drug court + Intensive judicial supervision (4) Gender responsive drug court
Comparison	(1) Intensive case management drug court (2) Treatment as usual (standard adjudication) (3, 4) Drug court as usual

1

2

3 **Table 169 Summary of findings table for drug court versus TAU for substance misuse**
4 **disorders**

5

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Drug court versus TAU (95% CI)
Days of substance use (12 month follow-up) - Alcohol	157 (1 study)	⊕⊕⊕⊖ MODERATE[1]		Control mean 85	MD 43.10 lower (46.80 to 39.40 lower)
Days of substance use (12 month follow-up) - Cocaine	157 (1 study)	⊕⊕⊕⊖ MODERATE[1]		Control mean 98.5	MD 43.70 lower (48.16 to 39.24 lower)
Days of substance use (12 month follow-up) - Heroine	157 (1 study)	⊕⊕⊕⊖ MODERATE[1]		Control mean 124.4	MD 54.50 lower (59.42 to 49.58 lower)
Rearrest (12 month follow-up)	157 (1 study)	⊕⊕⊖⊖ LOW[1,2]	RR 0.66 (0.49 to 0.89)	648 per 1000	220 fewer per 1000 (from 71 fewer to 330 fewer)
Maximum Crime Seriousness Scale (12 month follow-up)	157 (1 study)	⊕⊕⊕⊖ MODERATE[1]			1.12 lower (1.18 to 1.06 lower)

[1] Gottfredson 2005 - Unclear randomisation and allocation concealment; No blinding; Unclear analysis; Insufficient outcome report
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

6

7 **Table 170 Summary of findings table for drug court versus active intervention for**
8 **substance misuse at post-treatment**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Drug court versus active intervention (95% CI)
Removed from treatment due to unsatisfactory progress	150 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.84 (0.38 to 1.86)	154 per 1000	25 fewer per 1000 (from 95 fewer to 132 more)
Addiction Severity Index (ASI): alcohol composite	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,4]		Control	MD 0.02 lower (0.04 to 0.00 lower)

score (Scale from 0 to 9; lower better)				mean 0.02	
Addiction Severity Index (ASI): drug composite score (Scale from 0 to 9; lower better)	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3]		Control mean 0.03	MD 0.01 lower (0.04 lower to 0.02 higher)
Number of sanctions at post-treatment	150 (1 RCT)	⊕⊕⊖⊖ LOW[1,4]		Control mean 4	MD 0.90 lower (1.99 lower to 0.19 higher)
Number of sanctions resulting in jail detention	121 (1 RCT)	⊕⊕⊖⊖ LOW[1,4]		Control mean 2.4	MD 0.5 lower (0.99 to 0.01 lower)
Re-incarceration	131 (1 RCT)	⊕⊖⊖⊖ VERY LOW[5,6]	RR 0.78 (0.47 to 1.28)	368 per 1000	81 fewer per 1000 (from 195 fewer to 103 more)
Urine test positive for drugs	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,6]	RR 0.4 (0.08 to 1.91)	161 per 1000	97 fewer per 1000 (from 148 fewer to 147 more)

[1] Messina 2012 - Inappropriate randomisation with adequate allocation concealment; No blinding; low risk of attrition bias; appropriate selective outcomes
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[3] Dakof 2010 - Unclear randomisation and allocation concealment; No blinding; ITT analysis; insufficient outcome report
[4] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.
[5] Jones 2013 - Permuted block randomisation with unclear allocation concealment; No blinding; low risk of attrition bias; insufficient outcome report
[6] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

1

7.2.2022 Case Management and Opioid Substitution Therapy

2 Two RCTs (N=301) met the eligibility criteria for this review: Gorden2008/Kinlock2007/
3 Kinlock2009(Gordon et al., 2008; Kinlock et al., 2009; Kinlock et al., 2007) and McKenzie
4 2012(McKenzie et al., 2012). The two studies evaluated opioid substitution therapy, namely
5 methadone, in addition to case management. The case management in the two studies were
6 not the same. The case management in Gorden2008 group was counselling with financial
7 assistance with or without transfer whereas that in McKenzie 2012 was counselling with
8 transfer with or without financial assistance.

9 An overview of the trials included in the meta-analysis can be found in Table 171. Further
10 information about both included and excluded studies can be found in Appendix L.

11 Summary of findings can be found in Table 172. The full GRADE evidence profiles and
12 associated forest plots can be found in Appendices N and O, respectively.

13 No data were available for mental health outcomes.

Table 171: Study information table for trials included in the meta-analysis of opioid substitution therapy plus case management vs case management only for drug misuse disorders

	Opioid substitution therapy plus case management vs case management only
Total no. of studies (N[1])	2 (301)
Study ID	(1) Gorden2008/Kinlock2007/Kinlock2009 (2) McKenzie2012
Study design	RCT
Country	(1, 2) USA
Underlying Mental Health Disorders	(1, 2) Drug misuse
Diagnosis	(1) Clinical (2) Unclear
Age (mean/range) years	(1)40.3 (2)40.7
Gender (% female)	(1)0 (2)29
Ethnicity (% white)	(1)16 (2)73
Criminal justice setting	(1, 2) Initiated in prison and continued in the community
Treatment length (weeks)	(1)12 (2) Mean – 2.1 weeks (15 days)
Follow-up length (weeks)	(1)16 (2) Mean – 28.1 weeks
Intervention (mean dose; mg/day)	(1) Case management (Counselling plus methadone with financial assistance) - Counselling fixed at once per week. Methadone started at 5mg every 8th day to a maximum of 60 mg per day (2) Case management (Counselling plus methadone with financial assistance) - Counselling fixed at one of the session. Methadone started on 5mg per day and increased by 2mg daily until release or they reached their individualised target dose. Average dose prior to release was 33mg/day.
Comparison	(1) Case management (Counselling plus financial assistance with or without transfer) (2) Case management (Counselling plus transfer with or without financial assistance)

Table 172 Summary of findings table for opioid substitution therapy plus case management versus case management only for substance misuse disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with case management	Risk difference with opioid substitution therapy plus case management versus case management only (95% CI)
Completed jail treatment - Total	211 (1 RCT)	⊕⊕⊖⊖ LOW[1]	RR 0.96 (0.81 to 1.14)	636 per 1000	25 fewer per 1000 (from 121 fewer to 89 more)
Completed jail treatment -	63	⊕⊖⊖⊖	RR 0.97	871 per 1000	26 fewer per 1000

Female sample	(1 RCT)	VERY LOW[1,2]	(0.79 to 1.18)		(from 183 fewer to 157 more)
Completed jail treatment - Male sample	148 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.95 (0.7 to 1.29)	539 per 1000	27 fewer per 1000 (from 162 fewer to 156 more)
Urine test positive for cocaine - 1 month follow-up	200 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.79 (0.58 to 1.07)	562 per 1000	118 fewer per 1000 (from 236 fewer to 39 more)
Urine test positive for cocaine - 6 month follow-up	76 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2]	RR 0.9 (0.62 to 1.31)	667 per 1000	67 fewer per 1000 (from 253 fewer to 207 more)
Urine test positive for cocaine - 12 month follow-up	115 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2]	RR 0.63 (0.43 to 0.91)	690 per 1000	255 fewer per 1000 (from 62 fewer to 393 fewer)
Urine test positive for opioids - 1 month follow-up	200 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.53 (0.35 to 0.8)	515 per 1000	242 fewer per 1000 (from 103 fewer to 335 fewer)
Urine test positive for opioids - 6 month follow-up	57 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2]	RR 0.43 (0.16 to 1.19)	578 per 1000	329 fewer per 1000 (from 485 fewer to 110 more)
Urine test positive for opioids - 12 month follow-up	115 (1 RCT)	⊕⊕⊖⊖ LOW	RR 0.44 (0.26 to 0.77)	563 per 1000	315 fewer per 1000 (from 130 fewer to 417 fewer)
Days of substance use (12 month follow-up) - Cocaine	204 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,4]		Control mean 64.6	MD 27.40 lower (47.25 to 7.55 lower)
Days of substance use (12 month follow-up) - Heroin	204 (1 RCT)	⊕⊖⊖⊖ VERY LOW[3,4]		Control mean 143	MD 36.80 lower (74.30 lower to 0.70 higher)
Self-reported drug use in past 30 days (6 month follow-up) - Crack/Cocaine	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 0.41 (0.16 to 1.05)	463 per 1000	273 fewer per 1000 (from 389 fewer to 23 more)
Self-reported drug use in past 30 days (6 month follow-up) - Heroin	62 (1 RCT)	⊕⊕⊖⊖ LOW[5]	RR 0.27 (0.09 to 0.79)	537 per 1000	392 fewer per 1000 (from 113 fewer to 488 fewer)
Self-reported drug use in past 30 days (6 month follow-up) - Marijuana	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 0.43 (0.1 to 1.83)	220 per 1000	125 fewer per 1000 (from 198 fewer to 182 more)
Self-reported drug use in past 30 days (6 month follow-up) - Injection drug use	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 0.26 (0.07 to 1.03)	366 per 1000	271 fewer per 1000 (from 340 fewer to 11 more)
Drug overdose - 6 month follow-up	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 0.84 (0.24 to 2.91)	171 per 1000	27 fewer per 1000 (from 130 fewer to 326 more)
Drug overdose - 12 month follow-up	204 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.14 (0.01 to 2.51)	45 per 1000	39 fewer per 1000 (from 45 fewer to 68 more)
Rearrest - 6 month follow-up	62 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,5]	RR 1.24 (0.56 to 2.73)	268 per 1000	64 more per 1000 (from 118 fewer to 464 more)
Rearrest - 12 month follow-up	204 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.96 (0.74 to 1.25)	556 per 1000	22 fewer per 1000 (from 145 fewer to 139 more)
Self-reported days of criminal activity (12 months follow-up)	204 (1 RCT)	⊕⊕⊖⊖ LOW[3]		Control mean 85.17	MD 3.37 lower (35.27 lower to 28.53 higher)

[1] *Gordon 2014 - Permuted blocks with adequate allocation concealment, No blinding with potential of effect size bigger in*

intervention group; available case analysis; appropriate outcome report
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[3] Kinlock 2007/Kinlock 2009/ Gordon 2008 - Permuted block randomisation with unclear allocation concealment; No blinding; ITT analysis with incomparable drop-out rates
[4] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.
[5] McKenzie 2012 - Unclear randomisation and allocation concealment; No blinding with potential increased effect size in intervention arm; per protocol analysis; appropriate outcome report

7.2.2013 Automated Telephony

One RCT (N=108) met the eligibility criteria for this review: Andersson 2014(Andersson et al., 2014). In this study, paroled offenders under supervision of assigned paroled officers were contacted by the central computer programmed to monitor acute dynamic risk factors daily during 30 consecutive days following probation. The group in either arm had been assessed but the intervention group received daily feedback with recommendations and a daily report to their correctional officers about their progress.

An overview of the trials included in the analysis can be found in

Table 173. Further information about both included and excluded studies can be found in Appendix L.

The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcome of service utilization.

Table 173: Study information table for trials included in the analysis of Automated Telephony

	Automated Telephony with feedback vs Automated Telephony Alone
Total no. of studies (N)	1(112)
Study ID	Andersson 2014
Study design	RCT
Country	Sweden
Underlying Mental Health Disorders	Paroled offenders under supervision of assigned paroled officers
Diagnosis	Not reported
Age (mean/range) years	36.2
Gender (% female)	2.8
Ethnicity (% white)	Not reported
Criminal justice setting	In the community
Treatment length (weeks)	4.25
Follow-up length (weeks)	Not reported
Intervention	Daily automated telephony assessment with feedback (7 hours/week)

	Automated Telephony with feedback vs Automated Telephony Alone
(mean dose; mg/day)	
Comparison	Daily automated telephony assessment

Table 174 Summary of findings table for automated telephony with feedback compared with automated telephony alone for mental health disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Automated telephony alone	Risk difference with Automated telephony with feedback (95% CI)
Change in Arnetz and Hasson stress questionnaire (AHSS) (Scale from 0 to 63; higher better)	108 (1 RCT)	⨁⨁⨁⨀ MODERATE[1]		Control mean 0.7	MD 2.5 higher (1.13 lower to 6.13 higher)
Change in symptom checklist-8D (SCL-8D) (Scale from 0 to 8; lower better)	108 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]		Control mean (-)1.2	MD 4.5 higher (0.22 to 8.78 higher)
Change in daily stressor assessment (Scale from 0 to 9; higher better)	108 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]		Control mean (-)0.01	MD 1.91 higher (1.11 to 2.71 higher)
Alcohol Urge Questionnaires: reduction in alcohol urge (Scale from 0 to 9; higher better)	108 (1 RCT)	⨁⨁⨁⨀ MODERATE[1]		Control mean 0.1	MD 0.2 higher (0.35 lower to 0.75 higher)
Alcohol Urge Questionnaires: reduction in alcohol use (Scale from 0 to 9; higher better)	108 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]		Control mean 0.1	MD 0.8 higher (0.11 to 1.49 higher)
Alcohol Urge Questionnaires: reduction in drug use (Scale from 0 to 9; higher better)	108 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]		Control mean (-)0.1	MD 1 higher (0.41 to 1.59 higher)

Alcohol Urge Questionnaires: reduction in drug urge (Scale from 0 to 9; higher better)	108 (1 RCT)	⊕⊕⊖⊖ LOW[1,2]	Control mean (-)0.1	MD 0.3 higher (0.25 lower to 0.85 higher)

[1] Andersson 2014 - Unclear randomisation with unclear allocation concealment; No blinding; Low drop-out rate with available rate analysis
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

7.2.2014 Integrated Disorders Treatment Program (IDDT)

One RCT (N=182) met the eligibility criteria for this review: Chandler 2006(Chandler & Spicer, 2006). IDDT service was a multidisciplinary team which included integrated substance abuse specialist and used stage-wise interventions. It helped IDDT clients get access to comprehensive service and time unlimited outreach services. The interventions included were motivational interventions, substance abuse counselling, group treatment oriented to both disorders (substance misuse and mental health disorders), family psychoeducation regarding dual disorders, participations in substance abuse self-help group, appropriate pharmacological treatment, interventions to promote health and secondary interventions for treatment non-responders. The in-custody care was provided to all participants (including those in IDDT intervention group) and included intensive assessment, medications, treatment planning before discharge, consultation to jail staff, one-on-one consoling and crisis intervention. The post-custody care in treatment as usual (TAU) which provided to all participants included 'usual services', available up to 60 days of post-release case management and housing assistance. Usual services included referral to one of the country-operated service teams for case management and medications.

An overview of the trials included in the analysis can be found in Table 175. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 176. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the mental health outcomes and re-offending rate.

Table 175: Study information table for trials included in the analysis of IDDT versus TAU

	IDDT vs TAU
Total no. of studies (N)	1(182)
Study ID	Chandler 2006
Study design	RCT
Country	USA
Underlying Mental Health Disorders	Severe mental illness and substance misuse
Diagnosis	Clinical
Age (range)years	36 to 50

	IDDT vs TAU
Gender (% female)	71.8
Ethnicity (% white)	21.2
Criminal justice setting	Post custody care
Treatment length (weeks)	130
Follow-up length (weeks)	Not reported
Intervention (mean dose; mg/day)	Integrated Disorders Treatment Program (IDDT) – The duration of the intervention differed from one participant to another according to the time of entry to program up until the completion (2.5 years)
Comparison	TAU, post-custody care included 'usual services' and the availability of up to 60 days of post-release case management and housing assistance. Usual services included referral to one of the county-operated service teams for case management and medications.
Notes. N=Total number of participants; TAU=Treatment as usual	

Table 176 Summary of findings table for IDDT versus TAU for mental health disorders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with IDDT versus TAU (95% CI)
Rate of outpatient medication services	182 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]	RR 1.25 (1.03 to 1.51)	646 per 1000	161 more per 1000 (from 19 more to 329 more)
Number of days in hospital	182 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]	-	Control mean 12.52 days	MD 5.63 lower (9.59 to 1.67 lower)
Rate of crisis visits	182 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]	-	Control mean 2.74	MD 2.26 lower (3.82 to 0.7 lower)

[1] Chandler 2006 - Unclear randomization with unclear allocation concealment; Blinding was not reported; Analysis by imputation
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25. [3]
The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

7.2.2055 Housing First

One RCT (N=297) met the eligibility criteria for this review: Somer 2013(Somers et al., 2013). In scattered housing first model intervention, subjects were dispersed in market accommodation and were also served by assertive community treatment (ACT) team which services included psychiatry and primary health care, and social and vocational rehabilitation (24/7). In congregate housing first model intervention, subjects were supported in a single building and were provided with on-site supports intended to match the overall intensity and composition of ACT (e.g. multi-professional, available 24/7). Moreover, the congregate model

highlighted the promotion of community through activities such as on-site recreation (e.g. street hockey). Treatment as usual (TAU) consisted of the existing and generally available services and support for individuals experiencing homelessness and mental illness. These included emergency shelters, housing units with varying levels of support and various health and social service providers.

An overview of the trials included in the analysis can be found in Table 177. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 178. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for mental health outcomes.

Table 177: Study information table for trials included in the analysis of housing first versus TAU

	Housing first versus TAU
Total no. of studies (N)	1(297)
Study ID	Somers 2013*
Study design	RCT
Country	Canada
Underlying Mental Health Disorders	Current mental disorder assessed on the MINI International Neuropsychiatric Interview (MINI)
Diagnosis	Clinical
Age (range)years	40
Gender (% female)	26
Ethnicity (% white)	57
Criminal justice setting	In the community
Treatment length (weeks)	104
Follow-up length (weeks)	Not reported
Intervention (mean dose; mg/day)	(1) Scattered Site Housing First (services available 24/7) + Assertive Community Treatment (Not reported) (2) Congregate Housing First (services available 24/7) (3) Treatment as usual (Not reported)
Comparison	TAU: existing and generally available services and support for individuals experiencing homelessness and mental illness
*3-armed study	

Table 178 Summary of findings table for housing first program (scattered HF or Congregate HF) versus treatment as usual for mental health disorders

Outcomes	**No of Participants (studies) Follow up**	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with TAU	**Risk difference with Housing First versus TAU (95% CI)**
Any offence	297 (1 RCT)	⨁⨁⊖⊖ LOW[1,2]	RR 0.43 (0.23 to 0.82)	190 per 1000	108 fewer per 1000 (from 34 fewer to 146 fewer)
Any offence - Scattered HF+ACT	140 (1 RCT)	⨁⨁⨁⊖ MODERATE[1]	RR 0.3 (0.12 to 0.77)	220 per 1000	154 fewer per 1000 (from 51 fewer to 194 fewer)
Any offence - Congregate HF	157 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.58 (0.25 to 1.39)	160 per 1000	67 fewer per 1000 (from 120 fewer to 62 more)

[1] Somers 2013 - Unclear randomisation with unclear concealment; no blinding of participants and care administrators; ITT analysis
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

7.2.2026 Texas Implementation of Medication Algorithm

One RCT (N=60) met the eligibility criteria for this review: Ehret 2013(Ehret et al., 2013). Texas Implementation of Medication Algorithm (TIMA) for Bipolar disorder was a treatment guideline consisting of both treatment strategies and treatment tactics.

An overview of the trials included in the analysis can be found in Table 179. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 180. The full GRADE evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

No data were available for the outcomes of re-offending rate and service utilization.

Table 179: Study information table for trials included in the analysis of Texas Implementation of Medication Algorithm (TIMA) versus TAU for Bipolar disorders

	TIMA vs TAU
Total no. of studies (N)	1 (60)
Study ID	Ehret 2013
Study design	RCT
Country	USA
Underlying Mental Health Disorders	Bipolar disorders
Diagnosis	Diagnosis
Age (range)years	32.7
Gender (% female)	100
Ethnicity (% white)	74
Criminal justice setting	prison
Treatment length (weeks)	24
Follow-up length (weeks)	Not reported
Intervention (mean dose; mg/day)	Texas Implementation of Medication Algorithm (TIMA)
Comparison	TAU (not specified)

Table 180 Summary of findings table for TIMA versus TAU for bipolar disorders

Outcomes	**No of Participants (studies)** Follow up	**Quality of the evidence** (GRADE)	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with TAU	**Risk difference with TIMA** versus TAU(95% CI)
Bipolar Disorder Symptom Scale (BDSS) (Scale from 7 to 70; lower better)	60 (1 RCT)	⨁⨁⨀⨀ LOW[1,2]		Control mean 1.49	MD 0.27 lower (0.75 lower to 0.21 higher)

Brief Psychiatric Rating Scale (BPRS) (Scale from 18 to 126; lower better)	60 (1 RCT)	⨁⨁⊖⊖ LOW[1,2]	Control mean 28.51	MD 0.97 higher (1.78 lower to 3.72 higher)

[1] Ehret 2013 - inappropriate randomization with unclear concealment; no blinding; available case analysis
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

1

7.2.2027 Service Brokerage Intervention

3 One RCT (N = 1325) met the eligibility criteria for this review: Kinner 2013/2014a/2014b
4 (Cutcher et al., 2014b; Kinner et al., 2014a; Kinner et al., 2013). The intervention included
5 the provision of documents tailed to each participant at release with post-release telephone
6 support.

7 An overview of the trials included in the analysis can be found in Table 181. Further
8 information about both included and excluded studies can be found in Appendix L.

9 Summary of findings can be found in Table 182. The full GRADE evidence profiles and
10 associated forest plots can be found in Appendices N and O, respectively.

11 No data were available for the mental health outcomes and re-offending rate.

12 **Table 181: Study information table for trials included in the analysis of service**
13 **brokerage intervention for substance misuse disorders**

	Service brokerage intervention vs TAU
Total no. of studies (N)	1 (1325)
Study ID	Kinner 2013/2014a/2014b
Study design	RCT
Country	Australia
Underlying Mental Health Disorders	Substance misuse disorders
Diagnosis	Unclear
Age (range)years	17-89
Gender (% female)	21.1
Ethnicity (% white)	Not reported
Criminal justice setting	Initiated in prison and continued in the community
Treatment length (weeks)	Not reported
Follow-up length (weeks)	26
Intervention (mean dose; mg/day)	Service brokerage model (a total of 660 hours)
Comparison	Treatment as usual (not specified)

14

15 **Table 182 Summary of findings table for service brokerage intervention versus TAU**
16 **for substance misuse disorders**

Outcomes	**No of Participants (studies)**	**Quality of the evidence** (GRADE)	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk	**Risk difference with Service**

	Follow up			with TAU	brokerage intervention versus TAU (95% CI)
Number of participants in contact with MH service	1325 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 1.16 (0.8 to 1.69)	71 per 1000	11 more per 1000 (from 14 fewer to 49 more)
Number of participants who have seen GP	1325 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 1.6 (0.81 to 3.17)	20 per 1000	12 more per 1000 (from 4 fewer to 43 more)
Number of participants who attended alcohol or drug service	1325 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 1.05 (0.55 to 2.02)	26 per 1000	1 more per 1000 (from 12 fewer to 26 more)

[1] Kinner 2013/2014a/2014b - RCT with unclear allocation concealment; Blinding of care administrators; ITT analysis
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

7.2.20.8 Therapeutic communities for substance misuse

8 RCTs met the eligibility criteria for this review: Czuchry 2003, Messina 2010, Sacks 2004, Sacks 2008, Sacks 2012a, Sacks 2012b, Sullivan 2007 and Wexler 1999(Czuchry & Dansereau, 2003; Messina et al., 2010; Sacks et al., 2012a; Sacks et al., 2008; Sacks et al., 2012b; Sacks et al., 2004; Sullivan et al., 2007; Wexler et al., 1999).

Therapeutic communities are structured, therapeutic environments. They are employed in substance misuse treatment in an effort to improve the likelihood of long-term outcomes, by embedding new skills and ways of living into everyday life.

7.2.20.8.1 Therapeutic community versus waitlist for substance misuse

1 RCT (N=715) met the eligibility criteria for this review: Wexler 1999(Wexler et al., 1999).

An overview of the trials can be found in Table 183. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 184. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

This was a 2-armed trial with service users randomly allocated either to a therapeutic community or a waitlist control condition.

The evidence for this review was low to moderate quality. No data was available for the outcomes of mental health, service utilisation, adaptive functioning or rates of self-injury.

Table 183: Study characteristics for the comparison of therapeutic communities versus waitlist for substance misuse

	Therapeutic community versus waitlist
Total no. of studies (N[1])	1 (715)
Study ID	Wexler 1999
Study design	RCT
Country	USA
Diagnosis	Drug misuse
Age (mean)	30.0 years
Sex (% female)	NR
Ethnicity (% white)	37.0%

	Therapeutic community versus waitlist
Setting	Prison
Coexisting conditions/other treatments received during study	NR
Treatment length (weeks)	35-52 weeks
Intervention (mean dose; mg/day)	Therapeutic community
Delivery method	Individual and group
Comparison	Waitlist control
Notes. N=Total number of participants; NR =Not reported [1] Number randomised	

Table 184: Summary of findings table for the comparison of therapeutic communities versus waitlist control for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with waitlist control	Risk difference with Therapeutic community versus waitlist control (95% CI)
Days until reincarceration	341 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2]		Control mean 294.98	MD 83.58 higher (32.69 to 134.47 higher)

[1] Wexler 1999 - Unclear randomisation and allocation concealment; No blinding with potential of effect size bigger in intervention group; ITT analysis; appropriate outcome report
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

7.2.20.852 Modified therapeutic community versus CBT informed psychoeducation for substance misuse

2 RCTs (N=375) met the eligibility criteria for this review: Sacks 2004 and Sullivan 2007(Sacks et al., 2004; Sullivan et al., 2007).

