Cochrane Library

Cochrane Database of Systematic Reviews

Specialised early intervention teams for first episode psychosis (Protocol)

Puntis S, Minichino A, De Crescenzo F, Cipriani A, Lennox B

Puntis S, Minichino A, De Crescenzo F, Cipriani A, Lennox B.
Specialised early intervention teams for first episode psychosis.
Cochrane Database of Systematic Reviews 2019, Issue 3. Art. No.: CD013288.
DOI: 10.1002/14651858.CD013288.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Protocol]

Specialised early intervention teams for first episode psychosis

Stephen Puntis[1], Amedeo Minichino[1], Franco De Crescenzo[1], Andrea Cipriani[1], Belinda Lennox[1]

[1]Department of Psychiatry, University of Oxford, Oxford, UK

Contact address: Stephen Puntis, Department of Psychiatry, University of Oxford, Warneford Hospital, Warneford Lane, Oxford, OX3 7JX, UK. stephen.puntis@psych.ox.ac.uk.

Editorial group: Cochrane Schizophrenia Group.
Publication status and date: New, published in Issue 3, 2019.

Citation: Puntis S, Minichino A, De Crescenzo F, Cipriani A, Lennox B. Specialised early intervention teams for first episode psychosis. *Cochrane Database of Systematic Reviews* 2019, Issue 3. Art. No.: CD013288. DOI: 10.1002/14651858.CD013288.

ABSTRACT

This is a protocol for a Cochrane Review (Intervention). The objectives are as follows:

To compare early intervention in psychosis (EIP) specialised teams to usual community mental health care for the treatment of people with first episode psychosis (FEP).

BACKGROUND

Description of the condition

The lifetime prevalence of psychotic illness is estimated to be 4 per 1000 of the population, with first episode psychosis (FEP) incidence estimated at 34 new cases per 100,000 person-years (Kirkbride 2012; Kirkbride 2017). FEP can occur at any age, but most people develop it in late adolescence and early adulthood, with a mean age of onset in the early 20's (Kirkbride 2017). Features of psychosis include hallucinations, delusions and disordered thinking (referred to as positive symptoms) and social withdrawal, flat or blunted affect, and poverty of speech (referred to as negative symptoms) (APA 2013). Psychotic illness encompasses a range of diagnoses, including schizophrenia and schizoaffective disorder, bipolar affective disorder and psychotic depression. The impact on the individual is also significant; a psychotic illness has wide-ranging implications on quality of life and disability, including effects on physical health, social functioning, social inclusion, and education and employment (Mason 1995; Meltzer 2002).

Many of those with FEP reach remission of psychotic symptoms and functional recovery following a first episode of psychosis, but the majority will relapse, and as the number of relapses increases, the likelihood of remission decreases (Wiersma 1998; Morgan 2014; Lally 2017). Schizophrenia and related psychotic illnesses are major contributors to the global burden of disease, with the associated economic costs estimated to range between USD 94 million and USD 102 billion by country (Murray 1996; Chong 2016).

Description of the intervention

An early intervention in psychosis (EIP) service is a specialised multidisciplinary community mental health team that treats people experiencing their first episode of a psychotic illness (Fusar-Poli 2017). The objectives of EIP services are two-fold: first, they aim to intervene at an early stage of the illness, reducing the duration of untreated psychosis; second, they aim to provide a comprehensive package of treatment including medication, psychological therapies, and patient and family education, all backed by assertive case

management (NICE 2014). The aim of EIP is to reduce impairment and facilitate recovery, and in turn, improve prognosis.
EIP services are time-limited to two or three years of treatment (depending on region and health service provision), with the rationale that early intensive treatment will preclude the need for such intensive treatment on an ongoing basis (i.e. a secondary prevention approach). The service model is of standalone, multidisciplinary community teams that provide an assertive outreach model of care. Care co-ordinators have restricted caseload size to enable them to work intensively with patients and engage them in treatment (RCPsych 2016).