An overview of the trials can be found in Table 185. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 186. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

These were both 2-armed trials with service users randomly allocates to either a modified therapeutic community or CBT-based psychoeducational programme. The modifications made to the therapeutic community model included an emphasis on criminal thinking and behaviour, adjustments to comply with security guidelines and inclusion of security personnel on the treating team.

The evidence for this review was of very low quality. No data was available for the outcomes of service utilisation, adaptive functioning or rates of self-injury.

Table 185: Study characteristics table for the comparison of modified therapeutic communities versus active intervention for substance misuse

	Modified therapeutic community (MTC) versus active intervention
Total no. of studies (N[1])	2 (375)
Study ID	(1) Sacks 2004* (2) Sullivan 2007
Study design	RCT
Country	(1, 2) USA
Diagnosis	(1) SMI plus substance misuse (2) Substance misuse
Age (mean)	(1, 2) 34.3
Sex (% female)	(1, 2) 0.0
Ethnicity (% white)	(1, 2) 49.0
Setting	(1, 2) Prison
Coexisting conditions/other treatments received during study	(1, 2) NR
Treatment length (weeks)	(1) NR (2) 52 weeks
Intervention (mean dose; mg/day)	(1) Prison MTC with or without aftercare (NR) (2) MTC (20-25 hours per week)
Delivery method	(1, 2) Group
Comparison	(1) Mental health (MH) program (2) CBT-informed psychoeducation
Notes. N=Number of participants; NR=Not reported [1] Number randomised;*3-armed study	

Table 186: Summary of findings for the comparison of modified therapeutic communities versus active intervention for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Modified therapeutic community versus CBT informed psychoeducation (95% CI)
Substance use (12 month follow-up)	139 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.56 (0.37 to 0.84)	547 per 1000	241 fewer per 1000 (from 88 fewer to 345 fewer)
Alcohol use (12 month follow-up)	139 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.53 (0.31 to 0.93)	375 per 1000	176 fewer per 1000 (from 26 fewer to 259 fewer)
Drug use (12 month follow-up)	139 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.55 (0.34 to 0.89)	438 per 1000	197 fewer per 1000 (from 48 fewer to 289 fewer)
Criminal activity (12 month follow-up)	139 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.66 (0.5 to 0.89)	703 per 1000	239 fewer per 1000 (from 77 fewer to 352 fewer)

Re-incarceration (12 month follow-up)	139 (1 RCT)	⊕⊕⊖⊖ LOW[3]	RR 0.28 (0.13 to 0.63)	328 per 1000	236 fewer per 1000 (from 121 fewer to 285 fewer)
Alcohol/drug offence (12 month follow-up)	139 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.62 (0.43 to 0.9)	578 per 1000	220 fewer per 1000 (from 58 fewer to 330 fewer)

[1] Sullivan 2007 - unclear randomisation and allocation concealment; No blinding; unclear analysis; self-reported data
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[3] Sacks 2004 - Unclear randomisation and allocation concealment; Unclear blinding; Available case analysis; inadequate outcome report

7.2.20.823 Enhanced therapeutic community versus standard therapeutic community for substance misuse

1 RCT (N=452) met the eligibility criteria for this review: Czuchry 2003(Czuchry & Dansereau, 2003).

An overview of the trial can be found in Table 187. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 188. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

This was a 2-armed trial with service users allocated either to a cognitive-skills enhanced therapeutic community or a standard therapeutic community. The enhanced condition received motivational interventions and participated in node-link mapping (a counselling technique) in addition to the standard care.

The evidence for this review was of low to very low quality. No data were available for the outcomes of offending and reoffending, adaptive functioning or rates of self-injury.

Table 187: Study characteristics table for the comparison of enhanced therapeutic communities versus standard therapeutic communities for substance misuse

	Enhanced therapeutic community versus standard therapeutic community
Total no. of studies (N[1])	1 (452)
Study ID	Czuchry 2003
Study design	RCT
Country	USA
Diagnosis	Drug misuse
Age (mean)	29.9 years
Sex (% female)	31.0%
Ethnicity (% white)	58.0%
Setting	Initiated in prison and continued in the community
Coexisting conditions/other treatments received during study	n/r
Treatment length (weeks)	30 weeks

	Enhanced therapeutic community versus standard therapeutic community
Intervention (mean dose; mg/day)	Enhanced therapeutic community
Delivery method	Group of <=35 people
Comparison	Standard therapeutic community
Notes. N=Total number of participants [1] Number randomised	

Table 188: Summary of findings table for the comparison of enhanced therapeutic community versus standard therapeutic community for substance misuse at post-treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with standard therapeutic community	Risk difference with Enhanced therapeutic community versus standard therapeutic community (95% CI)
Engagement with treatment	451 (1 RCT)	⨁⨁⨀⨀ LOW[1]		Control mean 0.61	MD 0.03 higher (0.01 lower to 0.07 higher)
Negative mood (as rated by counsellor) (Scale from 2 to 14; lower better)	449 (1 RCT)	⨁⨀⨀⨀ VERY LOW[1,2]		Control mean 4.46	MD 1.79 lower (2.09 to 1.49 lower)

[1] Czuchry 2003 – unclear randomisation and allocation concealment; no blinding; unclear attrition
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.

7.2.20.951 Gender-responsive therapeutic community versus standard therapeutic community for substance misuse

1 RCT (N=115) met the eligibility criteria for this review: Messina 2010(Messina et al., 2010).

An overview of the trial can be found in Table 189. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 190. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

This was a 2-armed trial with women randomly allocated either to a gender-responsive therapeutic community, where all staff facilitating the groups and counselling the women were female, or a standard therapeutic community where staff were either male or female.

The evidence for this review was very low quality. No data were available for the outcomes of adaptive functioning or quality of life.

Table 189: Study characteristics table for the comparison of gender-responsive therapeutic community versus standard therapeutic community

	Gender-responsive therapeutic community versus standard therapeutic community
Total no. of studies (N[1])	1 (115)
Study ID	Messina 2010
Study design	RCT
Country	USA
Diagnosis	Drug misuse
Age (mean)	36.1 years
Sex (% female)	100
Ethnicity (% white)	48.0
Setting	Prison
Coexisting conditions/other treatments received during study	NR
Treatment length (weeks)	NR
Intervention (mean dose; mg/day)	Gender-responsive therapeutic community
Delivery method	Group
Comparison	Standard therapeutic community
Notes. N=Total number of participants. NR=Not reported. [1] Number randomised	

Table 190: Summary of findings for the comparison of gender-responsive therapeutic community versus standard therapeutic community for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with control	Risk difference with Gender-responsive therapeutic community versus standard therapeutic community (95% CI)
Addiction Severity Index (ASI): alcohol composite score (Scale from 0 to 9; lower better)	115 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2,3]		Control mean 0.07	MD 0.04 lower (0.08 lower to 0 higher)
Addiction Severity Index (ASI): psychological composite score	115 (1 RCT)	⨁⨁⊖⊖ LOW[1,2]		Control mean 0.24	MD 0.01 lower (0.1 lower to 0.08 higher)
Addiction Severity Index (ASI): drug composite score (Scale from 0 to 9; lower better)	115 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2,3]		Control mean 0.02	MD 0.02 higher (0 to 0.04 higher)
Addiction Severity Index (ASI): family composite score	115 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2,3]		Control mean 0.14	MD 0.04 lower (0.12 lower to 0.04 higher)
Participated in aftercare	115	⨁⊖⊖⊖	RR 0.86	545 per 1000	76 fewer per 1000

upon release	(1 RCT)	VERY LOW[1,2,4]	(0.6 to 1.23)		(from 218 fewer to 125 more)
Months spent in aftercare	115 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,3]		Control mean 3.4	MD 1.50 higher (0.29 to 2.71 higher)
Disciplinary removal from first residential treatment post-release	115 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,4]	RR 0.92 (0.37 to 2.28)	145 per 1000	12 fewer per 1000 (from 92 fewer to 186 more)
Re-incarceration (12 month follow-up)	115 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,4]	RR 0.66 (0.41 to 1.07)	455 per 1000	155 fewer per 1000 (from 268 fewer to 32 more)
Voluntarily dropped-out from first residential treatment post-release	115 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,4]	RR 0.54 (0.27 to 1.08)	309 per 1000	142 fewer per 1000 (from 226 fewer to 25 more)
Months until reincarceration	115 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2,3]		Control mean 5.9	MD 1.90 higher (0.5 to 3.3 higher)

[1] *Messina 2010 - high risk of selection bias; No blinding; available case analysis; unclear selective outcome report*
[2] *Evidence was downgraded by one level because study population of one study (Messina 2010) differed from the review question in that not all the participants met the proxy measure criteria for substance misuse disorder.*
[3] *Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or both boundaries of the defined minimally important difference (MID) for the outcome, respectively. For continuous outcomes, +/-0.5 (mean for 2 studies and median for 3 or more studies) times SD of the control group (if MD was used) were considered as MID boundaries.*
[4] *Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.*

7.2.20.912 Gender-specific therapeutic community versus psychoeducation for substance misuse

2 RCTs (N=782) met the eligibility criteria for this review: Sacks 2008 and Sacks 2012a (Sacks et al., 2012a; Sacks et al., 2008).

An overview of the trials can be found in Table 191. Further information about both included and excluded studies can be found in Appendix L.

Summary of findings can be found in Table 192. The full evidence profiles and associated forest plots can be found in Appendices N and O, respectively.

These were both 2-armed trials with service users randomised either to a gender-specific therapeutic community or to CBT-informed psychoeducation. All service users in these trials were female.

The evidence for this review was very low quality. No data was available for the outcomes of adaptive functioning or rates of self-injury.

Table 191: Study characteristics table for the comparison of gender-specific therapeutic communities versus psychoeducation for substance misuse

	Gender-specific therapeutic community versus psychoeducation
Total no. of studies (N[1])	2 (782)
Study ID	(1) Sacks 2008

	Gender-specific therapeutic community versus psychoeducation
	(2) Sacks 2012a
Study design	RCT
Country	(1, 2) USA
Diagnosis	(1, 2) Substance misuse
Age (mean)	(1) 35.6 years (2) 35.1 years
Sex (% female)	(1, 2) 100
Ethnicity (% white)	(1) 48.0 (2) 47.0
Setting	(1, 2) Prison
Coexisting conditions/other treatments received during study	(1, 2) Not reported
Treatment length (weeks)	(1) 28 weeks (2) 26 weeks
Intervention (mean dose; mg/day)	Gender specific therapeutic community: (1, 2) 40 hours per week
Delivery method	(1, 2) Individual and group
Comparison	CBT-informed psychoeducation; (1) Not reported (2) 6 hours per week
Notes. N=Total number of participants; [1] Number randomised	

1

2 **Table 192: Summary of findings for the comparison of gender-specific therapeutic**
3 **communities versus psychoeducation for substance misuse**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Gender-specific therapeutic community versus CBT informed psychoeducation (95% CI)
Beck Depression Inventory (BDI) total score at post-treatment (Scale from 0 to 63; lower better)	314 (1 RCT)	⨁⨁⨀⨀ LOW[1]		Control mean 14.48	MD 2.64 lower (5.26 to 0.02 lower)
Brief Symptom Inventory (BSI) total score at post-treatment (Scale from 0 to 212; lower better)	314 (1 RCT)	⨁⨁⨀⨀ LOW[1]		Control mean 55.1	MD 1.63 lower (4.45 lower to 1.19 higher)

Post-traumatic Symptom Scale (PSS) at post-treatment (Scale from 0 to 51; lower better)	314 (1 RCT)	⊕⊕⊖⊖ LOW[1]		Control mean 13.12	MD 2.90 lower (5.68 to 0.12 lower)
Self-reported criminal activity (any) - 6 month follow-up	702 (2 RCTs)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 0.77 (0.64 to 0.92)	454 per 1000	104 fewer per 1000 (from 36 fewer to 164 fewer)
Self-reported criminal activity (any) - 12 month follow-up	370 (1 study)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.85 (0.65 to 1.1)	411 per 1000	62 fewer per 1000 (from 144 fewer to 41 more)
Self-reported criminal activity (drugs) - 6 month follow-up	702 (2 RCTs)	⊕⊖⊖⊖ VERY LOW[1,2,3]	RR 0.86 (0.69 to 1.08)	329 per 1000	46 fewer per 1000 (from 102 fewer to 26 more)
Self-reported criminal activity (drugs) - 12 month follow-up	370 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.81 (0.61 to 1.09)	368 per 1000	70 fewer per 1000 (from 144 fewer to 33 more)
Self-reported criminal activity (sexual)	314 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.35 (0.09 to 1.29)	53 per 1000	34 fewer per 1000 (from 48 fewer to 15 more)
Receiving substance abuse treatment at follow-up	314 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.86 (0.75 to 0.98)	781 per 1000	109 fewer per 1000 (from 16 fewer to 195 fewer)
Receiving mental health treatment at follow-up	314 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.96 (0.73 to 1.25)	417 per 1000	17 fewer per 1000 (from 113 fewer to 104 more)
Alcohol use (follow-up NR)	314 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 1.31 (0.86 to 2)	192 per 1000	60 more per 1000 (from 27 fewer to 192 more)
Frequency of alcohol use (follow-up NR) (Scale from 0 to 8; lower better)	162 (1 RCT)	⊕⊕⊖⊖ LOW[1]		Control mean 0.97	MD 0.25 higher (0.42 lower to 0.92 higher)
Frequency of drug use (follow-up NR) (Scale from 0 to 8; lower better)	206 (1 RCT)	⊕⊕⊖⊖ LOW[1]		Control mean 1.51	MD 0.42 lower (1.14 lower to 0.30 higher)
Self-reported	702	⊕⊖⊖⊖	RR 0.77	265 per 1000	61 fewer per 1000

drug use - 6 month follow-up	(2 RCTs)	VERY LOW[1,2,3]	(0.59 to 1.01)		(from 109 fewer to 3 more)
Rearrest - 6 month follow-up	388 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.5 (0.29 to 0.85)	181 per 1000	90 fewer per 1000 (from 27 fewer to 128 fewer)
Rearrest - 12 month follow-up	370 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 1.65 (0.83 to 3.28)	67 per 1000	44 more per 1000 (from 11 fewer to 154 more)
Rearrest - Follow-up NR	314 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,2]	RR 0.73 (0.52 to 1.03)	351 per 1000	95 fewer per 1000 (from 168 fewer to 11 more)
Re-incarceration	468 (1 RCT)	⊕⊖⊖⊖ VERY LOW[2,3]	RR 0.82 (0.6 to 1.12)	280 per 1000	50 fewer per 1000 (from 112 fewer to 34 more)

[1] Sacks 2008 - unclear randomisation and allocation concealment; No blinding; analysis by regression technique; appropriate outcome report
[2] Evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or more boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.
[3] Sacks 2012a - unclear randomisation and allocation concealment; No blinding with potential of effect size bigger in intervention group; available case analysis

1

7.2.20.923 Re-entry modified therapeutic community versus treatment as usual

3 1 RCT (N=127) met the eligibility criteria for this review: Sacks 2012b(Sacks et al., 2012b).

4 An overview of the trials can be found in Table 193. Further information about both included
5 and excluded studies can be found in Appendix L.

6 Summary of findings can be found in Table 194. The full evidence profiles and associated
7 forest plots can be found in Appendices N and O, respectively.

8 This was a 2-armed trial with service users randomly allocated to either a re-entry modified
9 therapeutic community or treatment as usual, which consisted of parole supervision and case
10 management. The re-entry condition included components to address criminal thinking and
11 behaviour which were not provided to the TAU group.

12 The evidence for this review was of very low quality. No data was available for the outcomes
13 of mental health, service utilisation, adaptive functioning or rates of self-injury.

14 **Table 193: Study characteristics for the comparison of re-entry modified**
15 **therapeutic communities versus treatment as usual for substance misuse**

	Re-entry modified therapeutic community versus treatment as usual
Total no. of studies (N[1])	1 (127)
Study ID	Sacks 2012b
Study design	RCT
Country	USA
Diagnosis	SMI and substance misuse
Age (mean)	38.2 years

	Re-entry modified therapeutic community versus treatment as usual
Sex (% female)	0.0
Ethnicity (% white)	56.0
Setting	Community corrections facility
Coexisting conditions/other treatments received during study	NR
Treatment length (weeks)	26 weeks
Intervention (mean dose; mg/day)	Re-entry modified therapeutic community; 3-5 hours per day for 3-7 days
Delivery method	Individual and group
Comparison	TAU: Clinical supervisor conducted a weekly on-site group in relapse prevention, and case managers provided daily medication monitoring whereas community MH clinics supplied psychiatric and MH counselling services.
Note. N=Total number of participants; MH= mental health; NR=Not reported; SMI=Serious Mental Illness; TAU=Treatment as usual [1] Number randomised	

Table 194: Summary of findings table for the comparison of re-entry modified therapeutic communities versus treatment as usual for substance misuse

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Re-entry modified therapeutic community versus treatment as usual (95% CI)
Re-incarceration (12-month post prison release)	127 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.53 (0.29 to 0.94)	375 per 1000	176 fewer per 1000 (from 23 fewer to 266 fewer)
Criminal activity (12-months post prison release)	110 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.64 (0.44 to 0.94)	617 per 1000	222 fewer per 1000 (from 37 fewer to 346 fewer)
Alcohol/Drug offence (12-months post prison release)	110 (1 RCT)	⨁⊖⊖⊖ VERY LOW[1,2]	RR 0.64 (0.42 to 0.96)	574 per 1000	207 fewer per 1000 (from 23 fewer to 333 fewer)

[1] Sacks 2012b – inappropriate randomisation without allocation concealment; no blinding; ITT analysis; lack of outcome report on percentages of therapeutic community in prison
[2] The evidence was downgraded by one level and two levels if the confidence interval crossed or touched one or two boundaries of the defined minimally important difference (MID) for the outcome respectively. The MID boundaries for dichotomous outcomes (RR) were 0.8 to 1.25.

7.2.21 Economic evidence

The systematic search of the literature identified 18 studies that assessed the costs and benefits associated with the organisation and structure of services, for the assessment, intervention and management of mental health problems in people in contact with the criminal justice system. Of these:

- 6 studies (in 7 publications) examined the costs and benefits associated with jail diversion programmes in the UK, US and Canada (Hayhurst et al., 2015; Zarkin et al., 2015; Cowell et al., 2013; Hughes et al., 2012; Mitton et al., 2007; Steadman et al., 2005; Cowell et al., 2004) (Cowell et al., 2004; Cowell et al., 2013; Hayhurst et al., 2015; Hughes et al., 2012; Mitton et al., 2007; Steadman & Naples, 2005; Zarkin et al., 2015)
- 2 studies assessed the costs and benefits associated with mental health courts in the US (Kubiak et al., 2015; Ridgely et al., 2007)
- 4 studies examined the costs and benefits associated with drug court programmes in the US and Australia (Cheeseman et al., 2016; Carey et al., 2004; Logan et al., 2004; Shanahan et al., 2004) (Carey & Finigan, 2004; Cheesman et al., 2016; Logan et al., 2004; Shanahan et al., 2004)
- 1 study assessed the costs and benefits associated with street triage in the UK (Heslin et al., 2016a); and 1 study assessed the costs and benefits associated with street triage, providing mental health act assessment for all Section 136 detainees, and having a link worker present at custody sites in the UK (Heslin et al., 2016b)
- 1 study examined the costs and benefits associated with integrated dual disorders treatment in the US (Chandler & Spicer, 2006)
- 1 study examined the costs and benefits associated with the forensic assertive community treatment (FACT) in the US (Cusack et al., 2010)
- 3 studies examined the costs and benefits associated with prison-based therapeutic community and aftercare treatments in the US ((McCollister et al., 2003a; McCollister et al., 2003b; McCollister et al., 2004)
- 2 studies examined the costs and benefits associated with the probation and mandated treatment in the US (Alemi et al., 2006; Anglin et al., 2013)(Alemi et al., 2006; Anglin et al., 2013)
- 1 study examined the costs and benefits associated with inpatient medium security unit (MSU) and residential service in the UK for people with personality disorders (Fortune et al., 2011)(Fortune et al., 2011)

Details on the methods used for the systematic review of the economic literature are described in Chapter 3; full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix S. Completed methodology checklists of the studies are provided in Appendix R. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix T.

7.2.21.1 Jail diversion

7.2.21.2 Hayhurst and colleagues (2015)

Hayhurst and colleagues (2015) evaluated the cost-utility of diversion and aftercare programmes for adult opiate- and/or crack-using offenders who come into contact with the CJS using Class A drugs in the UK. The diversion and aftercare intervention was broadly defined to cover the variety of interventions provided in the UK. Standard care was defined as no formal diversion and aftercare programmes (that is, the usual CJS pathway of such offenders). This was a modelling study and decision tree was used to synthesize data.

The source of clinical effectiveness data included an observational study, other published studies, and assumptions. The perspective taken was that of public sector (healthcare, social care, and criminal justice). The study considered a range of costs including drug intervention programme, drug test, drug treatment, arrest, prison, costs associated with remaining in the community after arrest and conviction, and costs associated with the subsequent recorded offence. The resource use estimates were based on the observational study and other published sources. The unit costs were obtained from national sources and other published

studies. The measure of outcome for the economic analysis was the QALY with utility weights based on the SF-12/SF-6D questionnaire. The time horizon of the main analysis was 12 months. A 5- and 10- year time horizon was explored in the sensitivity analyses and discounting was applied at 3.5% as recommended by the NICE. Sensitivity analyses were used to explore the impact of varying the intensity and scope of the drug intervention programme on the probability of reoffending, costs and outcomes.

Diversion resulted in a greater number of QALYs when compared with standard care (0.655 versus 0.650, respectively; a difference of 0.005, 95% CI: -0.057 to 0.065). From a public sector perspective, the mean expected costs per person over 12 months were £14,404 for the diversion and £14,551 for standard care, a difference of -£147, 95% CI: –£17,573 to £16,317 (in 2012 prices). Based on the above findings the diversion was dominant when compared with standard care services. The cost effectiveness acceptability curve suggested that if decision-makers were willing to pay up to £30,000 to gain 1 additional QALY for arrested drug users to receive an intervention, there may be a 50% chance that diversion is cost-effective. The sensitivity analysis indicated that there was a substantial uncertainty. Under some sets of assumptions (that is when changing the assumptions pertaining to the population eligible for diversion) the cost per QALY was as high as £1,194,800.

The analysis was judged by the GC to be directly applicable to the NICE decision-making context. This was the UK-based study and the outcome measure was QALY (even though utility weights were based on the SF-12/SF-6D questionnaire). Overall, this was a well-conducted study and was judged by the GC to have only minor methodological limitations.

7.2.21.122 Zarkin and colleagues (2015)

Zarkin and colleagues (2015) examined the costs associated with a jail diversion programme in the US. The authors examined costs associated with 2 hypothetical policy scenarios. In Scenario 1 diversion eligible offenders had a 10% probability of being diverted from incarceration to treatment in the community and in Scenario 2 this probability was increased to 40%. The scenarios were compared with standard care defined as no diversion from prison or jail into the community. The study population comprised adult offenders with substance abuse problems. This was a modelling study (a discrete event simulation) with effectiveness data being taken from various published sources. The analysis was conducted from a public sector perspective (healthcare and criminal justice sectors). The study considered a range of cost categories including costs associated with crime victimisation, arrest, court, and incarceration; and health care services. The resource use estimates were based on published sources. The unit costs were obtained from national sources and published studies. The benefits associated with the programme were expressed in terms of earnings potential. The net benefit was calculated as the sum of the difference in the expected costs and benefits. The time horizon of the analysis was over lifetime. A discount rate of 3% was applied to both costs and outcomes.

Both modelled scenarios resulted in greater benefits (that is, earnings) when compared with standard care. The benefits associated with scenario 1 were $103,509, with scenario 2 $107,018, and with standard care $101,754; the difference in benefits was $1,754 between scenario 1 and standard care and $5,263 between scenario 2 and standard care (likely 2014 prices). The mean lifetime costs per person were $303,509 for scenario 1, $294,737 for scenario 2, and $308,772 for standard care; the difference in costs was -$5,263 between scenario 1 and standard care and -$14,035 between scenario 2 and standard care. The net savings per person over the lifetime from a public sector perspective were ~$7,018 and ~$19,298 associated with scenario 1 and scenario 2, respectively; in both cases the level of statistical significance was $p<0.01$. Under one-way sensitivity analyses results changed little and conclusions were robust.

The analysis was judged by the GC to be partially applicable to the NICE decision-making context since the study was conducted in the US. Overall, given the data limitations in this

area, this was a well-conducted study and was judged by the GC to have only minor methodological limitations.

7.2.21.133 Cowell and colleagues (2013)

Cowell and colleagues (2013) examined the costs associated with a pre-booking component of a jail diversion programme for adults with indications of serious mental illness (including major depression, bipolar disorder, schizophrenia, or schizoaffective disorder) in the US (Bexar County, Texas). The jail diversion programme was compared with no diversion alternative. The cost analysis was based on an observational case-control study (N=468). Clinical effectiveness data were derived from an observational study and various interlinked administrative databases. The time horizon of the economic analysis was 2 years, and its perspective was public sector, including healthcare and criminal justice sector costs. Cost elements comprised arrests, court, incarcerations, diversion, and treatment. Cost data were derived from observational study participants and various interlinked administrative databases, published and unpublished studies, and billing records. Regression analysis was used to adjust the cost differences for baseline differences in participant characteristics including race, living arrangements, education, time at risk, gender, marital status, and age.

According to the analysis the mean costs over 2 years per participant were $8,247 (SE $1,037) and $15,147 (SE $646) for diverted and non-diverted participants, respectively (in 2007 US dollars); the unadjusted difference was -$6,901 (SE $1,253), $p < 0.01$; and the adjusted difference was -$2,819 (SE $824), $p < 0.01$. Based on these results, the authors concluded that jail diversion for people with serious mental illness may be justified fiscally.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged by the GC to have potentially serious methodological limitations, including the short time horizon (2 years), the study design (observational case-control study), and the source of unit cost data was unclear.

7.2.21.134 Cowell and colleagues (2004)

In another US study, Cowell and colleagues (2004) and Steadman and colleagues (2005) evaluated the cost effectiveness of 4 jail diversion programmes in the US (Lane County, Oregon; Memphis, Tennessee; New York City; and Tucson, Arizona). The programmes in Lane County, New York City, and Tucson were post-booking, and in Memphis the programme was pre-booking. Jail diversion programmes at all four sites were compared with no diversion alternative. The economic analysis was based on an observational cohort study. Clinical effectiveness data were obtained from the study participants: Lane County (N=185), Memphis (N=609), New York (N=231), and Tucson (N=90). The time horizon of the economic analysis was up to 1 year, and its perspective was public sector, including healthcare, social care, and criminal justice sector costs. Cost elements comprised criminal justice (court, public defenders' and prosecutors' offices, police, and jail) and healthcare (mental health, residential substance abuse care, outpatient care [both substance abuse and mental health], emergency room [for substance abuse and mental health visits], mental health assessment or evaluation, and case management). Resource use data were also obtained from the study participants: Lane County (N=129), Memphis (N=609), New York (N=231), and Tucson (N=90). Unit cost data were obtained from key study stakeholders. Where necessary resource use data and unit cost data were supplemented with information from published studies and information from other sites where jail diversion programmes have already been implemented. The study used a number of outcome measures including: criminal behaviour (whether the person was arrested in the previous 30 days), service user quality of life (whether the respondent had been violently and/or non-violently victimised in the past 3 months), housing status stability, level of physical and mental health [as measured using the 12-Item Short Form Health Survey (SF-12) and Colorado Symptom Index (CSI) questionnaire], and substance use (whether the respondent abused alcohol/drugs at any time during the past 3 months). The costs were reported at 1 year, and outcomes were reported at 3 months and 1 year. The results were reported for each site separately.

Regression analysis was used to adjust the cost differences for baseline differences in participant characteristics including age, gender, race or ethnicity, whether the individual was mentally disturbed at baseline, whether the respondent was ever arrested as a juvenile, number of past arrests, and the severity of alcohol and drug use.

According to the analysis in Lane County the mean annual costs were $16,164 (SD $13,245) and $15,743 (SD $17,498), per diverted and non-diverted participant, respectively; the adjusted difference was $1,796 (SD $3,492), p = ns (in 1996 US dollars); in Memphis $8,740 (SD $14,911) and $3,685 (SD $8,352) per diverted and non-diverted participant, respectively; the adjusted difference was $5,855 (SD $1,158), p ≤ 0.001; in New York $13,366 (SD $17,114) and $18,480 (SD $17,629) per diverted and non-diverted participant, respectively; the adjusted difference was -$6,260 (SD $2,594), p ≤ 0.05; and in Tucson $11,976 (SD $15,048) and $11,119 (SD $2,155) per diverted and non-diverted participant, respectively; the adjusted difference was $447 (SD $3,551), p = ns.

According to the analysis at 3-months in Lane County the intervention resulted in an increase in the odds (OR) of being arrested (OR 3.24, p ≤ 0.1) and being non-violently victimised (OR 3.81, p ≤ 0.1); at 12-months the intervention resulted in the reduction in the odds of substance abuse (OR 0.21, p ≤ 0.05). In Memphis at 3-months the intervention resulted in an improvement on the CSI scale, p ≤ 0.1; and at 12-months there were no significant changes. In New York at 3-months there was a reduction in the odds of being seriously victimised (OR 0.37, p ≤ 0.1) and non-violently victimized (OR 0.27, p ≤ 0.05); and at 12-months, there were no significant changes. In Tucson at 3-months there was an increase in the odds of being non-violently victimized (OR 5.01, p ≤ 0.1) and an improvement in the CSI score, p ≤ 0.1; and also at 12-months there was an improvement in the CSI score, p ≤ 0.1.

In terms of cost effectiveness in Memphis at 3-months the incremental cost-effectiveness ratio (ICER) was $1,236 (95% CI, $492 to $17,728) per additional point change on the CSI scale. In Lane County at 12-months diversion reduced the probability of drug use by 80% at no greater cost. In Tucson at 12 months diversion resulted in an ICER of $190 per additional point change on the CSI scale; and in New York, at 12 months, diversion reduced the odds of non-violent victimization by approximately 70% and also resulted in cost savings (that is, the intervention was dominant).Based on the above results the authors concluded that taken together with the findings from previous studies on jail diversion the results of this study provide mounting evidence that jail diversion results in positive outcomes for individuals, systems, and communities (Steadman & Naples, 2005).

The study is only partially applicable to the NICE decision-making context. It has been conducted in the US. The measure of outcomes was not expressed in QALYs, which made interpretation of findings difficult. The study was judged to have potentially serious methodological limitations; clinical effectiveness data were obtained from an observational study, the time horizon was relatively short (1 year), and resource use and unit cost data were obtained from a mixture of national and local sources and published studies.

7.2.21.405 Hughes and colleagues (2012)

Hughes and colleagues (2012) assessed the costs associated with a jail diversion programme in the US (Travis County, Texas). The economic analysis was based on an observational cohort study and economic modelling. The study sample consisted of 422 adults with a serious non-specified mental illness. Clinical effectiveness data (that is, transition probabilities) were obtained from published literature and expert opinion. Resource use and unit cost data were derived from the observational study participants and interlinked administrative databases, claims data and, where necessary, were supplemented with an expert opinion. The time horizon of the economic analysis was 2 years, and its perspective was public sector, including healthcare, social care, and criminal justice sector costs. Cost elements comprised criminal justice sector (police, pre-trial services, court, jail, and probation) and healthcare and social care (residential care, emergency services, inpatient treatment, outpatient treatment, rehabilitation and support services).

According to the analysis the mean costs per person at year 1 were $9,163 and $8,343 for diverted and non-diverted group, respectively; the difference was $820 in favour of the non-diverted group (in likely 2006 US dollars). The mean costs per person over 2 years were $12,946 and $14,307 for diverted and non-diverted group, respectively; the difference was -$1,361 (suggesting savings for the diverted group). Based on these results, jail diversion seems to offer good value for money over no diversion from a public sector perspective.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged to have potentially serious methodological limitations, including the relatively short time horizon (2 years), the estimates of relative treatment effects were obtained from an observational study, some of the resource use data were based on expert opinion, and the unit costs were based on administrative and claims data.

7.2.21.136 Mitton and colleagues (2007)

Mitton and colleagues (2007) assessed the costs and consequences of the post-booking component of a jail diversion programme in Calgary, Canada. The economic analysis was based on an observational before-after study. The study sample consisted of 117 adults with a serious mental illness and co-occurring substance use disorder. Clinical effectiveness data were obtained from the observational pre-post study. The time horizon of the economic analysis was 18 months, and its perspective was public sector, including healthcare and criminal justice sector costs. Cost elements comprised programme provision, hospital admissions, other inpatient visits, emergency room visits, complaints, charges and court appearances. Cost data were obtained for observational study participants from interlinked health and police administrative databases, and where necessary were supplemented with data from other published studies. The measures of outcome utilised in the economic analysis were the total Brief Psychiatric Rating Scale (BPRS) scores and the quality of life (Wisconsin Quality of Life Questionnaire of service users). Costs were reported at 18-months pre- and post the diversion programme, and outcomes at baseline and at 3-months.

According to the analysis the mean cost per participant over 18-months was $9,542 and $7,820 pre-diversion and post-diversion, respectively; a difference of $1,721 (in likely 2006 Canadian dollars), $p = 0.201$ (in favour of the post-diversion group). In terms of effectiveness the mean BPRS scores were 45.78 (SD 12.03) and 35.02 (SD 8.96) at baseline and at 3-months, respectively; a difference of 10.76, $p \leq 0.001$. The mean total scores on Wisconsin Quality of Life scale were 0.29 (SD 0.95) and 1.06 (SD 0.84) at baseline and at 3-months, respectively; a difference of 0.77, $p < 0.01$. Based on these results, the authors concluded that jail diversion improved outcomes and reduced overall costs.

The study is only partially applicable to the NICE decision-making context. It has been conducted in Canada and has not considered QALYs. The study was judged to have potentially serious methodological limitations, including the relatively short time horizon (up to 18 months); the estimates of relative treatment effects were obtained from an observational before-after study, resource use data were obtained from a mixture of sources, and the source of unit cost data was unclear. Moreover, the GC expressed concerns on how well BPRS and Wisconsin Quality of Life measures captured health consequences; finally, costs and outcomes were measured at different time points.

7.2.21.142 Mental health courts

7.2.21.251 Kubiak and colleagues (2015)

Kubiak and colleagues (2015) evaluated the cost effectiveness of a mental health court programme compared with no mental health programme alternative in adult offenders with a diagnosis of mental illness (bipolar, depressive, schizophrenia, and other) in the US. The majority had a co-occurring substance abuse problem. The analysis was conducted alongside an observational cohort study (N=150). The measures of outcome for the

economic analysis included the residential care days, jail days, and prison days. The time horizon of the analysis was 12 months, and its perspective was public sector, including healthcare and criminal justice sector costs. Cost elements comprised mental health treatment (case management, medication reviews, individual/group therapy, intensive outpatient treatment, residential treatment, psychiatric hospitalization, crisis residential, or crisis centre, arrest and incarceration); substance abuse treatment (residential and outpatient treatment); arrests; jail; court; incarceration; and victimisation. The resource use estimates were derived from the observational cohort study and other published sources. The unit costs were obtained from local and national sources; and where necessary were supplemented with information from other published studies. The results were reported for successful and unsuccessful mental health court participants; combined cost data was not available.

For successful mental health court participants, the intervention resulted in fewer residential care days at 12 months compared with standard care (0.00 versus 21.47, respectively; difference -21.47, p<0.001); fewer jail days (4.73 versus 49.27; difference -44.54, p<0.001); and fewer prison days (5.38 versus 48.70; difference -43.32, p<0.001). There was also a reduction in the mean number of arrests, jail bookings, court cases, and victimisation cases. However, these reductions were non-significant. The mean total costs per person over 12 months were $16,964 for the intervention and $39,870 for standard care, a difference of -$22,906 (p = ns) in 2013 prices. Based on the above findings mental health court programme was found to be the dominant intervention.

For unsuccessful mental health court participants, the intervention resulted in fewer residential care days at 12 months compared with standard care (1.57 versus 21.47, respectively; difference -19.9, p<0.001); fewer jail days (23.20 versus 49.27; difference -26.07, p<0.001). However, mental health court participants had more prison days (130.00 versus 48.70; difference 81.3, p<0.001). There was also a reduction in the mean number of arrests, jail bookings, court cases, and victimisation cases. However, these reductions were non-significant. The mean total costs per person over 12 months were $32,258 for the intervention and $39,870 for standard care, a difference of -$7,612 (p=ns). Based on the above findings mental health court programme was found to be the dominant intervention using residential care days and jail days as outcome measures. Using prison days as an outcome measure standard care resulted in an ICER of $94 per additional prison day avoided.

The analysis was judged by the GC to be partially applicable to the NICE decision-making context since the study was conducted in the US. The authors did not attempt to estimate QALYs. However, this was not a problem for judging cost effectiveness as the intervention was found to be dominant for successful participants. Overall this study was judged by the GC to have potentially serious methodological limitations including a short time horizon and some of the unit cost estimates being from local sources.

7.2.21.202 Ridgely and colleagues (2007)

Ridgely and colleagues (2007) evaluated the costs of a mental health court programme versus standard care in the US (Allegheny County, Pennsylvania). Standard care was defined as a normal judicial process. The economic analysis was based on an observational before-after study. The study sample consisted of 365 adults with a diagnosis of mental illness (or co-occurring mental and substance abuse disorder). Clinical effectiveness data were obtained from the study participants and where necessary were supplemented with expert opinion. The time horizon of the economic analysis was 2 years, and its perspective was public sector, including healthcare and criminal justice sector costs plus transfer payments. Cost elements comprised mental health and substance abuse treatment, arrests, incarceration, probation, and cash assistance payments. Cost data were collected for the observational study participants from various interlinked information systems, claims data, other published studies and as necessary were supplemented with authors' assumptions.

Two methods were used to estimate standard care costs. Using the first method standard care costs were approximated with actual service use for study participants in years prior to the programme enrolment (or index arrest), and using the second method standard care costs were based on authors' assumptions (informed by sentencing guidelines) about the criminal penalties that participants would likely have experienced had there been no mental health court programme.

According to the analysis, mental health court programme resulted in an increase of $2,656 per participant in actual costs in year 1 following mental health court programme entry compared with hypothetical costs based on the sentencing guidelines. Based on a before-after comparison mental health court programme resulted in a decrease in costs of $1,804 per participant at year 1, and also in a decrease in costs of $7,780 per participant at year 2; with an overall decrease in costs of $9,584 per participant over 2 years (in likely 2006 US dollars). According to the deterministic sensitivity analysis when assuming higher offending rates mental health court programme resulted in an increase in healthcare costs from $2,656 to $2,824 per participant in year 1 following mental health court programme entry compared with hypothetical costs based on the sentencing guidelines. Similarly, assuming that in the absence of mental health court programme individuals would use 10% fewer mental health services resulted in an increase in the costs from $2,656 to $4,052 per participant in year 1 following mental health court programme entry compared with hypothetical costs based on the sentencing guidelines. Based on these results, mental health court programme may potentially be cost saving.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged by the GC to have potentially serious methodological limitations, including the relatively short time horizon (24 months), the estimates of relative treatment effects being obtained from an observational before-after stud, the resource use and unit cost data being based on a mixture of county and state sources and authors' assumptions, and significance levels were not reported.

7.2.21.3 Drug court programmes

7.2.21.3.1 Cheesman and colleagues (2016)

Cheesman and colleagues (2016) assessed the costs of drug court programmes versus standard care in the US (Virginia). Standard care was defined as a combination of jail, Prison, and/or probation). The economic analysis was based on an observational cohort study. The study sample consisted of 1,944 adult offenders with substance abuse problems. The time horizon of the economic analysis was 2 years, and its perspective was public sector, including healthcare and criminal justice sector costs. Cost elements comprised drug court (assessment, staffing and court sessions, court treatment, testing, court supervision), fees, arrest, pre-trial supervision, pre-trial confinement, general district court cost, circuit court costs, misdemeanour arrest, felony arrests, jail, prison, probation, and victimisation (property and person). Cost data were collected from an observational study, survey, and other interlinked administrative databases.

According to the analysis the mean cost per participant over 2 years was $44,249 and $63,483 for drug court and non-drug court participants, respectively; a difference of -$19,234 (in 2012 US dollars). Based on the above findings drug courts appear to be cost saving from a public sector perspective).

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged by the GC to have potentially serious methodological limitations, including the lack of reporting of levels of statistical significance, and the study design (observational cohort study) although the sample size was large, and the source of unit cost data was unclear.

7.2.21.312 Carey and colleagues (2004)

Carey and colleagues (2004) assessed the costs of a drug court programme versus no drug court programme in the US (Multnomah County, Oregon). The economic analysis was based on an observational cohort study. The study sample consisted of 1,173 adult offenders with substance abuse problems. The time horizon of the economic analysis was 30 months, and its perspective was public sector, including healthcare, social care and criminal justice sector costs. Cost elements comprised court, public defender, district attorney, law enforcement (arrests, bookings, and jail and court time), treatment and probation. Cost data were collected for observational study participants from an interlinked administrative databases and claims data.

According to the analysis the mean cost per participant over 30 months was $14,910 and $18,681 for drug court and non-drug court participants, respectively; a difference of -$3,770 (in 2002 US dollars). Based on the above findings the authors concluded that drug courts can be a cost-effective use of taxpayer resources (Carey & Finigan, 2004).

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged by the GC to have potentially serious methodological limitations, including the lack of reporting of levels of statistical significance, and the study design (observational cohort study), although it was acknowledged that the study sample was large.

7.2.21.313 Logan and colleagues (2004)

Logan and colleagues (2004) assessed the costs associated with 3 drug court programmes versus no programme for adults with substance abuse problems in the US, Kentucky. The economic analysis was based on an observational cohort study (N=745) and modelling. The time horizon of the economic analysis was 1 year, and its perspective was public sector, including health and social care, criminal justice and welfare costs. Cost elements comprised criminal justice (prison, jail, parole, probation, convictions, charges, and orders), healthcare (inpatient and outpatient mental health), social accidents, child support, and earnings. Cost data were collected for the observational study participants from various interlinked administrative databases and were supplemented as necessary with information from published studies. The costs were reported per graduate episode, per terminator episode, and per participant episode. Regression analysis was used to model the financial benefits.

According to the analysis, the 12-month programme cost and tangible benefits were $5,132 and $19,658 per graduate episode, respectively (in 1999 US dollars), resulting in cost-savings of $14,526; per terminator episode, the 12-month programme cost and tangible benefits were $1,791 and $2,022, respectively, resulting in cost-savings of $231; and per participant episode the 12-month programme cost and tangible benefits were $3,178 and $8,624, respectively, leading to savings of $5,446. Based on these results, the authors concluded that the drug court programme was associated with a reduction in incarceration, mental health services, and legal costs, as well as an increase in earnings and child support payments.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged by the GC to have potentially serious methodological limitations, including the relatively short time horizon (12 months), the study design (observational cohort study), and the source of unit cost data was unclear.

7.2.21.314 Shanahan and colleagues (2004)

Shanahan and colleagues (2004) assessed the cost effectiveness of a drug court programme versus standard care (defined as a normal judicial process) for adult criminal offenders addicted to illicit drugs in Australia. The economic analysis was undertaken alongside an RCT (N=468) included in the guideline systematic review (Shanahan 2004).

Clinical effectiveness data were obtained from the RCT and various interlinked administrative databases, and other local information systems. The time horizon of the economic analysis was 23 months, and its perspective was public sector, including healthcare and criminal justice sector costs. Cost elements comprised programme provision, court, assessment and detoxification, treatment, monitoring, and incarceration. Cost data for RCT participants were obtained from various administrative databases and other information systems. The primary measures of outcome utilised in the economic analysis were the time to the first offense and offending frequency per year.

According to the analysis the mean cost per day per participant was $144 and $152 for intervention and standard care groups, respectively; a difference of -$8 (in favour of the intervention), in 2003 Australian dollars. Drug court programme was shown to be the most effective intervention in terms of reduction in the time to the first offense and offending frequency per year. The mean days to the first drug-related offense was 325 and 279 for the participants in the intervention and standard care group, respectively; a difference of 46 days (p = 0.005). Similarly, the mean number of drug-related offenses per day was 0.009 and 0.012 for the intervention and standard care group, respectively; a difference of -0.004 (p = ns). Based on the above findings intervention was dominant when compared with standard care (more effective and less costly). According to the deterministic sensitivity analysis only when the proportion of sentence served was varied (assuming that only 66% of the sentence was served) was the cost per day for the intervention group higher than that for the standard care group. Based on these results, drug court programme seems to offer good value for money when compared with a standard judicial process.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in Australia and only non-health outcomes were considered. The study was judged by the GC to have potentially serious methodological limitations, no consideration of health outcomes, significance levels were not reported, and the source of unit cost data was unclear.

7.2.21.4 Street triage

7.2.21.4.1 Heslin and colleagues (2016)

Heslin and colleagues (2016A) evaluated the costs of a street triage model where a psychiatric nurse attended incidents with a police constable compared with usual care. The study was conducted in the UK (Sussex, South East England). Standard care was defined as only police attendance to all mental health incidents. The study population comprised adults with mental health problems who were detained under Section 136 or had contact with street triage. The economic analysis was based on observational before-after study and modelling. The area in which street triage was implemented covered 99,412 people, and for the rest of the county the population size was 688,654. The source of effectiveness data for the economic model was an observational before-after study and authors' assumptions. The main analysis was conducted from a public sector perspective (NHS and criminal justice sector). The results were also reported from NHS only and criminal justice sector only perspectives. The study considered a range of costs including the provision of street triage services (police constable and nurse), detention in custody (officer attendance, cost of time in custody, mental health act assessment, referral to GP), detention in hospital (officer attendance, inpatient bed day, mental health act assessment), GP visits, community mental health teams, A&E attendances, social worker attendances, and inpatient care. The resource use estimates were based on the observational before-after study, assumptions, and other published sources. The unit costs were obtained from national sources. The time horizon of the analysis was 1 day. The mean cost was estimated based on costs incurred by people seen by the services over a period of 6 months.

According to the economic modelling results, from a public sector perspective, the mean total costs per participant were £1,043 for the intervention and £1,077 for standard care, a difference of -£34 in 2013/14 prices. The mean NHS costs per participant were £574 for the intervention and £517 for standard care, a difference of £57. When considering only criminal justice sector costs the mean costs per participant were £470 for the intervention and £559 for standard care, a difference of -£89. Interestingly the intervention leads to an increase in NHS costs, but a reduction in criminal justice sector costs.

The analysis was judged by the GC to be partially applicable to the NICE decision-making context. The authors have not considered health outcomes and did not attempt to estimate QALYs. This study was judged by the GC to have potentially serious methodological limitations, including its short time horizon, and the fact that some model's inputs were based on authors' assumptions.

7.2.21.132 Heslin and colleagues (2016)

In another study Heslin and colleagues (2016B) evaluated the costs associated with 3 scenarios, including street triage; offering Mental Health Act assessments to all individuals detained under the Mental Health Act Section 136; and having a link worker present at custody suites. The scenarios were compared with standard care. The study was conducted in the UK. The study population comprised adults with mental health problems who are in contact with the criminal justice system. Standard care was defined as locally available services and did not include any of the above services (that is, street triage, Mental Health Act assessments for all individuals or a link worker at custody suites). The economic analysis was based on an observational cohort study (N=55) and further decision analytic modelling. The analysis was conducted from a public sector perspective that included NHS and criminal justice sector costs. The study considered a range of costs including mental health care (in-patient services; client contacts with mental health staff; meetings in the absence of client; and client assessments), police and other emergency services (police contacts/attendance, ambulance attendance at incident), custody services (length of stay in custody suite, Mental Health Act assessments, health care practitioner triage, forensic medical examiner, approved mental health practitioner, hospital attendance) and other services (transport, follow-up calls by police and escorting). The unit costs were obtained from national sources. The time horizon of the analysis was 1 year. The costs were reported per incident.

The mean total NHS and criminal justice sector costs per incident associated with the standard care pathways were £522. Offering street triage services resulted in total costs of £526 per incident (an increase of £4), offering Mental Health Act assessment for all Section 136 detainees resulted in total costs of £526 (an increase of £4), and having a link worker present at custody sites resulted in total costs of £534 (an increase of £12) in 2011/12 prices. Sensitivity analyses indicated that the estimated costs from the NHS and criminal justice sector were robust, with total costs associated with street triage ranging from £478 to £568; total costs associated with the Mental Health Act assessment for all Section 136 detainees ranged from £530 (including a forensic medical examiner in all custody suites) to £532 (including a forensic medical examiner contact and healthcare practitioner in all custody suites); and assuming a client contact duration of 3h with link worker rather than 1h increased costs to £557. Overall recommended enhancements to care pathways only marginally increased costs per incident.

The analysis was judged by the GC to be partially applicable to the NICE decision-making context. The authors did not attempt to measure health outcomes and to estimate QALYs. Overall, this study was judged by the GC to have potentially serious methodological limitations, including a small sample size, and the fact that resource use data were based on a small observational cohort study.

7.2.21.15 Integrated Disorders Treatment Program (IDDT)

7.2.21.52 Chandler & Spicer (2006)

Chandler and Spicer (2006) evaluated the cost effectiveness of an integrated dual disorders treatment programme compared with standard care in the US. The study population comprised adult jail recidivists with serious mental illness and substance use disorders. The economic analysis was conducted alongside a RCT (Chandler 2006) (N=182). The authors intended to adopt a public sector perspective (healthcare payer and criminal justice sector), however only 12-month outcome data were possible to report and comparable (12-month) cost data were available only from a healthcare payer perspective. The healthcare payer perspective included mental health service costs (outpatient and inpatient care, crisis visits, and psychiatric medications). The resource use estimates were based on the RCT (N used to estimate resource use is unclear). The unit costs were obtained from local sources (Alameda County, California). The measures of outcome for the economic analysis included the arrests, convictions, felony convictions, and jail days. The intended time horizon of the analysis was 18 months. However, comparable cost and outcome data were available only at 12 months.