How the intervention might work

The long-term prognosis for psychotic illnesses has traditionally been considered very poor and therefore treatment had been focused on managing those with the most severe, chronic illness and disability. In the last two decades, however, this viewpoint has changed. Follow-up studies of those at first episode of psychosis have shown an estimated rate of remission at 58% and rate of recovery at 38% (Lally 2017). The focus in early intervention has therefore been to enhance and improve this rate of recovery.
One of the strongest arguments for the development of early phase treatments is that there is evidence of a 'critical period' in FEP. This period, during the first few weeks and months of a psychotic illness, is potentially a period of rapid biological, psychological, and social changes, after which is followed by an eventual plateau of illness severity and functioning (Birchwood 1998). This trajectory of fluctuation of illness in the early years, followed by gradual deterioration has been found to be strongly predictive of later outcomes (Wiersma 1998; Harrison 2001). Standard community mental health teams had particular difficulty engaging this population, making it hard to deliver treatment (Birchwood 2014). EIP was developed primarily to improve engagement through assertive outreach, reducing the time to treatment (thereby reducing the duration of untreated psychosis) and potentially minimising the long-term burden of the illness (Fusar-Poli 2017).

Why it is important to do this review

EIP services are now considered the gold standard of care for people with FEP in America, Europe, and Australasia. In the UK it is the recommended treatment by the National Institute for Health and Care Excellence (NICE 2014), and timely access to EIP was the first National Health Service (NHS) waiting time standard for mental health care (NHS England 2015). Despite its popularity, a previous Cochrane Review of early interventions for psychosis only found one eligible randomised control trial (RCT) of specialist team interventions for FEP (Marshall 2011). A number of new RCTs comparing EIP to usual care have since been published (for example, Ruggeri 2015 and Kane 2016) and a recent meta-analysis of EIP in the treatment of 'early phase psychosis', which included trials that recruited participants with multiple acute psychotic episodes or patients who had already had lengthy community treatment (e.g. up to 5 years), found EIP superior to standard community mental health care in reducing treatment discontinuation, admission to hospital, and psychotic symptoms (Correll 2018). It is important to do this review to ensure that new evidence for EIP services is evaluated.

OBJECTIVES

To compare early intervention in psychosis (EIP) specialised teams to usual community mental health care for the treatment of people with first episode psychosis (FEP).

METHODS

Criteria for considering studies for this review

Types of studies

We will consider randomised controlled trials (RCTs) meeting our inclusion criteria and reporting useable data. We will include RCTs regardless of blinding, but will exclude quasi-randomised studies, such as those that allocate intervention by alternate days of the week. Given the nature of the intervention it would be difficult to blind participants and clinicians from whether they are receiving the intervention or control condition and so we will include both single- and double-blinded studies. Where people are given additional treatments as well as specialised early intervention treatment for psychosis, we will only include data if the adjunct treatment is evenly distributed between groups and it is only the EIP specialised teams that are randomised. We will not exclude studies offering alternative models of care, such as step-down care, following discharge from their EIP specialised team.

Types of participants

EIP services are designed to treat people in the early stages of psychosis. Exact eligibility criteria for services often differ both within and between regions and countries, but generally have a 'time since onset' criterion and a 'number of onsets' criterion. We will include participants with a first or second episode of psychosis within three years of the onset of their first psychosis. We will include participants who currently exhibit symptoms that match the criteria for primary psychotic diagnoses according to standardised criteria, such as the Diagnostic and Statistical Manual of Mental Disorders: DSM-III (APA 1980), DSM-III-R (APA 1987), DSM-

IV (APA 1994), DSM-IV-TR (APA 2000), DSM-5 (APA 2013), ICD-10 (WHO 2004), ICD-11 (WHO 2018) or Melbourne Criteria (Yung 2008). We will exclude trials where participants have organic psychoses or head injury, and studies that recruited participants with prodromal symptoms (also known as 'at-risk mental states') who have not yet transitioned to a psychotic episode. We will also exclude trials that include participants whose onset of illness is longer than three years, unless we can extract data on only eligible participants from the paper or such data are provided by trial authors.

Types of interventions

1.1. Early intervention in psychosis (EIP) specialised team care

These are multidisciplinary, standalone community-based mental health teams which take referrals for patients who have first episode psychosis (FEP). EIP specialised teams provide a specified package of comprehensive care to individuals with FEP usually structured around a combination of assertive community engagement, medication and psychological and social interventions to individuals and families/carers. These interventions are provided by, and co-ordinated by the EIP specialised teams. EIP specialised teams are an alternative to, rather than an addition to, standard psychiatric care.