The intervention resulted in a greater reduction in arrests (-0.68 versus -0.23, respectively; difference -0.45; a greater reduction in convictions (-0.10 vs 0.12, respectively; difference -0.22); and a greater reduction in jail days (-36.03 days vs -20.05 days, respectively; difference -15.98 days). When considering felony conviction standard care resulted in a greater reduction (0.02 versus 0.03; difference -0.01). From a healthcare payer perspective, the mean costs per person over 12 months were $5,620 for the intervention and $4,828 for standard care, a difference of $792 in likely 2005 prices. Levels of statistical significance were not reported for differences in costs and outcomes between the groups. Based on the above findings from a healthcare payer perspective (mental health service costs only) the intervention resulted in an ICER of $1,671 per additional arrest avoided; $3,418 per additional conviction avoided; and $47 per additional jail day avoided. When using felony convictions as an outcome measure standard care was the dominant option (that is, it resulted in lower costs and greater reduction in felony convictions).

The analysis was judged by the GC to be partially applicable to the NICE decision-making context since it was conducted in the US. The authors did not attempt to estimate quality-adjusted life years (QALYs) which made it difficult to interpret the cost-effectiveness results and to compare the findings with those of other studies. Overall, this study was judged by the GC to have potentially serious methodological limitations including a short time horizon, the consideration of mental health costs only (for total costs from a public sector perspective no comparable outcomes were reported), and the use of local unit costs.

7.2.21.16 Forensic assertive community treatment (FACT)

7.2.21.81 Cusack and colleagues (2010)

Cusack and colleagues (2010) assessed the cost effectiveness of forensic assertive community treatment (FACT) compared with treatment as usual (TAU) in the US. FACT comprised team-based mental health and substance abuse services, as well support for housing, employment assistance, benefits applications, and advocacy. TAU was defined as services routinely available in the county-operated public behavioural health system. The economic analysis was undertaken alongside a RCT included in the guideline systematic review (Cusack 2010). Clinical effectiveness data were obtained from the study participants. The study sample consisted of 134 adult detainees with a serious mental illness (a psychotic disorder including schizophrenia-spectrum or other psychotic disorders) in the county jail. The majority of detainees also had a co-occurring substance abuse problem. The time horizon of the economic analysis was 24 months, and its perspective was public sector, including healthcare and criminal justice sector costs. Cost elements comprised healthcare

(hospital admissions, psychiatric crisis contacts, outpatient services for both mental health and substance abuse) and criminal justice (bookings, convictions, and jail). Cost data were collected for RCT participants from various interlinked administrative databases, and claims and reimbursement databases. The measures of outcome utilised in the economic analysis were bookings, jail days, and convictions. Costs were reported for 2 time periods: 0-12 months and 13-24 months.

According to the analysis the mean cost per participant over the first 12 months was $20,859 (SD $26,494) and $17,475 (SD $31,163) for FACT and TAU group, respectively; a difference of $3,384 in likely 2009 US dollars. The mean cost per participant over 13-24 months was $14,182 (SD $25,680) and $14,436 (SD $28,869) for FACT and TAU group, respectively; a difference of -$254. In terms of effectiveness the mean bookings per participant were 1.21 and 2.31 for FACT and TAU group, respectively; $p < 0.01$. The mean jail days per participant were 39 and 65.8 for FACT and TAU group, respectively; p-value was unclear. The mean convictions per participant were 1.13 and 1.4 for FACT and TAU group, respectively; p = ns. Based on the above, the ICERs associated with the intervention are: $2,845 per additional booking avoided, $117 per additional jail day avoided, and $11,593 per additional conviction avoided

According to the authors a range of other costs were not included in the analysis (for example, court costs), that would have favoured FACT. The authors stated that FACT leads to reduced criminal justice involvement, reduced psychiatric hospitalisations, and reduced costs for offenders with serious mental illness and criminal justice involvement. However, due to the lack of QALYs the GC could not judge whether FACT represents value for money.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged to have potentially serious methodological limitations, including the relatively short time horizon (2 years), the lack of consideration of health outcomes, and the fact that resource use data were based on local administrative data.

7.2.21.7 Therapeutic community treatment

7.2.21.7.1 McCollister and colleagues (2003A)

McCollister and colleagues (2003A) evaluated the cost effectiveness of a work release therapeutic community and aftercare programme CREST for the management of adult drug-abusing criminal offenders in the US. CREST comprised a co-educational, 6-month programme and also aftercare that runs for 6 months and involves both group and individual counselling sessions weekly. CREST programme was compared with SC defined as standard work release programme. The economic analysis was undertaken alongside an RCT (N=836) included in the guideline systematic review (McCollister 2003). The time horizon of the economic analysis was 18 months, and it adopted the perspective of a local prison service provider. Cost elements comprised only the programme provision (personnel, programme supplies and materials, contracted services, and equipment). The sources of unit cost data were not reported. Some of the cost data was supplemented with information from published sources. The primary measure of outcome utilised in the economic analysis was the number of days incarcerated. Costs and outcomes associated with the intervention were reported for those who completed work release programme only, and for those who completed work release and aftercare programme.

According to the analysis the mean cost was $1,604 (SD $714) per participant completing only CREST work release programme, and $2,539 (SD $468) per participant completing both CREST work release and aftercare programme (in 1997/1998 US dollars). There were no additional costs associated with SC, consequently the cost of SC was $0 in the analysis. The differences between all groups were statistically significant, $p < 0.01$.

In terms of effectiveness the mean number of days incarcerated was 92 (SD 112) per participant completing only CREST work release programme, and 43 (SD 86) days per participant completing both CREST work release and aftercare programme. The mean number of days incarcerated for participants in the SC group was 104 (SD 128). The differences between all groups were statistically significant, $p < 0.01$.

Based on the above findings when comparing CREST work release participants with SC the mean cost per additional day of incarceration was $134, and then comparing CREST plus aftercare programme participants with CREST work release only participants the mean cost per additional day of incarceration avoided was $19 (95% CI, $14 to $28). The authors concluded that work release programme was not cost-effective since this cost per avoided incarceration day was actually slightly higher than the average daily cost of incarceration of $57. It seems that therapeutic community treatment (in particular CREST plus aftercare programme) represents reasonable value for money.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US and has adopted a narrow prison service provider perspective. The measure of outcomes was not expressed in QALYs and the study was judged by the GC to have potentially serious methodological limitations, including the relatively short time horizon (18 months) and the lack of consideration of health outcomes. The analysis has not considered wider healthcare, social care and criminal justice sector costs; and the source of unit cost data was unclear.

7.2.21.212 McCollister and colleagues (2003B)

McCollister and colleagues (2003B) evaluated the cost effectiveness of a prison-based therapeutic community and aftercare programme for the management of adult drug-abusing criminal offenders in the US. Therapeutic community and aftercare programme was compared to the therapeutic community programme only (that is, no aftercare) and to no treatment alternative. The economic analysis was undertaken alongside an RCT (N=715) included in the guideline systematic review (McCollister 2003B). The time horizon of the economic analysis was 1 year, and it adopted the perspective of a local prison service provider. Cost elements comprised only the programme provision (personnel, programme supplies and materials, contracted services, and equipment). The sources of unit cost data were not reported. Some of the cost data was supplemented with information from published sources. The primary measure of outcome utilised in the economic analysis was the number of days incarcerated.

According to the analysis the mean cost was $2,708 (95% CI: $2,568; $2,847) per participant in therapeutic community programme only and $6,985 (95% CI: $6,509; $7,489) per participant in therapeutic community and aftercare programme (in 1993 US dollars). There were no additional costs associated with SC, consequently the cost of SC was $0 in the analysis. The differences between all groups were statistically significant, $p < 0.001$.

In terms of effectiveness the mean number of days incarcerated was 118.4 (95% CI: 104; 133) per participant in therapeutic community programme only and it was 34.41 (95% CI: 22; 48) days per participant in therapeutic community and aftercare programme. The mean number of days incarcerated for participants in the SC group was 142.30 (95% CI: 126; 160). The differences between all groups were statistically significant, $p < 0.05$.

Based on the above findings when comparing therapeutic community programme with no treatment alternative the mean cost per additional day of incarceration was $113, and then comparing therapeutic community programme plus aftercare with therapeutic community treatment only the mean cost per additional day of incarceration avoided was $51. It seems that therapeutic community treatment (in particular therapeutic community treatment plus aftercare programme) represents reasonable value for money.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US and has adopted a narrow prison service provider perspective. The measure of outcomes was not expressed in QALYs and the study was judged by the GC to have potentially serious methodological limitations, including the relatively short time horizon (1 year) and the lack of consideration of health outcomes. The analysis has not considered wider healthcare, social care and criminal justice sector costs; and the source of unit cost data was unclear.

7.2.21.783 McCollister and colleagues (2004)

McCollister and colleagues (2004) evaluated the cost effectiveness of prison-based therapeutic community (TC) and post-release community based addiction treatment versus SC in the US, Southern California. SC was defined as no prison-based substance abuse treatment. The economic analysis was undertaken alongside a RCT included in the guideline systematic review (McCollister 2004). The study sample consisted of 576 adult drug abusing criminal offenders. The time horizon of the economic analysis was 5 years, and its perspective was that of the local prison service provider. Cost elements comprised programme provision and treatment including hospital inpatient, prison-based residential TC, community-based residential TC, day treatment (day care rehabilitative programmes), outpatient methadone maintenance, outpatient detoxification, and outpatient drug-free, other outpatient (private counselling), sober living, and self-help/12-step programmes. Cost data were collected for study participants from interlinked criminal justice records, and various other local and national sources. The source of unit costs was unclear. The primary measure of outcome utilised in the economic analysis was the number of days incarcerated. Costs and outcomes associated with the intervention were reported for those who completed the TC programme only and for those who completed the TC and aftercare programme.

According to the analysis the mean cost over 5 years was $3,356 (95% CI, $2,702 to $4,179) per participant completing only the prison TC programme, $15,325 (95% CI, $10,159 to $21,640) per participant completing the prison TC plus post-release treatment, in 2000 US dollars. The mean cost per participant in the SC group was $1,731 (95% CI, $1,084 to $2,713). The differences between all groups were statistically significant, $p < 0.01$.

In terms of effectiveness the mean number of days incarcerated over 5 years was 634 (95% CI, 565 to 690) days per participant completing only the prison TC programme, and 343 (95% CI, 261 to 438) days per participant completing the prison TC plus post-release treatment. The mean number of days incarcerated over 5 years in the SC group was 626 (95% CI, 565 to 690). The differences were statistically significant between participants in the SC group, and those in the prison TC only and the prison TC plus post-release treatment groups, $p<0.01$.

Based on the above findings, the prison TC only group was dominated by SC (less effective and higher costs). When comparing prison TC plus post-release treatment group with SC the cost per additional incarceration day avoided was $48.

Based on these results, the authors concluded that, when considering the average daily cost of incarceration in California ($72), offering substance abuse treatment in prison and then directing offenders into community based aftercare treatment was a cost-effective option (McCollister et al., 2004). Similarly, the GC considered the above and judged that therapeutic community treatment represents reasonable value.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US and the measure of outcome was not expressed in QALYs. The study was judged by the GC to have potentially serious methodological limitations. It has not considered health outcomes and criminal justice sector costs, and the resource use and cost data were based on a mixture of state-wide and local sources, and the source of unit cost data was unclear.

7.2.21.18 Probation and mandated treatment

7.2.21.821 Anglin and colleagues (2013)

Anglin and colleagues (2013) assessed the costs of mandated probation or continued parole with substance abuse treatment versus SC for adult offenders convicted of non-violent drug offenses and probation or parole violators in the US. SC was defined as a traditional probation where treatment is left to the client's choice. The economic analysis was based on a large observational cohort study (intervention N=47,355; control N=41,607). Clinical effectiveness data were obtained from observational study participants and other published sources. The time horizon of the economic analysis was 30 months, and its perspective was public sector, including healthcare and criminal justice sector costs. Cost elements comprised prison, jail, probation, parole, arrests, convictions (including adjudication costs), publicly funded healthcare use, and substance abuse treatment. Cost data were obtained for observational study participants from an interlinked administrative database, claims data, and other published sources. Regression analysis was used to adjust the cost differences for baseline differences in participant characteristics including individual-level characteristics (age, gender and race) and for country-level characteristics (baseline arrests per capita and change in arrests per capita).

According to the analysis for the intervention group unadjusted mean costs per participant were $16,935 (SD $21,412) and $25,251 (SD $24,894) over 30 months prior to and post the index conviction, respectively; a difference of -$8,316 (SD $24,712) in 2009 US dollars. Similarly, for the control group unadjusted mean costs per participant were $15,294 (SD $21,074) and $26,595 (SD $25,911) over 30 months prior to and post the index conviction, respectively; the difference of -$11,301 (SD $24,853). The unadjusted difference between control and intervention groups was –$2,681 (95% CI, –$3,007 to –$2,354), the adjusted difference for country-level characteristics was –$2,173 (95% CI, –$2,584 to –$1,762) and the adjusted difference for both individual-level and country-level characteristics was –$2,317 (95% CI, –$2,730 to –$1,905). Based on the above findings the authors concluded that the monetary benefits of the programme exceeded the additional cost of implementation and provision of treatment (Anglin et al., 2013).

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged by the GC to have minor methodological limitations, including the estimation of the relative treatment effects from a large cohort study and other published studies, the resource use data were obtained from a mixture of sources, and the source of unit cost data was unclear.

7.2.21.852 Alemi and colleagues (2006)

Alemi and colleagues (2006) assessed the costs of combining probation and substance abuse treatment versus traditional probation (where treatment is left to the client's choice) for substance abusing adult offenders in the US (Baltimore-Washington, DC). This study was based on an RCT and decision analytic modelling. Decision analytical modelling was used to synthesise the evidence. Probabilities of events and resource use data were obtained from the RCT (N=272) and published studies. The time horizon of the economic analysis was 2.75 years, and its perspective was public sector, including healthcare, social care, and criminal justice sector costs. Cost elements comprised programme provision, treatment (mental health and substance abuse), physical healthcare, arrests, re-offending, and legal costs, violation, conviction and sentencing, prison, tax earnings, and shelter accommodation. Cost data were obtained for RCT participants from interlinked state and county information systems, and as necessary were supplemented with information from other published sources, and authors' assumptions.

According to the analysis the expected daily mean cost per participant was $39 and $22 for the combined probation and substance abuse treatment, and traditional probation, respectively; the difference was $17 per day or $6,293 per year per participant (in favour of

the traditional probation), in 2004 US dollars. Deterministic sensitivity analysis showed that there was no change in a rate of any single adverse outcome (arrest, mental hospitalisation, or incarceration), which could make probation combined with substance abuse treatment cost saving. A reduction of more than 50% in all of the adverse outcome rates was required to make combined probation and substance abuse treatment more cost saving. Also, a minimum of 69% reduction in mental hospitalisation rates and incarceration rates or an 8-fold increase in the cost of arrest was required for the combined probation and substance abuse treatment to become the cost saving option. Based on these results, combining probation and substance abuse treatment does not appear to offer good value for money over the traditional probation.

The study is only partially applicable to the NICE decision-making context, as it has been conducted in the US. The study was judged by the GC to have potentially serious methodological limitations, including the fact that some model inputs were based on authors' assumptions and that resource use and unit cost data were based on a mixture of state and county sources.

7.2.21.9 Medium security units

7.2.21.9.1 Fortune and colleagues (2011)

Fortune and colleagues (2011) evaluated the cost effectiveness of an inpatient medium security unit (MSU) and a residential service managed by a local housing association compared with an impatient MSU and a community team, and an impatient MSU, a community team and a residential service in personality-disordered male offenders in the UK. Participants were grouped and compared according to whether they were being treated by MSUs or community/residential services. The economic analysis was based on an observational cohort study (N=54, N=42 at a 6-month follow-up, N=25 at a 24-month follow-up). The analysis was conducted from a public sector perspective (healthcare, social care, and criminal justice system). The study considered a range of costs including accommodation (hostels, MSU, low secure unit, prison, high secure hospital, bed and breakfast), health and community services (inpatient stay, outpatient appointments, A&E, GP, practice nurse, key worker, psychiatric nurse, psychiatrist, psychologist, counsellor/therapist, drug and alcohol worker, dentist, occupational therapist, social worker, day centre), and criminal justice services (probation, solicitor, police, police custody, court appearance). The resource use estimates were based on the observational cohort study (N=48). The unit costs were obtained from national sources. The measure of outcome for the economic analysis was an improvement in social functioning as measured on The Work and Social Adjustment Scale (WSAS). The time horizon of the main analysis was 2 years. Costs were reported as ranges.

The community and residential intervention resulted in a greater reduction on WSAS when compared with MSU services. The difference between baseline and a 6-month follow-up was -0.67 and -0.89 for MSU and community and residential treatment groups, respectively; a difference of -0.22 (p-value non-significant). Similarly, the difference between baseline and a 24-month follow-up was -3.5 and -5.92 for MSU and community and residential treatment groups, respectively; a difference of -2.42 (p-value non-significant).

The costs per service user per year ranged from £192,978 to £199,696 for MSU and from £111,943 to £162,752 for the community and residential care group; a difference of £36,944 to £81,035 in 2005/06 prices. Based on the above findings community and residential service is dominant however this is based on cost and outcomes reported over different time horizons.

The analysis was judged by the GC to be partially applicable to the NICE decision-making context. The authors did not attempt to estimate QALYs. This was not a problem since the intervention seems to be dominant. However, WSAS may be limited as an outcome measure

of overall HRQoL. This study was judged by the GC to have very serious methodological limitations including its short time horizon, lack of consideration of wider health outcomes, costs reported as ranges, and the study design (very small cohort study).

7.2.21.10 Cost analysis

7.2.21.10.1 Objective

A systematic review of the clinical evidence indicated that therapeutic community treatment delivered in prison setting may be effective in reducing future re-offending in people who have substance misuse disorders. No directly applicable economic evidence was identified assessing the cost-effectiveness of therapeutic community treatment for substance misuse in the UK. Given the lack of suitable outcome data to populate a full economic evaluation, a simple exploratory cost analysis was undertaken, which assessed the potential economic impact of therapeutic community treatment for the management of substance misuse problems in adults in a prison setting in the UK when compared with the 'no treatment' alternative. The cost analysis assessed whether the costs of providing therapeutic community treatment for substance misuse would be offset by future cost savings resulting from reduced incarcerations.

7.2.21.10.2 Methods

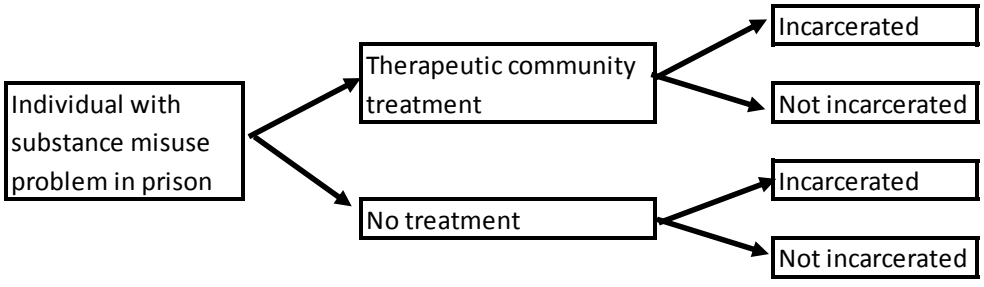
Intervention examined

Therapeutic community treatment for substance misuse delivered in a prison setting was modelled on the description of the programme in the RCT that assessed the intervention in the clinical review. The resource use information was modified by the GC to reflect the practice in the UK. The intervention was compared with the 'no treatment' alternative.

Model structure

A simple decision-tree was constructed using Microsoft Excel 2013 to estimate the costs of therapeutic community treatment for people with substance misuse problems. According to the model structure, adults with substance abuse problems in prison received either the therapeutic community treatment or 'no treatment'. During the duration of the model, individuals who received either intervention or no treatment could commit a crime and get incarcerated during the 12 month follow up. The time horizon of the model was 41.2 months (the treatment duration was 13 months, incarceration rates were considered at 12-month follow-up, and the duration of an incarceration was 16.2 months). A schematic diagram of the decision-tree is presented in Figure 1.

Figure 1. Schematic diagram of the structure of the economic model.



Costs considered in the analysis

People with substance misuse who relapse following the initial prison stay are likely to incur substantial costs to health and social care services and the criminal justice system. NICE recommends that economic analyses of interventions with health and non-health outcomes in public sector settings adopt a public sector perspective (NICE., 2014). According to the GC expert opinion, therapeutic community treatment is fully funded by the Ministry of Justice (MoJ). No data linking therapeutic community treatment for substance misuse and changes in the future health care needs were identified. As a result, the analysis adopted a narrow criminal justice perspective and considered only intervention and future incarceration costs. The exclusion of healthcare and social care costs is acknowledged as a factor reducing the applicability of the economic analysis to the guideline context. The GC felt that to enable informed decision making it was important to assess the cost implications even if only from the criminal justice system sector perspective. Discounting of costs was not undertaken due to the short time horizon of the analysis.

Model input parameters

Efficacy of therapeutic community treatment and baseline re-incarceration risk

Efficacy data regarding the relative effect of therapeutic community treatment versus 'no treatment' and the baseline effect of 'no treatment' alternative were approximated using data from 1 RCT (N=139) (Sacks 2004) of therapeutic community treatment for substance misuse that was included in the guideline systematic review, which reported offending outcome in the form of re-incarcerations at 12-month follow-up. This was the only RCT included in the systematic review that reported re-incarceration outcomes associated with the therapeutic community treatment at follow-up. The RCT compared a modified therapeutic community treatment versus CBT informed psychoeducation. However, in the model therapeutic community treatment is compared to 'no treatment' alternative. In effect, since the comparator in the RCT was active intervention the model is underestimating the cost effectiveness of therapeutic community treatment and provides a conservative estimate.

The RCT (Sacks 2004) found that the modified therapeutic community treatment relative to CBT informed psychoeducation reduced re-incarcerations, criminal activity in general, and specific alcohol and drug related offences at 12 month follow up (re-incarceration RR=0.28, 95% CI: 0.13 to 0.63). It was further estimated that of all adult offenders 25% reoffended, of these 35% received a custody or court order, and of these 34% got a determinate sentence of 12 months or more, and 10% received a lifetime sentence(Ministry of Justice., 2016a). This results in a baseline re-incarceration rate of 3.85% in the UK and this was utilised in the economic analysis.

Intervention cost

In the RCT (Sacks 2004) that informed the effectiveness of the therapeutic community treatment the intervention involved learning through self-help and community affiliation to foster change in themselves and others. The intervention had three fundamental modifications to the standard therapeutic community approach: (1) increased flexibility, decreased intensity, more individualisation; (2) emphasis on criminal thinking and behaviour, recognition and understanding of inter-relationship between substance misuse, mental illness and criminality; and (3) included medication, therapeutic interventions, psycho-educational classes and cognitive behaviour protocols. The resource use associated with the therapeutic community treatment was based on the GC expert opinion to reflect the provision of such an intervention in the UK. The economic analysis modelled therapeutic community treatment comprising induction, primary treatment, re-entry, one to one sessions, and morning community and social meetings. Induction was modelled to comprise 3 group sessions per week lasting 1.5 hours each for 6 weeks. Primary treatment was modelled as comprising 2 group sessions per week lasting 1.5 hours each for 36 weeks. Re-entry into the community

part of the programme was modelled as comprising 2 group sessions per week lasting 1.5 hours each for 12 weeks. The mean group size was assumed to be 8 people. One to one sessions comprised monthly sessions each lasting 1 hour for 13 months. Morning community and social activity group meetings comprised daily sessions each lasting 20 minutes for 52 weeks. These were assumed to be delivered in a communal setting comprising of 48 people. All sessions were assumed to be facilitated by a prison officer (Grade 4). The caseload per prison officer was assumed to be 8 people.

According to the GC expert opinion, people with substance misuse in prisons would usually get a short-term intervention lasting approximately 30-60 days; and only those who are on long-term sentences and for whom a short-term intervention hasn't worked would go on to receive a long-term intervention. Also, the approach adopted when managing substance misuse varies across prisons. As a result, in the model it was conservatively assumed that people with substance misuse who require long-term intervention at present do not receive any services. However, if it is found that therapeutic community treatment is cost-effective when compared with 'no treatment' alternative, and if people are actually receiving treatment then the cost-effectiveness of therapeutic community treatment has been underestimated (that is, the model provides a conservative estimate of therapeutic community treatment cost effectiveness).

The unit cost of prison officer was estimated to be £18.03 per hour (Prison Service Pay Review Body., 2016). This is based on a salary of £27,058 for a prison officer Grade 4 and 1,501 hours per year and includes 25 days annual leave and 10 statutory leave days.

Future incarceration costs

In order to estimate the costs of incarceration, the duration of incarceration and the cost of imprisonment was required. The average duration of a prison stay in the UK is approximately 16.2 months (Ministry of Justice., 2016b) and the average cost for holding one prisoner for a year is £34,087 (in 2013/14 prices) (Ministry of Justice., 2014). Costs were uplifted to 2014/15 UK pounds – using UK PPS hospital & community health services (HCHS) index (Curtis & Burns, 2015).

Table 195 presents the values of input parameters as well as the cost data that was used to populate the economic model.

Table 195: Values of input parameters as well as the cost data that was used to populate the economic model.

Input parameter	Value	Source of data – comments
Efficacy (reduction in incarcerations at 12 month follow-up) – therapeutic community treatment	RR 0.28 (95% CI: 0.13 to 0.63)	Sacks et al., 2004
Baseline rate of re-incarceration – 'no treatment' alternative	3.85%	It was estimated that of all adult offenders 25% reoffended, of these 35% received a custody or court order, and of these 34% got a determinate sentence of 12 months or more, and 10% received a lifetime sentence (MoJ, 2016).
Therapeutic community treatment - intervention cost	£815	GC expert opinion. Induction: 3 group sessions per week, 1.5 hours each for 6 weeks; primary treatment: 2 group sessions per week lasting 1.5 hours each for 36 weeks; re-entry: 2 group sessions per week lasting 1.5 hours each for 12 weeks. The mean

Input parameter	Value	Source of data – comments
		group size was: 8 people. One to one sessions: monthly, each lasting 1 hour for 13 months. Morning community and social activity group meetings: daily sessions each lasting 20 minutes for 52 weeks delivered in a group of 48 people. All sessions were assumed to be facilitated by a prison officer (Grade 4). The caseload per prison officer was assumed to be 8 people. The unit cost of prison officer was estimated to be £18.03 per hour (Prison Service Pay Review Body, 2016).
Duration of a prison stay	16.2 months	MoJ, 2016B
Prison costs	£34,087	MoJ, 2014. Costs were uplifted to 2014/15 UK pounds – using UK PPS hospital & community health services (HCHS) index (Curtis & Burns, 2015).

Sensitivity and threshold analyses

One-way sensitivity analyses were undertaken to explore the robustness of the results under the uncertainty characterising some model input parameters. The following parameters were tested in sensitivity analysis:

- intervention efficacy
- baseline incarceration rate
- duration of a prison stay
- intervention cost
- standard care cost

Threshold analyses were conducted to identify model input parameter values at which the conclusions might change.

Also, what if analysis was undertaken exploring a scenario where therapeutic community treatment is funded by the NHS and what would be the required quality adjusted life year (QALY) gain for the therapeutic community treatment to be considered cost effective (that is, for the ICER to be below the NICE lower cost-effectiveness threshold of £20,000 per QALY (NICE., 2008).

Validation of the economic model

The economic model (including the conceptual model and the Excel spreadsheet) was developed by the health economist working on this project and checked by a second modeller not working on the project. The model was tested for logical consistency by setting input parameters to null and extreme values and examining whether results changed in the expected direction. The assumptions and the results were discussed with the GC to confirm their plausibility.

7.2.21.1043 Results

Base-case analysis

It was estimated that the intervention cost per person for therapeutic community treatment for substance misuse is £815. The costs per person associated with the future incarcerations are £489 and £1,747 for therapeutic community treatment for substance misuse and ‘no treatment’ option, respectively. Therapeutic community treatment for substance misuse is

associated with the cost savings of £1,258 per person due to the reduction in re-incarceration costs. Based on the above findings, therapeutic community treatment for substance misuse is associated with the overall cost savings of £443 per person. Full results of the base-case analysis are reported in Table 196.

Table 196. Results of the economic analysis of therapeutic community intervention versus 'no treatment' alternative in people with substance misuse problems – mean costs per person

Intervention	Total intervention costs	Total future incarceration costs	Total costs	Incremental cost (intervention versus SC)
Therapeutic community treatment	£815	£489	£1,304	- £443 (intervention is cost saving)
No treatment alternative	£0	£1,747	£1,747	

Sensitivity and threshold analyses

The RCT (Sacks 2004) found that the modified therapeutic community treatment relative to the CBT informed psychoeducation reduced the re-incarcerations (RR=0.28; 95% CI: 0.13 to 0.63). Cost savings estimated using the lower and the upper estimate for RR were £705 and -£168 per person, respectively. This indicates that when using the upper estimate (RR 0.63, lower efficacy) of RR the therapeutic community treatment is associated with a slight increase in costs. The threshold analysis indicated that the RR would need to be 0.53 for the therapeutic community treatment and 'no treatment' option costs to break-even.

The model was found to be very sensitive to the baseline re-incarceration rate at 12 month follow up. The baseline re-incarceration rate was estimated to be 4%. Doubling this rate increases the cost savings associated with the therapeutic community treatment to £1,799 per person. The RCT from which the efficacy rate was taken (Sacks 2004) reports the baseline re-incarceration rate to be 32.8% at 12 month follow-up. Using this rate, the cost savings associated with the therapeutic community treatment would be approximately £10,000 per person. However, this was a US-based study with a very different criminal justice system setup. The threshold analysis indicated that the baseline re-incarceration rate would need to be 2.49% for the therapeutic community treatment and 'no treatment' option costs to break-even.

There is a high uncertainty surrounding the resource use estimates associated with the therapeutic community treatment. According to the sensitivity analysis, reducing and increasing the intervention cost by 50% resulted in the cost savings of £851 and £36 per person, respectively. The threshold analysis indicated that the intervention cost would need to be as highs as £1,258 per person for the therapeutic community treatment and 'no treatment' option costs to break-even.

Similarly, the model is quite sensitive to the assumptions pertaining to the comparator. At baseline it was assumed that people with substance misuse do not receive a long term prison-based intervention. However, assuming that standard care costs are £300 per person (equivalent to approximately 10 sessions with a prison-based psychologist at £33 per hour) increases the cost savings associated with a prison based therapeutic community treatment to £743 per person.

There is also high uncertainty surrounding the duration of a prison sentence estimate. At base-case analysis it was assumed that the prison sentence would be approximately 16.2 months. When reducing the duration of a prison sentence to 14 months (the minimum sentence required for an individual to be able to complete therapeutic community treatment) the intervention results in the cost savings of £272 and when it is increased, for example, to

23 months the therapeutic community treatment results in the cost savings of £1,002 per person.

Assuming that therapeutic community treatment is financed by the NHS the QALY gain would need to be 0.041 for the therapeutic community treatment to be considered cost-effective (that is, for the cost per QALY to be below NICE's lower cost-effectiveness threshold of £20,000 per QALY); plus, there would be £1,258 per person savings to the MoJ. The required QALY gain of 0.041 is relatively small and would be equivalent to an individual being 15 days in full health over the duration of the model.

7.2.21.1094 Conclusions

Based on the above findings it seems that therapeutic community treatment is likely to be cost saving from a CJS perspective. The therapeutic community treatment for substance misuse had higher intervention costs but resulted in savings in CJS costs. Even though prison therapeutic community treatment for substance misuse is intensive treatment, most of it is delivered in a group setting resulting in relatively low per person costs. According to the GC, good mental health care in prisons could potentially reduce health care costs when people are released back into the community. Due to the unavailability of appropriate data, it was impossible to quantify such cost savings. Nevertheless, sensitivity analyses indicated that assuming that therapeutic community treatment is provided by the NHS a relatively small QALY gain would be required for the intervention to be considered cost effective (that is, for a cost per QALY to be below NICE's lower cost effectiveness threshold of £20,000 per QALY).

The cost analysis is characterised by a number of limitations, including efficacy data from 1 RCT that was conducted in the US. The comparator in the RCT was an active intervention. However, in the model the comparator was 'no treatment'. This was due to the lack of suitable data on the standard care treatment in the UK. This is likely to have underestimated the cost savings associated with the therapeutic community treatment in the model. Also, in the economic analysis therapeutic community treatment was modelled as approximately 270 hours of treatment (mostly group), whereas the efficacy was derived from an RCT where therapeutic community treatment was much more intense. However, the GC reviewed the study and concluded that the structure of the therapeutic community treatment in the RCT (Sacks 2004) is similar to that provided in the UK. Given the lack of better data on the effectiveness the GC expressed their view that this RCT should provide a reasonable approximation. Also, the baseline re-incarceration rate used in the model is for the general prison population and it is likely to be much higher in the population with mental health and substance misuse problems underestimating the cost savings associated with the therapeutic community treatment. Overall, the GC judged this analysis to provide a conservative estimate of cost savings associated with the therapeutic community treatment.

The analysis has considered only intervention costs and the resource use data was based on the GC expert opinion. Due to the unavailability of suitable data re-incarceration costs included only prison stay costs and hasn't considered costs associated with the imprisonment (such as, police and court costs). Excluding such costs have potentially underestimated the cost savings associated with the therapeutic community treatment. Also, it was modelled that the intervention will be delivered by prison officers. The estimate of prison officer salary included only basic salary and hasn't considered salary on-costs, qualification costs, and overheads. However, the threshold analysis indicated that the intervention cost could increase by as much as 54% for therapeutic community treatment and 'no treatment' option costs to break-even.

Notwithstanding the above limitations, this analysis indicates that therapeutic community treatment may potentially be cost saving and cost-effective treatment option for people with substance misuse problems in prison settings. There is a need for further research on effectiveness and cost-effectiveness of therapeutic community treatment for substance

misuse in the UK and in particular on assessing its effect on future health outcomes and associated costs.

7.2.22 Clinical evidence statements

7.2.224.1 Street triage

Very low quality of evidence from one before-after study with nine street triage scheme (n=200,000) showed that there was no difference in the total number of detentions per 100,000 people under section 136 between before and after street triage.

Low quality evidence from two before-after studies (n=49914) showed clinically important difference that street triage pilot scheme effectively reduced the number of detentions under section 136 in custody whereas very low quality evidence from three before-after studies (n=49953) reported clinically important increase in the number of detentions under section 136 in health board places of safety (a desirable outcome) with street triage scheme.

7.2.2232 Diversion Services

Diversion services versus No diversion services

Very low quality of evidence from two before and after studies (*n*=611) indicated uncertainty about the difference between before and after court diversion for duration of stay between remand and mental health assessment. However, one before and after study (n=565) of very low quality suggested clinically significant difference that total time on remand in days was reduced among participants after diversion programme compared with before the programme.

Very low quality evidence from one controlled cohort study (n=220) indicated uncertainty about the effectiveness of assessment by a doctor or nurse compared to no assessment before appearing at magistrate courts in terms of proportion of prisoners on bail release, attendance at alcohol and drug treatment programmes, OPD attendance rate for those on bail and registration of care programmes and supervision registration.

Court diversion versus Community diversion

Evidence from one controlled cohort study (n=428) indicated clinically important difference that the rate of re-incarceration within 2 years after discharge from hospital was higher whereas the 100% attendance rate of appointment was lower among participants in court diversion programme than those in community diversion programme. However, there was no difference in the number of days in hospital and the number of diverted participants with no mental health disorders between court and community diversion services. The evidence was of low to very low quality.

7.2.2253 Patient Navigation intervention

Very low quality evidence from one RCT (*N*=18) reported that there was no difference in the number of participants who used drugs, those who used alcohol to intoxication and average days when mental health was not good in the last 30 days between those in patient navigation intervention and those in facilitated enrolment groups.

7.2.2.214 Neighbourhood outreach

Very low quality evidence from one before and after observational study (N=506) showed clinically important difference that there was a decrease in proportion of crime contacts with policing team escalated to court between before and after neighbourhood outreach.

7.2.2.225 Drug Rehabilitation Program

Very low quality evidence from one controlled cohort study (N=52) showed that there was no clinically significant different between DRR (formerly DTTO) and TAU (mainstream services) for Maudsley Addiction Profile (MAP) total scores and Health of National Outcome Scales (HoNOS) scores whereas there was a clinically significant effect of DRR compared with TAU for overall satisfaction.

7.2.2.216 Case Management

Case management versus treatment as usual for substance misuse disorders

Treatment effects were not clinically important for re-arrest [at post-treatment [1RCT; N=504) and at 3-months follow-up (1RCT; N=462)], re-incarceration [at post-treatment (1RCT; N=504), at 3-months follow-up (1RCT; N=462) and at 12-months follow-up (1RCT; N=862)] and reconviction at post-treatment (1RCT; N=504). The quality was very low to low.

Very low to low quality evidence suggested that the treatment effect for self-reported alcohol use was not clinically significant during treatment (1RCT; N=288), post-treatment (1RCT; N=680) and 12-months follow-up (1RCT; N=862). However, there was clinically important difference at 12 month follow-up where either men or women in the case management condition were less likely to report alcohol use than those in the treatment as usual conditions; this effect was much larger for women than men (female sample: RR=0.18 [0.07, 0.50]; *N*=154; male sample: RR=0.83 [0.70, 0.99]; *N*=708). The quality of evidence was of moderate for female samples and low for male samples.

Very low to low quality evidence suggested that treatment effects for self-reported drug use (marijuana or hard drugs) during treatment (1RCT; N=288), post-treatment (1RCT; N=680) and 12-months follow-up [1RCT; either male (N=708) or female (N=154)] were not clinically significant for self-reported drug use. However, low quality evidence from one RCT (*N*=862) suggested clinically important difference that at 12 month follow-up, total participants (both male and female samples) in the case management condition were less likely to report drug use than those in the treatment as usual conditions. Similarly, one RCT (N=462) of very low quality reported no clinical difference in injection drug use between case management and treatment as usual at post-treatment.

Very low to low quality evidence reported for no clinically different effect in abstinence at either during treatment (1RCT; N=283) or at post-treatment (1RCT; N=462) between case management and treatment as usual.

Case management versus active intervention among participants with substance misuse disorders

Very low quality evidence from one RCT (*N*=369) reported clinically important difference that those in the case management condition were more likely to remain in treatment than those in the urine testing condition.

There was very low quality evidence from one RCT (N=369) for no clinically important difference for re-arrest, re-conviction and re-incarceration between case management plus urine monitoring and urine monitoring only at post-treatment.

There was very low quality evidence from one RCT (N=511) for no clinically important difference for re-arrest for any crime, re-arrest for drug crime, re-conviction, re-incarceration, any self-reported drug use and positive hair test for either crack/cocaine or marijuana between case management plus intensive discharge planning and discharge planning only at 3-months follow-up.

Assertive community treatment versus treatment as usual among participants with substance misuse disorders

Very low quality evidence from one RCT (*n*=119) suggested no clinically importance difference between assertive community treatment and treatment as usual for positive urine test for drug use, self-reported injection drug use, self-reported drug use and re-incarceration during treatment.

Case management versus treatment as usual among participants with mental health disorders other than substance misuse

Very low quality evidence from 2 RCTs (N=223) reported no clinically important difference between case management and treatment as usual for service utilization rates at post-treatment.

Similarly, very low quality evidence from 3 RCTs (N=432) suggested no clinically important difference between case management and treatment as usual for re-offending rates at post-treatment.

Very low quality evidence from 2 RCTs (N=369) reported clinically significant difference that participants in case management group stayed shorter duration in jail than those in TAU according to up to 24-months follow-up data.

Very low quality evidence from 1 RCT (N=92) reported no clinically important difference in the quality of life between assertive community treatment plus mental health treatment court (MHTC) and MHTC only at post-treatment.

7.2.2.267 Drug Courts

Drug court versus treatment as usual

Very low quality evidence from one RCT (*N*=157) reported clinically important difference that those in the drug court condition were less likely to be arrested and committed less serious crimes as measured by the maximum crime seriousness scale than those in the treatment as usual condition at 12-months follow-up. Similarly, moderate quality evidence from the same RCT suggested clinically significant difference that the number of days of substance use (alcohol or cocaine or heroin) were reduced among those in the drug court than those in the treatment as usual at 12-months follow-up. Low to moderate quality evidence of treatment effects for attrition in gender responsive drug court relative to drug court as usual were not clinically significant.

Drug court versus active intervention

There was very low to low quality evidence from one RCT (*N*=150) for no clinically important difference between gender responsive drug court and drug court as usual for number of participants being removed from treatment due to unsatisfactory progress, number of sanctions and number of sanctions resulting in jail diversion at post-treatment.

Very low quality evidence from one RCT (N=62) suggested no clinically important difference between engaging mum drug course and intensive case management drug court for alcohol

and drug composite scores measured by ASI and number of drug positive urine tests at post-treatment.

There was very low quality evidence from one RCT (N=62) for no clinically important difference between drug court plus intensive judicial supervision and drug court as usual for re-incarceration at post-treatment.

7.2.2.8 Case Management and Opioid Substitution Therapy

Very low to low quality evidence from one RCT (N=211) reported no clinically important difference between methadone with case management and case management alone for completed jail treatment.

There was very low quality evidence from one RCT for no clinically important difference between methadone plus case management and case management for cocaine positive urine test at 1 and 6 months follow-up (N=200 and N=76). However, the effect was clinically significant at 12 months follow-up (N=115) with reduction in cocaine positive urine test among participants with counselling plus methadone with financial assistance.

Similarly, very low to low quality evidence from one RCT suggested clinically significant reduction in opioid positive urine test with methadone plus case management at 1 month follow-up (N=200) and 12 month follow-up (N=115), in comparison to case management only. However, the effect was not significant at 6 month follow-up (N=57) and it was of very low quality evidence.

Very low quality evidence from one RCT (N=204) reported that there was no clinically important difference between methadone plus case management and case management only for average days of cocaine or heroin use at 12-months follow-up. On the other hand, one RCT of very low to low quality (N=62) reported clinically significant reduction in the risk of self-reported heroin use in past 30 days at 6 month follow-up although the effect was not significant in the risk of self-reported crack/cocaine or marijuana or injection drug use.

Very low quality evidence reported no clinically important difference between management plus methadone therapy and case management only for drug overdose and re-arrest at 6 month follow-up (N=62)and 12 month follow-up (N=204). Similarly, low quality evidence from one RCT (N=204) reported no clinically important difference in self-reported days of criminal activity between case management plus methadone therapy or case management only.

7.2.2.9 Automated Telephony

Low quality evidence from one RCT (N=108) suggested clinically important effect of automated telephony with feedback for depressive symptoms measured by SCL-8D or daily stressor assessment while moderate quality evidence reported no clinically important difference for depressive symptoms measured by AHSS questionnaires with automated telephony with feedback, as relative to automated telephony alone. Similarly, low quality evidence found clinically important difference effect for reduction in alcohol use or drug use although no clinical effect for reduction in alcohol urge (moderate quality evidence) or drug urge (low quality evidence) with automated telephony treatment.

7.2.2.10 Integrated Disorders Treatment Program (IDDT)

Low quality evidence from one RCT (N=182) suggested clinically important difference effect of an increase in rate of outpatient medication services with IDDT as relative to TAU. Moreover, there was no clinically important difference in number of days in hospital and rate of crisis visits between IDDT and TAU.

7.2.22.11 Housing First

Low quality evidence from one randomized study (n=297) reported clinical important effect of housing first for any offence rate as relative to treatment as usual. Looking at the breakdown figures between scattered housing first with ACT and congregate housing first, it was suggested from moderate quality evidence for clinical important difference with scattered HF plus ACT whereas very low quality found no clinical important difference in comparison with treatment as usual.

7.2.22.12 Texas Implementation of Medication Algorithm

There was low quality evidence from one RCT (n=60) for no clinical important effect of Texas Implementation of Medical Algorithm (TIMA) as compared to treatment as usual for bipolar and psychiatric symptoms measured by BDSS and BPRS respectively.

7.2.22.13 Service Brokerage Intervention

Very low quality evidence from one RCT (*N*=1325) reported that there was no clinical important difference in the number of participants who were in contact with mental health services, who had seen GP and who attended alcohol or drug service between those in service brokerage intervention and those in TAU groups.

7.2.22.14 Therapeutic Communities for substance misuse

Therapeutic communities versus wait-list controls

Very low quality evidence from one RCT (N=341) suggested no clinically important difference in the number of days until re-incarceration between therapeutic communities versus wait-list control.

Modified Therapeutic communities versus active intervention

Very low quality evidence from one RCT (N=139) reported clinically important difference between modified therapeutic communities and CBT-informed psychoeducation for reduction in substance use, alcohol use, drug use and criminal activity at 12-months follow-up.

Likewise, very low to low quality evidence from one RCT (N=139) suggested clinically important difference between prison modified therapeutic communities with or without aftercare versus mental health program only for reduction in the rate of re-incarceration and alcohol/drug offence.

Enhanced therapeutic community versus standard therapeutic community

Low to very low quality evidence from 1 RCT (N=451) showed clinically important difference between enhanced therapeutic communities and standard therapeutic communities on decreased negative mood as rated by counsellor. However, there was no difference in treatment engagement between enhanced and standard therapeutic communities.

Gender responsive therapeutic community versus standard therapeutic community

Very low quality evidence from 1 RCT (N=115) showed clinically important difference with gender-responsive therapeutic communities on increased time spent in aftercare and increased time to re-incarceration as relative to standard therapeutic communities. However, there was no clinically important difference for drug or alcohol use as well as psychological improvement and self-efficacy measured by ASI, aftercare participation rate upon release, disciplinary removal rates, re-incarceration rates and voluntary drop-out rates between two groups.

Gender-specific therapeutic community versus psychoeducation

Low quality evidence from one RCT (N=314) suggested no clinical difference between gender specific therapeutic communities and CBT-informed psychoeducation for mental health symptoms measured by BDI or BSI or PSS scales.

Very low quality evidence from two RCTs (N=702) reported clinical important effect of gender-specific therapeutic community on self-reported any criminal activity at 6-months follow-up. However, the effect was not significant at 12-months follow-up. Moreover, there was no difference in criminal activities related to drugs at 6-month (2RCTs; N=702) and 12-months (1 RCT; N=370) follow-up as well as sexually related criminal activity (1 RCT; N=314) at post-treatment.

Very low quality evidence from one RCT (N=314) suggested no clinically important difference between gender specific therapeutic community and CBT-informed psychoeducation for receiving mental health and substance abuse treatment at follow-up.

Very low to low quality evidence also reported no clinically important difference between gender specific therapeutic community and CBT-informed psychoeducation for alcohol use at follow-up (1 RCT; N=314), self-reported drug use at 6-months follow-up (2 RCTs; N=702), frequency of alcohol use (1 RCT; N=162) and frequency of drug use (1 RCT; N=206) at follow-up.

Very low quality evidence from one RCT (N=388) suggested clinically important difference between gender specific therapeutic community and CBT-informed psychoeducation for re-arrest rates at 6-months follow-up. However, the effect was not significant at 12-weeks follow-up. Similarly, very low quality evidence from another RCT (N=314) reported no difference in re-arrest rates at follow-up. Moreover, one RCT (N=468) of very low quality reported no difference in re-incarceration rates between the two groups.

Re-entry modified therapeutic communities versus treatment as usual

Very low quality evidence from one RCT (N=127) showed clinically important difference between re-entry modified therapeutic communities and treatment as usual for decreased re-incarceration rates, criminal activity and alcohol/drug offence at 12-months post prison release.

7.2.23 Economic evidence statements

7.2.23.1 Jail diversion

The evidence from 1 UK cost-utility analysis based on economic modelling found diversion plus treatment and/or aftercare programme when compared with no diversion to be dominant (that is, it resulted in lower public sector costs and a greater QALY gain). The sensitivity analyses indicated a high level of uncertainty about the parameter estimates used. Given the limitations with the data and high uncertainty in the results GC found it difficult to draw any conclusions about the cost-effectiveness of diversion for adult substance abusing offenders who come into contact with the CJS. The study was characterised only by minor methodological limitations (some of the model inputs being based on assumptions).

The remainder of the evidence is from the US and Canada. Three studies found diversion to be cost-saving from the public sector perspective. The conclusions from the remainder of the studies were unclear. Generally, in these studies the diversion resulted in higher public sector costs but also improvements on various scales (such as the BPRS and the Wisconsin quality of life scale). However, since none of the health outcomes were expressed in QALYs it was difficult for the GC to assess whether improvements in health outcomes were adequate to justify the increase in the public sector costs. With the exception of 1 US study, which was characterised by minor methodological limitations (model inputs based on a single

published study), the rest of the studies were characterised by potentially serious methodological limitations including short time horizons, being based on observational study designs, and having small study samples.

7.2.23.2 Mental health courts

There was evidence from 1 cost-effectiveness analysis and 1 cost analysis conducted in the US. The cost-effectiveness analysis was based on an observational cohort study (N=150) and found mental health court programme to be dominant for successful participants using residential and jail days, and prison days as outcome measures. However, the cost-difference was not significant. The cost analysis was based on an observational before-after study (N=365) and found that mental health court programme may potentially be cost saving. Both studies were conducted in the US and are only partially applicable to the NICE decision-making context, and both are characterised by potentially serious limitations, including study designs, small study samples and the lack of use of national unit costs.

7.2.23.3 Drug court programmes

There was evidence from 3 US cost analyses based on cohort studies (N=1,944; N=1,173; N=745). All 3 studies found drug court programmes to be cost-saving when compared with no such programmes in adults with substance abuse problems from a public sector perspective. There is also evidence from an Australian cost-effectiveness analysis based on an RCT (N=468). It found drug court programme when compared with no drug court programme to be dominant from a public sector perspective (that is, it resulted in lower costs and also better outcomes [it took longer to the first drug-related offence and there were fewer drug-related offences per day]). This is non-UK evidence so it is only partially applicable to the NICE decision-making context. In addition, no QALYs were measured. All studies are characterised by potentially serious limitations, including their study design (3 were observational cohort studies), lack of consideration of health outcomes, and lack of reporting of statistical significance levels.

7.2.23.4 Street triage

There was evidence from 2 UK studies. One cost analysis was based on an observational before-after study (N=99,412 for street triage, N=688,654 for the rest of the county) and decision analytic modelling and found that from the NHS and criminal justice sector, as well as from a criminal justice sector perspective only, street triage was cost-saving, but from the NHS perspective only street triage was associated with a slight increase in costs. Another cost analysis was also based on an observational cohort study (N=55) and decision-analytic modelling. It found that street triage, conducting Mental Health Act assessments for all Section 136 detainees, and having a link worker present at custody suites only marginally increased public sector costs. This evidence, although derived from 2 UK studies, is partially applicable to the NICE decision-making context as studies did not consider health outcomes and did not estimate QALYs. Both studies are characterised by potentially serious limitations, including short time horizons, 1 study had a very small study sample, and some model inputs being based on authors' assumptions.

7.2.23.5 Integrated Disorders Treatment Program (IDDT)

There was evidence from 1 economic analysis conducted alongside an RCT (N=182) in the US. It found that integrated treatment when compared with standard care resulted in an increase in health care costs but there was a reduction in arrests, convictions, and jail days over 12 months. However, there was an increase in felony convictions. This evidence was derived from a US study and is only partially applicable to the NICE decision-making context. It did not report outcomes in the form of QALYs so judgements on cost effectiveness were difficult to make, and is characterised by potentially serious limitations, including a short time

horizon, the consideration of mental health costs only (for total costs from a public sector perspective no comparable outcomes were reported), and the use of local unit costs

7.2.233 Forensic assertive community treatment (FACT)

There was evidence from 1 US (N=134) cost-effectiveness analysis. FACT resulted in an increase in public sector costs and in a reduction in bookings, jail days, and convictions when compared with treatment as usual defined as services routinely available in the county-operated public behavioural health system. However, health outcomes were not considered and QALYs were not estimated which made it difficult for the GC to draw any conclusions pertaining to the cost-effectiveness of FACT in adult detainees with serious mental illness with co-occurring substance abuse problems. The study was only partially applicable to the NICE decision-making context, and is characterised by potentially serious limitations, including the relatively short time horizon (2 years), the lack of consideration of health outcomes, and the fact that resource use data were based on local administrative data.

7.2.234 Therapeutic community treatment

There was evidence from 3 existing cost-effectiveness analyses of prison-based therapeutic community treatment for substance misuse. All 3 economic evaluations were conducted in the US and were based on RCTs (N=836, N=715, N=576). In 2 studies, the work release component of therapeutic community treatment (when compared with SC or no treatment) resulted in an ICER of $113-34 per day of incarceration avoided, and the therapeutic community treatment plus aftercare combined resulted in an ICER of $19-51 per day of incarceration avoided (when compared with the work release component only). In another study, prison therapeutic community treatment only was dominated by the SC treatment, and prison based therapeutic treatment and post-release care combined (when compared with SC) resulted in an ICER of $48 per additional incarceration day avoided. The GC considered the above ICERs and concluded that therapeutic community treatment may potentially be cost effective for the treatment of substance misuse in prisons. This evidence is US-based and is only partially applicable to the NICE decision-making context and is characterised by potentially serious limitations. None of the evaluations considered health outcomes, and wider health care, and social care costs, 2 studies adopted time horizon of less than 2 years, and the source of unit costs was unclear in all studies.

A cost analysis conducted for this guideline found that therapeutic community treatment for substance misuse delivered in prison setting may potentially be cost saving when compared with ‘no treatment’ alternative. The therapeutic community treatment results in higher intervention costs, but it is associated with the reduction in re-incarcerations and associated reduction in the criminal justice sector costs. The cost analysis is only partially applicable to the NICE decision-making context since it has not considered health outcomes and has not estimated QALYs. Due to the lack of the relevant data the perspective of the criminal justice sector only was adopted. The analysis was characterised by potentially serious limitations including efficacy data from a single US-based study, and resources use data based on US study and GC expert opinion.

7.2.235 Probation and mandated treatment

There was evidence from 2 US-based cost analyses. One cost analysis was based on a large observational cohort study (intervention N=47,355; control N=41,607) and found probation and mandated treatment when compared with the SC to be cost saving at 30 months from a public sector perspective. Another cost analysis was based on an RCT (N=272) and modelling. The intervention resulted in a cost increase at 2.75 year follow up from a public sector perspective. This evidence was derived from the US and is only partially applicable to the NICE decision-making context. One study is characterised by minor limitations and the other by potentially serious limitations including the lack of consideration

of health outcomes, the estimation of the relative treatment effects from observational studies (1 from a large cohort study), and the resource use data were obtained from a mixture of local and national sources.

7.2.2349 Medium security units

There was evidence from 1 cost-effectiveness analysis based on an observational cohort study (N=54). From a public sector perspective community and residential service was dominant when compared with an inpatient medium secure unit and a residential service for personality-disordered male offenders. This evidence, although derived from a UK study, is partially applicable to the NICE decision-making context as it did not report outcomes in the form of QALYs. The measure of outcome was an improvement on The Work and Social Adjustment Scale, which made interpretation of the results difficult. This study is characterised by very serious limitations, including the study design (very small cohort study), and lack of reporting of statistical significance levels for costs and outcomes. Due to its very serious limitations, this study was not considered by the GC when making recommendations.

7.3 Recommendations and link to evidence

Recommendations	**49. Practitioners should consider referral to a therapeutic community specifically for substance misuse for people in prison with a minimum 18-month sentence who have an established pattern of drug misuse.** **50. When setting up therapeutic community programmes in prison settings in a separate wing of a prison for people with substance misuse problems, aim to:** • **include up to 50 prisoners in the programme** • **provide treatment for between 12 and 18 months, made up of:** o twice-weekly group therapy sessions (mean group size of 8) o daily (5 days only) community meeting for all wing residents o daily (5 days only) social activity groups for all wing residents o a once-weekly individual review meeting (20 minutes). **51. Commissioners and providers of criminal justice services and healthcare services should consider developing systems for police custody and court custody that provide prompt access to the following:** • **the effective identification and recognition of mental health problems** • **a comprehensive mental health assessment** • **advice on immediate care and management** **52. Providers of criminal justice services and healthcare services should consider diverting people from standard**

	courts to dedicated drug courts if the offence is linked to substance misuse and was non-violent. **53. Commissioners and providers of criminal justice services and healthcare services should consider establishing joint working arrangements between healthcare, social care and police services for managing urgent and emergency mental health presentations in the community (for example, street triage). Include:** • **joint training for police, healthcare and social care staff** • **agreed protocols for joint working developed and reviewed by a multi-agency group** • **agreed protocols for effective communication within and between agencies** • **agreed referral pathways for urgent and emergency care and routine care.** **54. Commissioners and providers of criminal justice services and healthcare services should ensure effective identification, assessment, coordination and delivery of care for all people with a mental health problem in contact with the criminal justice system (including the National Probation Service or Community Rehabilitation Company). In particular, ensure that:** • **all people with a severe or complex mental health problem have a designated care coordinator** • **during transitions between services care plans are shared and agreed between all services** • **effective protocols are in place to support routine data sharing between health and criminal justice agencies to reduce unnecessary duplication of assessments.**
Relative values of different outcomes	The GC considered more effective service utilisation to be the critical outcome. Service level interventions aimed to provide changes in service designed which through increased access or the more appropriate use of services lead to better mental health outcomes, reduced reoffending and improved treatment engagement to be important outcomes in this area. This combination of benefits is likely to be particularly efficacious as greater engagement with treatment would be expected to have a positive impact upon mental health and reoffending, and consequently upon long-term service use, adaptive functioning and quality of life. Critical outcomes also varied across service level interventions for example Street Triage aimed to reduce the number of people admitted under section 136 to a Health Based Place of Safety.
Trade-off between clinical benefits and harms	**Street triage –** The GC noted that the evidence was of low quality and drawn from cohort studies and also there was not one specific model for street triage scheme developed, but the benefits (for example, reduced use of s136, increased use of Health Based Places of Safety, increased access to mental health treatments) were reported in services which shared a number of common characteristics. The GC did not identify any harms associated with the model. Thus, drawing on the available evidence and their own knowledge and experience the GC made a recommendation based around the key characteristics which the GC saw as underpinning effective Street Triage models identified in the studies.

	Other UK service delivery systems – the GC reviewed a number of other service delivery models which stressed increased prompt access and coordination of care e.g. the neighbourhood outreach program in Cornwall which linked together both street triage and court diversion. Drug Rehabilitation Requirements (DRR), mental health courts, drug courts and custody diversion and liaison services all had some limited low quality evidence to suggest reductions in re-offending rates but often with no effect on people's mental health and social functioning. The GC considered that these interventions taken together (they had a broadly shared objective of diverting people from management by the criminal justice system into more appropriate health care settings) these programmes seemed obtain better engagement for individuals with mental health services, often diverting them away from expensive and potentially distressing criminal procedures, with a positive impact on offending. The GC could not identify any significant hams associated with these approaches. **Case management** - the very low quality evidence from a number of RCTs which focused in the majority of cases on people with substance misuse did not demonstrate a clear or consistent benefit in either service utilisation or clinical outcomes for case management across the range of people contact with the criminal justice system. However, given the importance placed on the coordination of care and their knowledge of a high drop-out rate for treatment in people in contact with the criminal justice system the GC decided to make a recommendation based on the available evidence which using informal consensus they extrapolated to apply to all people in contact with the criminal justice system for the coordination of care. Again drawing on their knowledge and experience they also recommended a number of key components of care management which they believed were associated with improved engagement in services. **Therapeutic communities -** therapeutic communities to be a clinically, although intensive, intervention, in particular for people with a significant history of drug misuse with evidence from 7 RCTs. The impact was demonstrated on offending and on mental health measure with some indication of a reduction in service utilisation in some studies. The GC were mindful of the relationship between substance misuse and offending and considered the impact The GC discussed the fact that as it is a long-term intervention that they recommended treatment only for those with significant drug abuse problems who have a minimum 18-month sentence. They considered the fact that the included studies followed a similar format with a combination of individual and group work and a range of day time activities and they used this information to develop advice of the delivery of the intervention as they wanted to ensure fidelity to the model (duration of treatment, frequency of group therapy and individual reviews). There do not appear to be any significant clinical harms associated with therapeutic communities.
Trade-off between net health benefits and resource use	**Developing systems for police custody and court custody that provide prompt access to effective identification, comprehensive assessment, and advice on immediate care and management** The GC considered limited UK evidence showing that street triage is cost-saving or may only marginally increase public sector costs. Also, according to the GC developing systems for street triage and police custody and court custody that provide prompt access to the effective identification of mental health problems, a comprehensive assessment, and advice on immediate care and management may have resource implications in terms of the extra time required to facilitate these service structures. However, the GC expressed the view that if such service structures lead to prompt identification of mental health needs and that this results in subsequent treatment and management of any mental health problems at an early stage,

before individuals require more resource intensive management, then the additional costs associated with facilitating such service structures might be expected to result in improved mental health outcomes in the longer term with potential future cost savings to the healthcare system (delays in treatment exacerbate symptoms) and criminal justice system (improvement in mental health may prevent future reoffending) that outweigh the costs associated with facilitating such service systems.

Diverting people from standard courts to dedicated drug courts

There was non-UK evidence that drug court programmes are potentially cost-saving when compared with standard courts in adults with substance abuse problems from a public sector perspective. There was also evidence from a non-UK cost-effectiveness analysis that found a drug court programme when compared with no drug court programme to be dominant from a public sector perspective (that is, it resulted in lower costs and also better outcomes [it took longer to the first drug-related offence and there were fewer drug-related offences per day]).

The GC considered existing economic evidence and the economic consequences arising from the presence of substance misuse in people who are in contact with the criminal justice system. The GC considered an increase in the incidence of people in this population and the additional pressure it imposes on healthcare and criminal justice sectors. The GC also considered the pressure on the facilities such service users place and the high costs of imprisonment. For example, in the UK to keep an individual in prison costs as much as £34,087 per annum, but the data suggests that someone going through the drug court programme would incur only a fraction of this cost. Moreover, the GC considered the cyclical relationship between the drugs and non-violent crime, and the fact that many offenders have frequent interactions with the criminal justice sector. This significantly increases public sector costs associated with the sentencing and potential imprisonment costs. Moreover, standard judicial process takes a considerable time, and the GC considered that this population would significantly benefit from early, prompt access to appropriate treatment. The GC considered that if drug courts result in a prompt subsequent treatment and management of any mental health problems, before they require more resource intensive management, then drug courts might be expected to result in improved mental health outcomes in the longer term and potential future cost savings to the healthcare system (delays in treatment exacerbate symptoms) and criminal justice system (improvement in mental health may prevent future reoffending) that outweigh the costs associated with the provision of drug court programmes.

Facilitating joint working arrangements between healthcare, social care and police services for managing urgent and emergency mental health presentations in the community

There was evidence from one UK study indicating that establishing joint working arrangements between healthcare, social care and police services for managing urgent and emergency mental health presentations in the community (that is, Street Triage) was cost-saving from NHS and criminal justice sector perspective, but not from the NHS perspective only. Similarly, another cost analysis found that Street Triage, conducting Mental Health Act assessments for all Section 136 detainees, and having a link worker present at custody suites only marginally increased public sector costs.

The GC considered the above existing economic evidence and also expressed the view that such working arrangements stop people entering the criminal justice system inappropriately and ending up in custody. Also, according to the GC expert opinion, custody is used far too frequently and underlying mental health problems may not be addressed. The GC

	expressed the view that if such joint working arrangements lead to a better care, prompt identification of mental health needs and individuals are signposted more quickly to the appropriate services before they require more resource intensive management (such as, expensive crisis care), then the additional costs associated with facilitating such service structures might be expected to result in improved mental health outcomes in the longer term and potential future cost savings to the healthcare system (delays in treatment exacerbate symptoms that may require expensive crisis care) and criminal justice system (improvement in mental health may prevent future reoffending) that outweigh the costs associated with facilitating such joint working arrangements. **Therapeutic community treatment** There was evidence from three existing US cost-effectiveness analyses of prison-based therapeutic community treatment for substance misuse. All economic evaluations found the therapeutic community treatment (plus aftercare) to be potentially cost-effective. Economic analysis conducted for this guideline indicated that therapeutic community treatment when compared with 'no treatment' alternative resulted in an increase in costs but also in a reduction in re-offending rates and associated criminal justice sector costs. Sensitivity analysis indicated that when assuming that therapeutic community treatment is funded by the NHS the QALY gain required for the intervention to be considered cost effective (that is, to result in a cost per QALY below NICE's lower cost-effectiveness threshold of £20,000) would need to be relatively small. The GC considered the ICERs associated with the therapeutic community treatment (from existing studies) together with the findings from the cost analysis conducted for this guideline and concluded that therapeutic community treatment may potentially be cost effective for the treatment of substance misuse in prisons. This evidence is only partially applicable to the NICE decision-making context and is characterised by potentially serious limitations. None of the evaluations considered health outcomes, and wider health care, and social care costs; 2 existing economic studies adopted time horizon of less than 2 years, and the source of unit costs was unclear in all existing studies. **Effective identification, assessment, coordination and delivery of care for all people with a mental health problems in contact within the criminal justice system (having designated care coordinator, care plans are shared and agreed between all services, effective protocols are in place)** There was no evidence on the cost effectiveness of having service structures such as, having designated care coordinator for people with severe and complex mental health problems, making sure that care plans are shared and agreed between all services during transitions between services, and having effective protocols in place to support routine data sharing. However, the GC expressed their view that if such service structures lead to better care and improvements in treatment and management of people with mental health problems who are in contact with criminal justice system then the additional costs associated with facilitating such service structures might be expected to result in improved mental health outcomes in the longer term and potential future cost savings to the healthcare system (delays in treatment exacerbate symptoms) and criminal justice system (improvement in mental health may prevent future reoffending) that outweigh the costs associated with facilitating such service structures.
Quality of evidence	The GC members were aware that the majority of RCTs reviewed were from non-UK and given the importance of the wider health care environment in influencing the outcome of service level interventions the GC went down the evidence hierarchy to review observational studies in UK settings. The GC also noted that several of the trials in therapeutic communities had all-

	female cohorts, which was of particular interest as women are listed within the protocol as a group to receive special consideration during the guideline development process. **Street Triage** – The quality of evidence was from very low to low. The majority of the evidence came from an observational study of a national evaluation of nine pilot schemes across England. **Other UK service delivery systems** – The evidence was of low to very low quality. In most cases, the evidence came from a single observational studies leading to uncertainty about the benefits of the interventions. **Drug courts** – The quality of evidence was very low. The outcomes reported and drug court programs were different trials and could not meta-analyse the data and the RCTs included had small sample sizes. **Therapeutic communities -** The evidence ranged in quality from very low to moderate. However, these RCTs were generally based on relatively large population sizes and had reasonable effect sizes.
Other considerations	**Therapeutic communities** The GC had concerns about whether the funding of the programme would fall under the remit of the Health Department or NOMS. The GC also expressed the view that fully realising the benefits of the service level intervention reviewed in this question require clear criteria for access, effective communication and defined roles and responsibility of all those providing services involved. The GC suggested that this would be best achieved through the establishment of care pathways for these populations. The GC were also concerned about the poor coordination of care experienced by many people in contact with the criminal justice system and therefore decided to develop a research recommendation to identify the best models to support effective care coordination for people with mental health problems in contact with the criminal justice system.

7.321 Research recommendations

8. What models for the coordination and delivery of care for people in contact with the criminal justice provide for the most effective and efficient coordination of care and improve access and uptake of services?

There is low quality evidence for a range of systems for the delivery and coordination of care in the criminal justice system (for example drug or mental health courts, and case management). However, there is clear evidence of poor engagement, uptake and retention in treatment for people with mental health problems in contact with the criminal justice system. A number of models for example, case management and collaborative care have shown benefit for people with common and sever mental health problems in routine healthcare settings. A programme of research and development is required which will (a) develop and test in small feasibility studies different models of care coordination for the delivery of care and (b) test those models which have shown promise in the feasibility studies in large scale randomised clinical trials in the criminal justice system.

Important outcomes could include:

- Improved access and uptake of services
- Improved mental health outcomes
- Reductions in offending and reoffending

1

8 Abbreviations

ACT	Acceptance and Commitment Therapy
ADHD	attention deficit hyperactivity disorder
AGREE	Appraisal of Guidelines for Research and Evaluation Instrument
ASD	autism spectrum disorder
AUC	area under the curve
CBT	cognitive behavioural therapy
CI	confidence interval
CPN	community psychiatric nurse
DSM(-III, -IV, -5, -R, -TR)	Diagnostic and Statistical Manual of Mental Disorders (3rd edition, 4th edition, 5th edition, Revised, Text Revision)
Embase	Excerpta Medica Database
GAD	generalised anxiety disorder
GC	Guideline Committee
GP	general practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HRQoL	health-related quality of life
HTA	Health Technology Assessment
ICD-10	International Statistical Classification of Diseases and Related Health Problems – 10th revised edition
ICER	incremental cost-effectiveness ratio
IQ	intelligence quotient
k	number of studies (K=Kappa statistics)
MD	mean difference
MEDLINE	Medical Literature Analysis and Retrieval System Online
n	number of participants
N	total number of participants
n/a	not applicable
n/r	Not reported
NCCMH	National Collaborating Centre for Mental Health
NGA	National Guideline Alliance
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NoMs	National Offender Management Service
OCD	obsessive–compulsive disorder
OIS	optimal information size
OR	odds ratio
PCL(-R, -SV)	Psychopathy Checklist (– Revised, -Screening Version)
PICO	Population, Intervention, Comparison and Outcome
PsycINFO	Psychological Information Database
PTSD	post-traumatic stress disorder
QALY	quality-adjusted life year
QUADAS-II	Quality Assessment of Diagnostic Accuracy Studies - Revised
RCT	randomised controlled trial
RQ	review question
RR	risk ratio
SD	standard deviation
SE	standard error
SMD	standardised mean difference

2

9 References

Abracen J, Looman J, Ferguson M, Harkins L, Mailloux D. Recidivism among treated sexual offenders and comparison subjects: Recent outcome data from the Regional Treatment Centre (Ontario) high-intensity Sex Offender Treatment Programme. Journal of Sexual Aggression. 2011;17:142-52.

AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Quality and Safety in Health Care. 2003;12:18-23.

Alan J, Burmas M, Preen D, Pfaff J. Inpatient hospital use in the first year after release from prison: a Western Australian population-based record linkage study. Australian and New Zealand Journal of Public Health. 2011;35:264-9.

Alemi F, Haack M, Nemes S, Harge A, Baghi H. Impact of online counseling on drug use: a pilot study. Quality management in health care. 2010.[cited;19: Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/972/CN-00732972/frame.html.

Alemi F, Taxman FS, Baghi H, Vang J, Thanner M, Doyon V. Costs and benefits of combining probation and substance abuse treatment. Journal of Mental Health Policy and Economics. 2006;9:57-70.

Alphs L, Benson C, Cheshire-Kinney K, Lindenmayer JP, Mao L, Rodrigues SC, et al. Real-world outcomes of paliperidone palmitate compared to daily oral antipsychotic therapy in schizophrenia: a randomised, open-label, review-board-blinded 15 months study. Journal of clinical psychiatry. 2015;76:554-61.

Anderson Varney TJ. An evaluation of a treatment program for imprisoned child sex offenders. 1991.[cited:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/545/CN-00735545/frame.html.

Andersson C, Vasiljevic Z, Hoglund P, Ojehagen A, Berglund M. Daily automated telephone assessment and intervention improved 1-month outcome in paroled offenders. International Journal of Offender Therapy and Comparative Criminology. 2014

Anglin MD, Nosky B, Jaffe A, Urada D, Evans E. Offender diversion into substance use disorder treatment: the economic impact of California's proposition 36. American journal of public health. 2013;103:1096-102.

Annis HM. Group treatment of incarcerated offenders with alcohol and drug problems: A controlled evaluation. Canadian Journal of Criminology. 1979.[cited;21:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/387/CN-00426387/frame.html.

Ascher-Svanum H, Nyhuis A, Faries D, Ball D, Kinon B. Involvement in the US criminal justice system and cost implications for persons treated for schizophrenia. . BMC psychiatry. 2010;10.

Aytes KE, Olsen SS, Zakrajsek T, Murray P, Ireson R. Cognitive/behavioral treatment for sexual offenders: an examination of recidivism. Sexual abuse : a journal of research and treatment. 2001;13:223-31.

Baksheev GN, Ogloff J, Thomas S. Identification of mental illness in police cells: A comparison of police processes, the Brief Jail Mental Health Screen and the Jail Screening Assessment Tool. Psychology, Crime & Law. 2012;18:529-42.

Barnett GD, Manderville-Norden R, Rakestrow J. The Good Lives Model or Relapse Prevention: What Works Better in Facilitating Change? Sexual Abuse: Journal of Research and Treatment. 2014;26:3-33.

1 Bates A, Williams D, Wilson C, Wilson RJ. Circles South East: The first 10 years 2002-2012.
2 International Journal of Offender Therapy and Comparative Criminology. 2014;58:861-85.

3 Battaglia C, di Cagno A, Fiorilli G, Giombini A, Borrione P, Baralla F, et al. Participation in a
4 9-month selected physical exercise programme enhances psychological well-being in a
5 prison population. Crim Behav Ment Health. 2015;25:343-54.

6 Bechelli M, Caudy M, Gardner T, Huber A, Mancuso D, Samuels P, et al. Case studies from
7 three states: breaking down silos between health care and criminal justice. Health Affairs.
8 2014;33:474-81.

9 Beggs SM, Grace RC. Assessment of dynamic risk factors: An independent validation study
10 of the violence risk scale: Sexual offender version. Sexual Abuse: Journal of Research and
11 Treatment. 2010;22:234-51.

12 Berlin JA. Does blinding of readers affect the results of meta-analyses? . Lancet.
13 2001;350:185-86.

14 Berman AH, Lundberg U, Krook AL, Gyllenhammar C. Treating drug using prison inmates
15 with auricular acupuncture: a randomized controlled trial. Journal of substance abuse
16 treatment. 2004.[cited;26:Available from:
17 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/676/CN-00467676/frame.html.

18 Bernstein SJ, Laouri M, Hilborne LH, Leape LL, Kahan JP, Park R, et al. Coronary
19 angiography: a literature review and ratings of appropriateness and necessity. Santa Monica,
20 USA: RAND; 1992.

21 Biggam FH, Power KG. A controlled, problem-solving, group-based intervention with
22 vulnerable incarcerated young offenders. International Journal of Offender Therapy and
23 Comparative Criminology. 2002.[cited;46:Available from:
24 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/330/CN-00412330/frame.html.

25 Bilderbeck AC, Farias M, Brazil IA, Jakobowitz S, Wikholm C. Participation in a 10-week
26 course of yoga improves behavioural control and decreases psychological distress in a
27 prison population. Journal of psychiatric research. 2013.[cited;47:Available from:
28 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/700/CN-00991700/frame.html.

29 Binswanger IA, Whitley E, Haffey PR, Mueller SR, Min SJ. A patient navigation intervention
30 for drug-involved former prison inmates. Substance abuse. 2015;36:34-41.

31 Bond GR, Kim SJ, Becker DR, Swanson SJ, et al. A controlled trial of supported employment
32 for people with severe mental illness and justice involvement. Psychiatric Services.
33 2015;66:1027-34.

34 Booth A. Chapter 6: Formulating research questions. In: Booth A, Brice A, eds. Evidence
35 based practice: A handbook for information professionals. London: Facet; 2003.

36 Bradley RG, Follingstad DR. Group therapy for incarcerated women who experienced
37 interpersonal violence: a pilot study. Journal of traumatic stress. 2003.[cited;16:Available
38 from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/244/CN-
39 00474244/frame.html.

40 Bradley. L. The Bradley Report. Lord Bradley's review of people with mental health problems
41 or learning disabilities in the criminal justice system 2009.

42 Brown RA. Conventional education and controlled drinking education courses with convicted
43 drunken drivers. Behavior therapy. 1980.[cited;11:Available from:
44 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/309/CN-00180309/frame.html.

Carey SM, Finigan MW. A detailed cost analysis in a mature drug court setting: A cost-benefit evaluation of the Multnomach County Drug Court. Journal of Contemporary Criminal Justice. 2004;20:315-38.

Carroll KM, Easton CJ, Nich C, Hunkele KA, Neavins TM, Sinha R, et al. The use of contingency management and motivational/skills-building therapy to treat young adults with marijuana dependence. Journal of consulting and clinical psychology. 2006.[cited;74:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/741/CN-00572741/frame.html.

Carroll KM, Nich C, Lapaglia DM, Peters EN, Easton CJ, Petry NM. Combining cognitive behavioral therapy and contingency management to enhance their effects in treating cannabis dependence: less can be more, more or less. Addiction (Abingdon, England). 2012.[cited;107:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/284/CN-00854284/frame.html.

Centre for Mental Health. Diversion. A better way for criminal justice and mental health. 2009.

Centre for Mental Health., Rethink., Royal College of Psychiatrists. Diversion: the business case for action 2010.

Chambers C, Rix KJ. A controlled evaluation of assessments by doctors and nurses in a magistrates' court mental health assessment and diversion scheme. Medicine, science, and the law. 1999;39:38-48.

Chandler DW, Spicer G. Integrated treatment for jail recidivists with co-occuring psychiatric and substance use disorders. Community mental health journal. 2006;42:405-25.

Cheesman FL, Graves SE, Holt K, Kunkel TL, Lee CG, White MT. Drug court effectiveness and efficiency: Findings for Virginia. Alcoholism treatment quarterly. 2016;34:143-69.

Chen XJ, Hannibal N, Gold C. Randomised trial of group music therapy with Chinese prisoners. Impact on anxiety, depression, self-esteem. International Journal of Offender Therapy and Comparative Criminology. 2015;12.

Clayton A, O'Connell MJ, Bellamy C, Benedict P, Rowe M. The Citizenship Project part II: impact of a citizenship intervention on clinical and community outcomes for persons with mental illness and criminal justice involvement. American journal of community psychology. 2013.[cited;51:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/721/CN-00918721/frame.html.

Cochrane Collaboration. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration; 2014.

Coid JW, Yang M, Ullrich S, Hickey N, Kahtan N, Freestone M. Psychiatric diagnosis and differential risks of offending following discharge. International Journal of Law & Psychiatry. 2015;38:68-74.

Cole KL, Sarlund-Heinrich P, Brown L. Developing and assessing effectiveness of a time-limited therapy group for incarcerated women survivors of childhood sexual abuse. Journal of trauma & dissociation. 2007.[cited;8:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/897/CN-00611897/frame.html.

Cornish JW, Metzger D, Woody GE, Wilson D, McLellan AT, Vandergrift B, et al. Naltrexone pharmacotherapy for opioid dependent federal probationers. Journal of substance abuse treatment. 1997.[cited;14:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/946/CN-00146946/frame.html.

Cosden M, Ellens J, Schnell J, Yamini-Diouf Y. Efficacy of a Mental Health Treatment Court with assertive community treatment. Behavioral sciences & the law. 2005.[cited;23:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/389/CN-00521389/frame.html.

Cosden M, Ellens JK, Schnell JL, Yamini-Diouf Y, Wolfe MM. Evaluation of a mental health treatment court with assertive community treatment. Behavioral sciences & the law. 2003.[cited;21:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/758/CN-00439758/frame.html.

Coviello DM, Cornish JW, Lynch KG, Alterman AI, O'Brien CP. A randomized trial of oral naltrexone for treating opioid-dependent offenders. American journal on addictions / American Academy of Psychiatrists in Alcoholism and Addictions. 2010.[cited;19:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/056/CN-00770056/frame.html.

Cowell AJ, Broner N, Dupont R. The cost-effectiveness of criminal justice diversion programs for people with serious mental illness co-occuring with substance abuse four case studies. Journal of Contemporary Criminal Justice. 2004;20:292-315.

Cowell AJ, Hinde JM, Broner N, Aldridge AP. The impact on taxpayer costs of a jail diversion program for people with serious mental illness. Evaluation and Program Planning. 2013;41.

Cox J, Morschauser P, Banks S, Stone J. A five-year population study of persons involved in the mental health and local correctional systems: Implications for service planning. The Journal of Behavioral Health Services & Research. 2001;28:177–87.

Craissati J, McClurg G. The Challenge Project: A treatment program evaluation for perpetrators of child sexual abuse. Child Abuse and Neglect. 1997;21:637-48.

Craissati J, South R, Bierer K. Exploring the effectiveness of community sex offender treatment in relation to risk and re-offending. Journal of Forensic Psychiatry and Psychology. 2009;20:769-84.

Crane CA, Eckhardt CI, Schlauch RC. Motivational enhancement mitigates the effects of problematic alcohol use on treatment compliance among partner violent offenders: Results of a randomised clinical trial. Journal of consulting and clinical psychology. 2015b;83:689-95.

Cropsey KL, Lane PS, Hale GJ, Jackson DO, Clark CB, Ingersoll KS, et al. Results of a pilot randomized controlled trial of buprenorphine for opioid dependent women in the criminal justice system. Drug and alcohol dependence. 2011.[cited;119:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/423/CN-00843423/frame.html.

Curtis L, Burns A. Unit costs of health and social care 2015. The University of Kent: Personal Social Services Research Unit; 2015.

Cusack KJ, Morrissey JP, Cuddeback GS, Prins A, Williams DM. Criminal justice involvement, behavioural health service use and costs of forensic assertive community treatment: A randomised trial. Community mental health journal. 2010;46.

Cutcher Z, Degenhardt AR, Kinner SA. Poor health and social outcomes for ex-prisoners with a history of mental disorder: a longitudinal study. Socioeconomic status and health. 2014b;38:424-29.

Czuchry M, Dansereau DF. Cognitive skills training: impact on drug abuse counseling and readiness for treatment. American journal of drug and alcohol abuse. 2003.[cited;29:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/695/CN-00558695/frame.html.

1 Dakof GA, Cohen JB, Henderson CE, Duarte E, Boustani M, Blackburn A, et al. A
2 randomized pilot study of the Engaging Moms Program for family drug court. Journal of
3 substance abuse treatment. 2010.[cited;38:Available from:
4 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/042/CN-00750042/frame.html.

5 Daley M, Love CT, Shepard DS, Petersen CB, White KL, Hall FB. Cost-effectiveness of
6 Connecticut's In-Prison Substance Abuse Treatment. Journal of Offender Rehabilitation.
7 2004;39:69-92.

8 Davis TM, Baer JS, Saxon AJ, Kivlahan DR. Brief motivational feedback improves post-
9 incarceration treatment contact among veterans with substance use disorders. Drug and
10 alcohol dependence. 2003.[cited;69:Available from:
11 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/041/CN-00417041/frame.html.

12 Dennis JA, Khan O, Ferriter M, Huband N, Powney MJ, Duggan C. Psychological
13 interventions for adults who have sexually offended or are at risk of offending. Cochrane
14 Database Syst Rev. 2012;12:CD007507.

15 Department of Health. Statistical update on suicide Health Analytical Team. In: HMSO,
16 editor. London2014.

17 Di Fazio R, Abracen J, Looman J. Group versus individual treatment of sex offenders: A
18 comparison. Forum on Corrections Research. 2001;13:56-59.

19 Dolan KA, Shearer J, MacDonald M, Mattick RP, Hall W, Wodak AD. A randomised
20 controlled trial of methadone maintenance treatment versus wait list control in an Australian
21 prison system. Drug and alcohol dependence. 2003.[cited;72:Available from:
22 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/129/CN-00462129/frame.html.

23 Dolan KA, Shearer J, White B, Zhou J, Kaldor J, Wodak AD. Four-year follow-up of
24 imprisoned male heroin users and methadone treatment: mortality, re-incarceration and
25 hepatitis C infection. Addiction (Abingdon, England). 2005.[cited;100:Available from:
26 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/728/CN-00511728/frame.html.

27 Donato R, Shanahan M. The economics of child sex-offender rehabilitation programs:
28 beyond Prentky & Burgess. American journal of Orthopsychiatry. 2001;71:131-39; discussion
29 40-41.

30 Drake R, Mueser K. Psychosocial Approaches to Dual Diagnosis. Schizophrenia Bulletin.
31 2000;26 105-18.

32 Duwe G. Can Circles of Support and Accountability (COSA) work in the United States?
33 Preliminary results from a randomized experiment in Minnesota. Sexual abuse.
34 2013.[cited;25:Available from:
35 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/007/CN-00919007/frame.html.

36 Duwe G, Goldman RA. The impact of prison-based treatment on sex offender recidivism:
37 Evidence from Minnesota. Sexual Abuse: Journal of Research and Treatment. 2009;21:279-
38 307.

39 Dyfed Powys Police and Powys Teaching Health Board. Powys Sec 136 Mental Health
40 Authorisation Pilot Twelve Month Review. Bwrdd lechyd Addysgu Powys 2015 April 2014 -
41 March 2015.

42 Earl F, Cocksedge K, Rheeder B, Morgan J, Palmer J. Neighbourhood outreach: a novel
43 approach to Liaison and Diversion. Journal of Forensic Psychiatry and Psychology.
44 2015;26:573-85.

Easton C, Swan S, Sinha R. Motivation to change substance use among offenders of domestic violence. Journal of substance abuse treatment. 2000.[cited;19:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/770/CN-00297770/frame.html.

Easton CJ, Mandel DL, Hunkele KA, Nich C, Rounsaville BJ, Carroll KM. A cognitive behavioral therapy for alcohol-dependent domestic violence offenders: an integrated substance abuse-domestic violence treatment approach (SADV). American journal on addictions / American Academy of Psychiatrists in Alcoholism and Addictions. 2007c.[cited;16:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/266/CN-00579266/frame.html.

Eccles M, Freemantle N, Mason J. North of England evidence based guideline development project: methods of developing guidelines for efficient drug use in primary care. British Medical Journal. 1998;316:1232-35.

Ehret MJ, Shelton D, Barta W, Trestman R, Maruca A, Kamath J, et al. Medication adherence among female inmates with bipolar disorder: results from a randomized controlled trial. Psychological services. 2013.[cited;10:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/473/CN-00919473/frame.html.

Evans C, Psych M, Brinded P, Simpson AI, Frampton C, Mulder RT. Validation of brief screening tools for mental disorders among New Zealand prisoners. Psychiatric Services. 2010.

Exworthy T, Parrott J. Comparative evaluation of a diversion from custody scheme. Journal of Forensic Psychiatry. 1997;8:406-16.

Fazel S, Yu R. Psychotic disorders and repeat offending: systematic review and meta-analysis. Schizophrenia Bulletin. 2011;37:800-10.

Ford JD, Chang R, Levine J, Zhang W. Randomized clinical trial comparing affect regulation and supportive group therapies for victimization-related PTSD with incarcerated women. Behavior therapy. 2013.[cited;44:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/797/CN-00875797/frame.html.

Ford JD, Trestman RL, Wiesbrock V, Zhang W. Development and validation of a brief mental health screening instrument for newly incarcerated adults. Assessment. 2007;14:279-99.

Ford JD, Trestman RL, Wiesbrock VH, Zhang W. Validation of a brief screening instrument for identifying psychiatric disorders among newly incarcerated adults. Psychiatric Services. 2009;60:842-46.

Forsberg LG, Ernst D, Sundqvist K, Farbring C. Motivational Interviewing delivered by existing prison staff: a randomized controlled study of effectiveness on substance use after release. Substance use & misuse. 2011B.[cited;46:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/722/CN-00810722/frame.html.

Fortune Z, Barrett B, Armstrong D, Coid J, Crawford M, Mudd D, et al. Clinical and economic outcomes from the UK pilot psychiatric services for personality-disordered offenders. International review of psychiatry. 2011;23:61-69.

Friedmann PD, Green TC, Taxman FS, Harrington M, Rhodes AG, Katz E, et al. Collaborative behavioral management among parolees: drug use, crime and re-arrest in the Step'n Out randomized trial. Addiction (Abingdon, England). 2012.[cited;107:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/974/CN-00832974/frame.html.

Friedmann PD, WIlson D, Knudsen H, et al. Effect of an Oranizational Linkage Intervention for staff perceptions of medication-assisted treatment and referral intentions in community corrections. Journal of substance abuse treatment. 2015;50:50-58.

Friendship C, Mann RE, Beech AR. Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. Journal of Interpersonal Violence. 2003;18:744-59.

Frottier P, Koenig F, Seyringer M, Matschnig T, Fruehwald S. The distillation of "VISCI": Towards a better identification of suicidal inmates. Suicide and Life-Threatening Behavior. 2009;39:376-85.

George DT, Phillips MJ, Lifshitz M, Lionetti TA, Spero DE, Ghassemzedeh N, et al. Fluoxetine treatment of alcoholic perpetrators of domestic violence: a 12-week, double-blind, randomized, placebo-controlled intervention study. Journal of clinical psychiatry. 2011.[cited;72:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/150/CN-00770150/frame.html.

Ginsberg Y, Hirvikoski T, Grann M, Lindefors N. Long-term functional outcome in adult prison inmates with ADHD receiving OROS-methylphenidate. European archives of psychiatry and clinical neuroscience. 2012b.[cited;262:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/902/CN-00859902/frame.html.

Ginsberg Y, Lindefors N. Methylphenidate treatment of adult male prison inmates with attention-deficit hyperactivity disorder: randomised double-blind placebo-controlled trial with open-label extension. British Journal of Psychiatry. 2012a.[cited;200:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/063/CN-00882063/frame.html.

Gisev N, Shanahan M, Weatherburn DJ, Mattick RP, Larney S, Burns L, et al. A cost-effectiveness analysis of opioid substitution therapy upon prison release in reducing mortality among people with a history of opioid dependence. Addiction. 2015:No Pagination specificed.

Goodman A, Ruggiero V. Crime, Punishment, and Ethnic Minorities in England and Wales Race/Ethnicity: Multidisciplinary Global Contexts. 2008;2:53-68.

Gordon MS, Kinlock TW, Schwartz RP, Fitzgerald TT, O'Grady KE, Vocci FJ. A randomized controlled trial of prison-initiated buprenorphine: Prison outcomes and community treatment entry. Drug and alcohol dependence. 2014.[cited;142:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/518/CN-00998518/frame.html.

Gordon MS, Kinlock TW, Schwartz RP, O'Grady KE. A randomized clinical trial of methadone maintenance for prisoners: findings at 6 months post-release. Addiction (Abingdon, England). 2008.[cited;103:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/566/CN-00701566/frame.html.

Gottfredson DC, Kearley BW, Najaka SS, Rocha CM. The Baltimore City Drug Treatment Court: 3-year self-report outcome study. Evaluation review. 2005.[cited;29:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/612/CN-00503612/frame.html.

Gowin JL, Green CE, Alcorn JL, Swann AC, Moeller FG, Lane SD. Chronic tiagabine administration and aggressive responding in individuals with a history of substance abuse and antisocial behavior. Journal of psychopharmacology (Oxford, England). 2012.[cited;26:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/448/CN-00843448/frame.html.

Grubin D, Carson D, Parsons S. Report on the New Prison Reception Health Screen and the Results of a Pilot Study in 10 Prisons. London: HM Prison Service; 2002.

Gussak D. The effects of art therapy on male and female inmates: advancing the research base. The Arts in Psychotherapy. 2008;36:5-12.

1 Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1.
2 Introduction-GRADE evidence profiles and summary of findings tables. Journal of Clinical
3 Epidemiology. 2011;64:383-94.

4 Guydish J, Chan M, Bostrom A, Jessup MA, Davis TB, Marsh C. A randomised trial of
5 probation case management of drug-involved women offenders. Crime & Delinquency.
6 2011;57:167-98.

7 Hakvoort L, Bogaerts S, Thaut MH, Spreen M. Influence of music therapy on coping skills
8 and anger management in forensic psychiatric patients: an exploratory study. International
9 Journal of Offender Therapy and Comparative Criminology2013

10 Hall SM, Loeb P, Coyne K, Cooper J. Increasing employment in ex-heroin addicts I: criminal
11 justice sample. Behavior therapy. 1981.[cited;12:Available from:
12 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/009/CN-00183009/frame.html.

13 Hanlon TE, Bateman RW, O'Grady KE. The relative effects of three approaches to the parole
14 supervision of narcotic addicts and cocaine abusers. Prison journal. 1999.[cited;79:Available
15 from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/884/CN-
16 00691884/frame.html.

17 Hanlon TE, McCabe OL, Savage C, Kurland AA. Narcotic antagonist treatment of addict
18 parolees--the failure of an effective approach. Comprehensive psychiatry.
19 1977.[cited;18:Available from:
20 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/488/CN-00241488/frame.html.

21 Hanson R, Steffy RA, Gauthier R. Long-term recidivism of child molesters. Journal of
22 consulting and clinical psychology. 1993;61:646-52.

23 Hanson RK, Broom I, Stephenson M. Evaluating community sex offender treatment
24 programs: A 12-year follow-up of 724 offenders. . Canadian Journal of Behavioural Sciences.
25 2004; 36:87-96.

26 Hanson RK, Helmus L, Thornton D. Predicting recidivism amongst sexual offenders: a multi-
27 site study of static-2002. Law and Human Behavior. 2010;34:198-211.

28 Hanson RK, Thornton D. Improving risk assessments for sex offenders: A comparison of
29 three actuarial scales. Law and Human Behavior. 2000;24:119-36.

30 Harkins L, Flak VE, Beech AR, Woodhams J. Evaluation of a Community-Based Sex
31 Offender Treatment Program Using a Good Lives Model Approach. Sexual Abuse: Journal of
32 Research and Treatment. 2012;24:519-43.

33 Harrison KS, Rogers R. Axis I screens and suicide risk in jails: A comparative analysis.
34 Assessment. 2007;14:171-80.

35 Hartwell S. Short-term outcomes for offenders with mental illness released from
36 incarceration. International Journal of Offender Therapy and Comparative Criminology.
37 2003;47:145–58.

38 Hayhurst KP, Leitner M, Davies L, Flentje R, Millar T, Jones A, et al. The effectiveness and
39 cost-effectiveness of diversion and aftercare programmes for offenders using class a drugs:
40 A systematic review and economic evaluation. Health Technology Assessment. 2015;19:1-
41 198.

42 Helmus L, O Ciardha C, Seto MC. The Screening Scale for Pedophilic Interests (SSPI):
43 Construct, predictive, and incremental validity. Law and Human Behavior. 2015;39:35-43.

Heslin M, Callagham L, Barrett B, Susan L, Eick S, Morgan J, et al. Costs of the police service and mental healthcare pathways experienced by individuals with enduring mental health needs. The British Journal of Psychiatry. 2016b:1-8.

Heslin M, Callagham L, Packwood M, Badu V, Byford S. Decision analytic model exploring the cost and cost-offset implications of street triage. BMJ open. 2016a;6:1-11.

Higgins JP, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011. Available from http://www.cochrane-handbook.org. ; 2011.

Hopkins R. An evaluation of communication and social skills groups for sex offenders at HMP Frankland. Psychology Conference of Prison Service. 1991.[cited:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/844/CN-00592844/frame.html.

Horton M, Wright N, Dyer W, Wright-Hughes A, Farrin A, Mohammed Z, et al. Assessing the risk of self-harm in an adult offender population: An incidence cohort study. Health Technology Assessment. 2014;18:i-xxv+1-151.

Howells C, Allen S, Gupta J, Stillwell G, Marsden J, Farrell M. Prison based detoxification for opioid dependence: a randomised double blind controlled trial of lofexidine and methadone. Drug and alcohol dependence. 2002.[cited;67:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/669/CN-00397669/frame.html.

Hucker S, Langevin R, Bain J. A double blind trial of sex drive reducing medication in pedophiles. Sexual Abuse: A Journal of Research and Treatment. 1988; 1:227-42.

Hughes D, Steadman HJ, Case B, Griffin PA, Leff H. A simulation modeling approach for planning and costing jail diversion programs for persons with mental illness. Criminal Justice and Behaviour. 2012;39:434-46.

Institute of Medicine. Crossing the Quality of Chasm: A New Health System for the 21st Century. Washington DC: National Academy Press; 2001.

Ivanoff A, Jang SJ. The role of hopelessness and social desirability in predicting suicidal behavior: A study of prison inmates. Journal of consulting and clinical psychology. 1991;59:394-99.

Jadad A, Moore R, Carroll D, Jenkinson C, Reynolds D, Gavaghan D, et al. Assessing the quality of reports of randomised clinical trials: is blinding necessary? Controlled Clinical Trials. 1996;17:1-12.

James DV, Farnham F, Moorey H, al. e. Outcome of Psychiatric Admission through the courts. London: Home Office; 2002.

Jarrett M, Thornicroft G, Forrester A, Harty M, Senior J, King C, et al. Continuity of care for recently released prisoners with mental illness: a pilot randomised controlled trial testing the feasibility of a Critical Time Intervention. Epidemiology and psychiatric sciences. 2012.[cited;21:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/018/CN-00852018/frame.html.

Johnson JE, Friedmann PD, Green TC, Harrington M, Taxman FS. Gender and treatment response in substance use treatment-mandated parolees. Journal of substance abuse treatment. 2011.[cited;40:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/749/CN-00787749/frame.html.

Johnson JE, Zlotnick C. Pilot study of treatment for major depression among women prisoners with substance use disorder. Journal of psychiatric research. 2012.[cited;46:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/322/CN-00845322/frame.html.

Jones CGA. Early-phase outcomes from a randomised trial of intensive judicial supervision in an Australian drug court. Criminal Justice and Behavior. 2013;40:453-68.

Kent J, Gunasekaran S. Mentally disordered detainees in the police station: the role of the psychiatrist Advances in Psychiatric Treatment. 2010;16:115-23.

Khan O, Ferriter M, Huband N, Powney MJ, Dennis JA, Duggan C. Pharmacological interventions for those who have sexually offended or are at risk of offending. Cochrane Database Syst Rev. 2015:CD007989.

Kingston DA, Seto MC, Firestone P, Bradford JM. Comparing indicators of sexual sadism as predictors of recidivism among adult male sexual offenders. Journal of consulting and clinical psychology. 2010;78:574-84.

Kinlock TW, Gordon MS, Schwartz RP, Fitzgerald TT, O'Grady KE. A randomized clinical trial of methadone maintenance for prisoners: results at 12 months postrelease. Journal of substance abuse treatment. 2009.[cited;37:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/870/CN-00720870/frame.html.

Kinlock TW, Gordon MS, Schwartz RP, O'Grady K, Fitzgerald TT, Wilson M. A randomized clinical trial of methadone maintenance for prisoners: results at 1-month post-release. Drug and alcohol dependence. 2007.[cited;91:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/265/CN-00701265/frame.html.

Kinner SA, Dooren K, Boyle FM, et al. Development of an intervention to increase health service utilization in ex-prisoners. Health and Justice. 2014a;2:1-8.

Kinner SA, Lennox N, Williams GM, et al. Randomised controlled trial of a service brokerage intervention for ex-prisoners in Australia. Contemporary Clinical Trials. 2013;36:198-206.

Konefal J, Duncan R, Clemence C. The impact of the addition of an acupuncture treatment program to an existing metro-Dade County outpatient substance abuse treatment facility. Journal of addictive diseases. 1995.[cited;13:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/563/CN-00113563/frame.html.

Konstenius M, Jayaram-Lindstrom N, Guterstam J, Beck O, Philips B, Franck J. Methylphenidate for ADHD and Drug Relapse in Criminal Offenders with Substance Dependence: A 24-week Randomized Placebo-controlled Trial. Addiction (Abingdon, England). 2013.[cited;109:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/591/CN-00874591/frame.html.

Krameddine YI, DeMarco D, Hassel R, Silverstone PH. A novel training program for police officers that improves interactions with mentally ill individuals and is cost-effective. Frontiers in Psychiatry. 2013:4-9.

Kubiak S, Roddy J, Comartin E, Tillander E. Cost analysis of long-term outcomes of an urban mental health court. Evaluation and Program Planning. 2015;52:96-106.

Lamb H, Weinberger L, Gross B. Persons in the criminal justice system: Some perspectives. The Psychiatric Quarterly. 2004;75:107–26.

Langevin R, Paitich D, Hucker S, Newman S, Ramsay G, Pope S. The effect of assertiveness training, Provera and sex of therapist in the treatment of genital exhibitionism. . Journal of Behavior Therapy and Experimental Psychiatry. 1979;10:275-82.

Lanza PV, Garcia PF, Lamelas FR, Gonzalez-Menendez A. Acceptance and commitment therapy versus cognitive behavioral theapy in the treatment of substance use disorder with incarcerated women. Journal of clinical psychology. 2014;70:644-57.

Lee J, Friedmann P, Kinlock T, Nunes E, Boney T, Hoskinson R, et al. Extended-release Naltrexone to prevent opioid relapse in criminal justice offenders. The New England Journal of Medicine. 2016;374:1232-42.

Lee JD, McDonald R, Grossman E, al. e. Opioid treatment at release from jail using extended-release naltrexone: a pilot proof-of-concept randomized effectiveness trial. Addiction. 2015;110:1008-14.

Lin I, Muser E, Munsell M, Benson C, Menzin J. Economic impact of psychiatric relapse and recidivism among adults with schizophrenia recently released from incarceration: a Markov model analysis. Journal of Medical Economics. 2015;18:219-29.

Lobmaier PP, Kunøe N, Gossop M, Katevoll T, Waal H. Naltrexone implants compared to methadone: outcomes six months after prison release. European addiction research. 2010.[cited;16:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/216/CN-00762216/frame.html.

Logan T, Hoyt WH, McCollister KE, French MT, Leukefeld C, Minton L. Economic evaluation of drug court: methodology, results and policy implications. Evaluation and Program Planning. 2004;27:381-96.

Looman J, Abracen J, Nicholaichuk TP. Recidivism among treated sexual offenders and matched controls: Data from the Regional Treatment Centre (Ontario). . Journal of Interpersonal Violence. 2000; 15:279-90.

Loper AB. Improving the emotional adjustment and communication patterns of incarcerated mothers: effectiveness of a prison parenting intervention. Journal of Child and Family studies. 2011;20:89-101.

Louden JE, Skeem JL, Blevins A. Comparing the predictive utility of two screening tools for mental disorder among probationers. Psychological Assessment. 2013;25:405-15.

Lowden K, Hetz N, Harrison L, English K, Pasini-Hill D. Evaluation of Colorado's Prison Therapeutic Community for Sex Offenders: A Report of Findings. 2003. p. 143.

Magura S, Lee JD, Hershberger J, Joseph H, Marsch L, Shropshire C, et al. Buprenorphine and methadone maintenance in jail and post-release: a randomized clinical trial. Drug and alcohol dependence. 2009.[cited;99:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/336/CN-00687336/frame.html.

Marques IK, Day DM, Nelson C, West Mary A. Effects of cognitive-behavioral treatment on sex offender recidivism: Preliminary results of a longitudinal study. Special Issue: The assessment and treatment of sex offenders. Criminal-Justice-and-Behavior. 1994a.[cited;21:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/832/CN-00241832/frame.html.

Marshall LE, Marshall WL, Fernandez YM, Malcolm PB, Moulden HM. The rockwood preparatory program for sexual offenders: Description and preliminary appraisal. Sexual Abuse: Journal of Research and Treatment. 2008;20:25-42.

Marshall WL, Barbaree HE. The long-term evaluation of a behavioral treatment program for child molesters. Behaviour research and therapy. 1988b;26:499-511.

Marshall WL, Eccles A, Barbaree HE. The treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. Behaviour research and therapy. 1991;29:129-35.

Martin MS, Wamboldt AD, O'Connor SL, Fortier J, Simpson AIF. A comparison of scoring models for computerised mental health screening for federal prison inmates. Criminal Behaviour and Mental Health. 2013;23:6-17.

Maunder L, Cameron L, Moss M, Muir D, Evans N, Paxton R, et al. Effectiveness of self-help materials for anxiety adapted for use in prison a pilot study. Journal of mental health (Abingdon, England). 2009.[cited;18:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/627/CN-00754627/frame.html.

McCollister KE, French MT, Inciardi JA, Butzin CA, Martin SS, Hooper RM. Post-release substance abuse treatment for criminal offenders: a cost-effectiveness analysis. Journal of Quantitative Criminology. 2003a;19:389-407.

McCollister KE, French MT, Prendergast M, Wexler H, Sacks S, Hall E. Is In-prison treatment enough? A cost-effectiveness analysis of prison-based treatment and aftercare services for substance-abusing offenders. Law & Policy. 2003b;25:63-82.

McCollister KE, French MT, Prendergast ML, Hall E, Sacks S. Long-term cost-effectiveness of addiction treatment for criminal offenders. Justice Quarterly. 2004;21:569-679.

McConaghy N, Armstrong MS, Blaszczynski A. Expectancy, covert sensitization and imaginal desensitization in compulsive sexuality. Acta psychiatrica Scandinavica. 1985;72::176–87.

McConaghy N, Blaszczynski A, Kidson W. Treatment of sex offenders with imaginal desensitization and/or medroxyprogesterone. Acta psychiatrica Scandinavica. 1988.[cited;77:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/535/CN-00053535/frame.html.

McGrath RJ, Cumming G, Livingston JA, Hoke SE. Outcome of a treatment program for adult sex offenders: From prison to community. Journal of Interpersonal Violence. 2003;18:3-17.

McGrath RJ, Cumming GF, Hoke SE, Bonn-Miller MO. Outcomes in a community sex offender treatment program: A comparison between polygraphed and matched non-polygraphed offenders. Sexual Abuse: Journal of Research and Treatment. 2007;19:381-93.

McGrath RJ, Hoke SE, Vojtisek JE. Cognitive-behavioral treatment of sex offenders. Criminal Justice and Behavior. 1998;25:203-25.

McGuire TJ. Correctional institution based sex offender treatment: A lapse behavior study. Behavioral Sciences and the Law. 2000;18:57-71.

McKenzie M, Zaller N, Dickman SL, Green TC, Parihk A, Friedmann PD, et al. A randomized trial of methadone initiation prior to release from incarceration. Substance abuse : official publication of the Association for Medical Education and Research in Substance Abuse. 2012.[cited;33:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/918/CN-00832918/frame.html.

McKinnon I, Grubin D. Evidence-Based Risk Assessment Screening in Police Custody: The HELP-PC Study in London, UK. Policing. 2014;8:174-82.

McKinnon I, Thorp J, Grubin D. Improving the detection of detainees with suspected intellectual disability in police custody. Advances in Mental Health and Intellectual Disabilities. 2015;9:174-85.

McPherson W. Managing the mental health population at the Broward's Sheriff's office. Corrections Today. 2008;70:62–67.

Menting ATA, Castro BO, Wijngaards-de Meij L, Matthys W. A Trial of Parent Training for Mothers Being Released From Incarceration and Their Children. Journal of clinical child and adolescent psychology. 2014.[cited;43:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/623/CN-00998623/frame.html.

Messina N, Calhoun S, Warda U. Gender-responsive drug court treatment: a randomised controlled trial. Criminal Justice and Behavior. 2012;39:1539-58.

Messina N, Grella CE, Cartier J, Torres S. A randomized experimental study of gender-responsive substance abuse treatment for women in prison. Journal of substance abuse treatment. 2010.[cited;38:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/683/CN-00732683/frame.html.

Miller PM. A behavioral intervention program for chronic public drunkenness offenders. Archives of general psychiatry. 1975.[cited;32:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/176/CN-00705176/frame.html.

Ministry of Justice. Statistical Notice. Ministry of Justice policy statement on revisions.: 2009.[cited:Available from.

Ministry of Justice. An overview of sexual offending in England and Wales. Home Office and Office for National Statistics2013a.

Ministry of Justice. Transforming rehabilitation. A strategy to reform. 2013b.

Ministry of Justice. Costs per place and costs per prisoner. National offender management service annual report and accounts 2013-14 Management Information Addendum: Ministry of Justice Information Release; 2014.

Ministry of Justice. Criminal Justice Statistics Quarterly Update to June 2015. England and Wales 2015a.

Ministry of Justice. Offender management statistics prison population 2015. 2015b.

Ministry of Justice. Proven Reoffending Statistics Quarterly Bulletin July 2013 to June 2014. England and Wales2016a.

Ministry of Justice. Criminal justice statistics 2015. Ministry of Justice bulletin. England and Wales2016b.

Ministry of Justice., HM Prison Service. Population bulletin: monthly December 2015 2016.

Mitton C, Simpson L, Gardner L, Barnes F, McDougall G. Diversion Program: A community-based alternative to incarceration for mentally ill offenders. Journal of Mental Health Policy and Economics. 2007;10:145-51.

Mokros A, Schilling F, Eher R, Nitschke J. The Severe Sexual Sadism Scale: cross-validation and scale properties. Psychological Assessment. 2012;24:764-69.

Moll A. Losing track of time: Dementia and the ageing prison population: treatment challenges and examples of good practice. London: Mental Health Foundation; 2013.

Morgan D. Powys Section 136 Mental Health Evaluation November 2015 Summative Evaluation Report. Cyfarwyddiaeth Gwybodaeth & Chudd Wybodaeth 2015.

Murphy MK, Black NA, Lamping D, McKee C, Sanderson C, Askham J, et al. Consensus development methods, and their use in clinical guideline development. Health Technology Assessment. 1998;2.

Mustafa R, Santesso N, Brozek J, Akl E, Walter S, Norman G, et al. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. Journal of Clinical Epidemiology. 2013;66:736-42.

Naeem F, Bhatti F, Pickering R, Kingdon D. A controlled trial of the effectiveness of drug treatment & testing orders (DTTO) with standard care. Journal of Substance Use. 2007;12:253-65.

Naud H, Daigle MS. Predictive validity of the suicide probability scale in a male inmate population. Journal of Psychopathology and Behavioral Assessment. 2010;32:333-42.

Naud H, Daigle MS. How to improve testing when trying to predict inmate suicidal behavior. International Journal of Law and Psychiatry. 2013;36:390-98.

Needels K, James-Burdumy S, Burghardt J. Community case management for former jail inmates: its impacts on rearrest, drug use, and HIV risk. Journal of urban health. 2005.[cited;82:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/908/CN-00528908/frame.html.

NICE. Process and methods guides. The Guidelines Manual. London 2012.

NICE. Social Value Judgements. Principles for the Development of NICE Guidance. London: NICE; 2008.

NICE. Process and methods guides. Developing NICE guidelines: the manual. Manchester: National Institute of Health and Care Excellence; 2014.

Nitschke J, Osterheider M, Mokros A. A cumulative scale of severe sexual sadism. Sexual Abuse: A journal of research and treatment. 2009;21:262-78.

O'Reilly G, Carr A, Murphy P, Cotter A. A controlled evaluation of a prison-based sexual offender intervention program. Sexual Abuse: Journal of Research and Treatment. 2010;22:95-111.

Olver ME, Nicholaichuk TP, Gu D, Wong SCP. Sex Offender Treatment Outcome, Actuarial Risk, and the Aging Sex Offender in Canadian Corrections: A Long-Term Follow-up. Sexual Abuse: A journal of research and treatment. 2013a;25:396-422.

Ostermann M, Matejkowski J. Estimating the Impact of Mental Illness on Costs of Crimes A Matched Samples Comparison. Criminal Justice and Behavior. 2013.

Perry AE, Gilbody S. Detecting and predicting self-harm behaviour in prisoners: A prospective psychometric analysis of three instruments. Social Psychiatry and Psychiatric Epidemiology. 2009;44:853-61.

Perry AE, Olason DT. A new psychometric instrument assessing vulnerability to risk of suicide and self-harm behaviour in offenders: Suicide concerns for offenders in prison environment (Scope). International Journal of Offender Therapy and Comparative Criminology. 2009;53:385-400.

Peters JJ, Pedigo J, Steg J, McKenna J, Jr. Group psychotherapy of the sex offender. . Federal Probation. 1968;32:41-45.

Petrila J, Andel R, Constantine R, Robst J. Public expenditures related to the criminal justice system and to services for arrestees with a serious mental illness. Psychiatric Services. 2010.

Pickard H, Fazel S. Substance abuse as a risk factor for violence in mental illness: some implications for forensic psychiatric practice and clinical ethics. Current Opinion in Psychiatry. 2013;26:349-54.

Prendergast M, Frisman LK, Sacks JY, et al. A multi-site, randomised study of strengths-based case management with substance-abusing parolees. Journal of Experimental Criminology. 2011;7:225-53.

Prendergast M, Hall E, Grossman J, Veliz R, Gregorio L, Warda U, et al. Effectiveness of using incentives to improve parolee admission and attendance in community addiction treatment Criminal Justice and Behavior. 2015;42:1008-31.

1 Prison Reform Trust. Prison: The Facts. Bromley Briefings Summer 2013.

2 Prison Service Pay Review Body. Fourteenth Report on England and Wales 2015. 2016.

3 Proctor SL, Hoffmann NG, Allison S. The effectiveness of interactive journaling in reducing
4 recidivism among substance-dependent jail inmates. International Journal of Offender
5 Therapy and Comparative Criminology. 2012.[cited;56:Available from:
6 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/370/CN-00852370/frame.html.

7 Redondo Illescas S, Garrido Genoves V. Efficacy of a psychological treatment for sex
8 offenders. Psicothema. 2008;20:4-9.

9 Reveruzzi B, Pilling S. Street triage: Report on the evaluation of nine pilot schemes in
10 England 2016.

11 Reveruzzi B, Pilling S, Dervin S. Evaluating joint partnerships between police and mental
12 health across England. submitted for publication. 2016.

13 Revolving Doors Agency. Prisons: Britain's 'Social Dustbins'. Proposals to keep people with
14 mental health problems out of prison. 2007.

15 Rich J, McKenzie M, Larney S, Wong J, Tran L, Clarke J, et al. Methadone continuation
16 versus forced withdrawal on incarceration in a combined US prison and jail: a randomised,
17 open-label trial. Lancet. 2015;386:350-59.

18 Ridgely MS, Engberg J, Greenberg MD, Turner S, DeMartini C, Dembosky JW. RAND study
19 first to document costs and fiscal impact of a mental health court. Psychiatric Services.
20 2007;58.

21 Robertson A, Swanson J, Lin H, Easter M, Frisman L, Swartz M. Influence of Criminal
22 Justice Involvement and Psychiatric Diagnoses on Treatment Costs Among Adults With
23 Serious Mental Illness. Psychiatric Services. 2015.

24 Rossman S, Sridharan S, Gouvis C, Buck J, Morley E. Impact of the Opportunity to Succeed
25 (OPTS) aftercare program for substance-abusing felons: Comprehensive final report.
26 Washington DC: Urban Institute, 1999. 1999.[cited:Available from:
27 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/501/CN-00714501/frame.html.

28 Roy L, Crocker AG, Nicholls TL, Latimer EA, Ayllon AR. Criminal behaviour and victimization
29 among homeless individuals with severe mental illness: a systematic review. Psychiatric
30 Services. 2014;65:739-50.

31 Ruddijs F, Timmerman H. The Stichting ambulante preventie projecten method: A
32 comparative study of recidivism in first offenders in a Dutch outpatient setting. International
33 Journal of Offender Therapy and Comparative Criminology. 2000;44:725-39.

34 Sacks JY, McKendrick K, Hamilton Z. A randomized clinical trial of a therapeutic community
35 treatment for female inmates: outcomes at 6 and 12 months after prison release. Journal of
36 addictive diseases. 2012a.[cited;31:Available from:
37 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/214/CN-00845214/frame.html.

38 Sacks JY, Sacks S, McKendrick K, Banks S, Schoeneberger M, et al. Prison therapeutic
39 community treatment for female offenders: Profiles and preliminary findings for mental health
40 and other variables (crime, substance use and HIV risk). Journal of Offender Rehabilitation.
41 2008;46:233-26.

42 Sacks S, Chaple M, Sacks JY, McKendrick K, Cleland CM. Randomized trial of a reentry
43 modified therapeutic community for offenders with co-occurring disorders: crime outcomes.
44 Journal of substance abuse treatment. 2012b.[cited;42:Available from:
45 http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/123/CN-00845123/frame.html.

Sacks S, Melnick G, Coen C, Banks S, Friedmann PD, Grella C, et al. CJDATS Co-occurring Disorders Screening Instrument for Mental Disorders (CODSI-MD): A pilot study. The Prison Journal. 2007a;87:86-110.

Sacks S, Melnick G, Coen C, Banks S, Friedmann PD, Grella C, et al. CJDATS Co-Occurring Disorders Screening Instrument for Mental Disorders: A validation study. Criminal Justice and Behavior. 2007b;34:1198-215.

Sacks S, Sacks JY, McKendrick K, Banks S, Stommel J. Modified TC for MICA offenders: crime outcomes. Behavioral sciences & the law. 2004.[cited;22:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/460/CN-00490460/frame.html.

Schünemann H, Brożek J, Oxman A. GRADE handbook for grading quality of evidence and strength of recommendation. Version 3.2 [updated March 2009]. : The GRADE Working Group; 2009.

Schünemann HJ, Best D, Vist G, Oxman A, Group GW. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. . Can Med Assoc J. 2003;169:677-80.

Scott CK, Dennis ML. The first 90 days following release from jail: findings from the Recovery Management Checkups for Women Offenders (RMCWO) experiment. Drug and alcohol dependence. 2012.[cited;125:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/205/CN-00845205/frame.html.

Scott L. It Never, Ever Ends: The Psychological Impact of Wrongful Conviction. American University Criminal Law Brief. 2010;5:10-22.

Seto MC, Harris GT, Rice ME, Barbaree HE. The screening scale for pedophilic interests predicts recidivism among adult sex offenders with child victims. Archives of sexual behavior. 2004;33:455-66.

Shanahan M, Donato R. Counting the cost: Estimating the economic benefit of pedophile treatment programs. Child abuse & neglect. 2001;25:541-55.

Shanahan M, Lancsar E, Hass M, Lind B, Weatherburn D, Chen S. Cost-effectiveness analysis of the New South Wales adult drug cout program. Evaluation review. 2004;28:3-27.

Sheard L, Wright NM, El-Sayeh HG, Adams CE, Li R, Tompkins CN. The Leeds Evaluation of Efficacy of Detoxification Study (LEEDS) prisons project: a randomised controlled trial comparing dihydrocodeine and buprenorphine for opiate detoxification. Substance abuse treatment, prevention, and policy. 2009.[cited;4:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/753/CN-00683753/frame.html.

Shearer J, White B, Gilmour S, Wodak AD, Dolan KA. Hair analysis underestimates heroin use in prisoners. Drug and alcohol review. 2006.[cited;25:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/893/CN-00571893/frame.html.

Sinha R, Easton C, Renee-Aubin L, Carroll KM. Engaging young probation-referred marijuana-abusing individuals in treatment: a pilot trial. American journal on addictions / American Academy of Psychiatrists in Alcoholism and Addictions. 2003.[cited;12:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/520/CN-00471520/frame.html.

Sjostedt G, Grann M. Risk assessment: What is being predicted by actuarial prediction instruments? The International Journal of Forensic Mental Health. 2002b;1:179-83.

Sleed M, Baradon T, Fonagy P. New Beginnings for mothers and babies in prison: a cluster randomized controlled trial. Attachment & human development. 2013.[cited;15:Available

from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/561/CN-00960561/frame.html.

Solomon P, Draine J, Meyerson A. Jail recidivism and receipt of community mental health services. Hospital & community psychiatry. 1994.[cited;45:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/441/CN-00107441/frame.html.

Somers JM, Rezansoff SN, Moniruzzaman A, Palepu A, Patterson M. Housing First Reduces Re-offending among Formerly Homeless Adults with Mental Disorders: Results of a Randomized Controlled Trial. PloS One. 2013.[cited;8:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/533/CN-01000533/frame.html.

Song L, Lieb R. Washington State sex offenders: Overview of recidivism studies (No. 95-02-1101). Olympia, WA: Washington State Institute for Public Policy; 1995.

Spurgeon A, McCarthy-Tucker S, Waters TF. Developing a substance abuse relapse screening questionnaire for adults on intensive probation: A pilot study. Journal of Offender Rehabilitation. 2000;32:167-80.

Stalans LJ, Seng M, Yarnold P, Lavery T, Swartz J. Process and initial impact evaluation of the Cook County Adult Probation Department's Sex Offender Program: Final and summary report for the period of June, 1997 to June, 2000. . Chicago, IL: Illinois Department of Corrections; 2001.

Steadman HJ, Naples M. Assessing the effectiveness of jail diversion programs for persons with serious mental illness and co-occuring substance use disorders. Behavioral sciences & the law. 2005;23:163-70.

Steadman HJ, Robbins PC, Islam T, Osher FC. Revalidating the brief jail mental health screen to increase accuracy for women. Psychiatric Services. 2007;58:1598-601.

Steadman HJ, Scott JE, Osher F, Agnese TK, Robbins PC. Validation of the brief jail mental health screen. Psychiatric Services. 2005;56:816-22.

Stevens H, Agerbo E, Dean K, Nordentoft M, Nielsen PR, Mortensen PB. Offending prior to first psychiatric contact: a population-based register study. Psychological Medicine. 2012;42:2673-84.

Stuart GL, Shorey RC, Moore TM, Ramsey SE, Kahler CW, O'Farrell TJ, et al. Randomized clinical trial examining the incremental efficacy of a 90-minute motivational alcohol intervention as an adjunct to standard batterer intervention for men. Addiction (Abingdon, England). 2013.[cited;108:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/506/CN-00960506/frame.html.

Sullivan CJ, McKendrick K, Sacks S, Banks S. Modified therapeutic community treatment for offenders with MICA disorders: substance use outcomes. American journal of drug and alcohol abuse. 2007.[cited;33:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/870/CN-00620870/frame.html.

Sumter MT, Monk-Turner E, Turner C. The benefits of meditation practice in the correctional setting. Journal of correctional health care. 2009.[cited;15:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/778/CN-00702778/frame.html.

Teplin LA, McClelland GM, Abram KM, Weiner DA. Crime Victimization in Adults with Severe Mental Illness comparison with the National Crime Victimization Survey. Archives of general psychiatry. 2005 Aug;62:911-21.

Teplin LA, Swartz J. Screening for severe mental disorder in jails. The development of the referral decision scale. Law and Human Behavior. 1989;13:1-18.

Thomas E, Degenhardt L, Alati R, Kinner S. Predictive validity of the AUDIT for hazardous alcohol consumption in recently released prisoners. Drug and alcohol dependence. 2014;134:322-29.

Thornicroft G, Rose D, Kassam A. Discrimination in healthcare aganist people with mental illness. International review of psychiatry. 2007;19.

Turner BW, Bingham JE, Andrasik F. Short-term community-based treatment for sexual offenders: Enhancing effectiveness. Sexual Addiction & Compulsivity. 2000;7:211-23.

Valentine PV, Smith TE. Evaluating traumatic incident reduction therapy with female inmates: a randomised controlled clinical trial. Research on Social Work Practice. 2001;11:40-52.

Villagra Lanza P, Menendez AG. Acceptance and Commitment Therapy for drug abuse in incarcerated women. Psicothema. 2013.[cited;25:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/799/CN-00994799/frame.html.

Wakeling HC, Howard P, Barnett G. Comparing the validity of the RM2000 scales and OGRS3 for predicting recidivism by internet sexual offenders. Sexual Abuse: Journal of Research and Treatment. 2011a;23:146-68.

Wang EA, Hong CS, Shavit S, et al. Engaging individuals recently released from prison into primary care: a randomised trial. Research and Practice. 2012;102:e22-e29.

Warren E, Viney R, Shearer J, Shanahan M, Wodak A, Dolan KA. Value for money in drug treatment: economic evaluation of prison methadone. Drug and alcohol dependence. 2006;84:160-66.

Weaver T, Taylor F, et al. The Bentham Unit: a pilot remand and assessment service for male mentally disordered remand prisoners II: Report of an independent evaluation. British Journal of Psychiatry. 1997;170:462-66.

Webster JM, Staton-Tindall M, Dickson MF, Wilson JF, Leukefeld CG. Twelve-month employment intervention outcomes for drug-involved offenders. American journal of drug and alcohol abuse. 2014.[cited;40:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/090/CN-00988090/frame.html.

Werb D, Elliott R, Fischer B, Wood E, Montaner J, Kerr T. Drug treatment courts in Canada: an evidence-based review. HIV/AIDS policy & law review/Canadian HIV/AIDS Legal Network. 2007;12:12-17.

Wexler HK, Melnick G, Lowe L, Peters J. 3-year reincarceration outcomes for Amity in-prison therapeutic community and aftercare in California. The Prison Journal. 1999;79:321-36.

Whiting P, Rutjes, A., Westwood, M., Mallett, S., Deeks, J., Leeflang, M., et al. . QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine. 2011;155:529-36.

Wichmann C, Serin R, Motiuk LL. Predicting suicide attempts among male offenders in federal penitentiaries. In: Branch R, editor.: Correctional Service Canada; 2000.

Wilson GL. Psychotherapy with depressed incarcerated felons: a comparative evaluation of treatments. Psychological reports. 1990.[cited;67:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/306/CN-00073306/frame.html.

Wilson RJ, Cortoni F, McWhinnie AJ. Circles of support & accountability: A canadian national replication of outcome findings. Sexual Abuse: Journal of Research and Treatment. 2009;21:412-30.

Wilson RJ, Picheca JE, Prinzo M. Evaluating the effectiveness of professionally-facilitated volunteerism in the community-based management of high-risk sexual offenders: Part two--A comparison of recidivism rates. Howard Journal of Criminal Justice. 2007b;46:327-37.

Witkiewitz K, Warner K, Sully B, Barricks A, Stauffer C, Thompson BL, et al. Randomized trial comparing mindfulness-based relapse prevention with relapse prevention for women offenders at a residential addiction treatment center. Substance use & misuse. 2014.[cited;49:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/233/CN-01022233/frame.html.

Wright NM, Sheard L, Adams CE, Rushforth BJ, Harrison W, Bound N, et al. Comparison of methadone and buprenorphine for opiate detoxification (LEEDS trial): a randomised controlled trial. British journal of general practice. 2011.[cited;61:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/479/CN-00814479/frame.html.

Zarkin GA, Cowell AJ, Hicks KA, Mills MJ, Belenko S, Dundlap LJ, et al. Lifetime benefits and costs of diverting substance-abusing offenders from state prison. Crime & Delinquency. 2015;48:57-62.

Zlotnick C, Johnson J, Najavits LM. Randomized controlled pilot study of cognitive-behavioral therapy in a sample of incarcerated women with substance use disorder and PTSD. Behavior therapy. 2009.[cited;40:Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/748/CN-00730748/frame.html.