1.2. Standard care

Standard care (or treatment as usual) for people with first episode psychosis differs by country, but usually consists of a community-based or outpatient mental health team which does not provide specialist, phase-specific (i.e. centred on the early phase of a psychotic illness) treatment.

Types of outcome measures

Timing of outcomes assessment

We will record post-treatment outcomes and any available outcomes during treatment and at reported follow-up. The duration of EIP treatment can differ substantially between trials, so where appropriate, and if the data are available, we will categorise treatment outcomes into short-term (up to 12 months), medium-term (13 to 24 months), and long-term (longer than 24 months). We will also divide follow-up post-treatment outcomes into short-term (up to 12 months post-treatment), medium-term (13 to 60 months post-treatment), and long-term (longer than 60 months post-treatment).

Primary outcomes

1. Global state
1.1 Recovery, as defined by the study
2. Service use
2.1 Disengagement from services, as defined by individual study

Secondary outcomes

1. Service use
1.1 Admission to hospital
1.2 Readmission
1.3 Number of days in hospital
2. Mental state
1.1 General
2.1.1 Clinically important change in general mental state
2.1.2 Any change in general mental state
2.1.3 Average endpoint/change score on a general mental state scale
2.2 Specific
2.2.1 Clinically important change in positive symptoms (delusions, hallucinations, disordered thinking), as defined by individual studies
2.2.2 Any change in positive symptoms (delusions, hallucinations, disordered thinking), as defined by individual studies
2.2.3 Clinically important change in negative symptoms (avolition, poor self-care, blunted affect), as defined by individual studies
2.2.4 Any change in negative symptoms (avolition, poor self-care, blunted affect), as defined by individual studies
2.2.5 Clinically important change in depression, as defined by individual studies
2.2.6 Any change in depression, as defined by individual studies
2.2.7 Average endpoint/change score on specific symptoms mental state scale/subscale
3. Behaviour

3.1 Specific

3.2.1 Occurrence of violent incidents (to self, others or property)
4. Adverse effects/events
4.1 General
4.1.1 At least one adverse effect/event
4.1.2 Average endpoint/change score on adverse effect scale

4.2 Specific

4.2.1 Incidence of any specific adverse effects, as defined by individual studies
4.3 Death
4.3.1 Suicide or natural cause

5. Leaving the study early

5.1 For any reason
5.2 Due to adverse effect

6. Quality of life (recipient or informal carers or professional carers)

6.1 Overall

6.1.1 Clinically important change in overall quality of life
6.1.2 Average endpoint/change score on quality of life scale

7. Functioning

7.1 General
7.1.1 Clinically important change in general functioning
7.1.2 Average endpoint/change score on general functioning scale

7.2 Specific (including social, cognitive, life skills)

7.2.1 Clinically important change in specific functioning
7.2.2 Average endpoint/change score on specific functioning scale
7.2.3 Any change in educational status
7.2.4 Any change in employment status

8. Satisfaction with care (including subjective well-being and family burden)

8.1 Recipient
8.1.1 Recipient satisfied with care
8.1.2 Average endpoint/change score on satisfaction scale
8.2 Carers
8.1.1 Carer satisfied with care
8.1.2 Average endpoint/change score on satisfaction scale

'Summary of findings' table(s)

We will use the GRADE approach to interpret findings (Schünemann 2011); and will use GRADEpro GDT to export data from our review to create a 'Summary of findings' table(s). A 'Summary of findings' table provides outcome-specific information concerning the overall certainty of evidence from each included study in the comparison, the magnitude of effect of the interventions examined, and the sum of available data on all outcomes we rate as important to patient care and decision making. We aim to select the following main outcomes for inclusion in the 'Summary of findings' table(s).

1. Global state: recovery, as defined by each study (at end of treatment).
2. Service use: disengagement from services, as defined by each study (at end of treatment).
3. Service use: admission at end of treatment.
4. Service use: number of days in hospital at end of treatment.
5. Mental state: clinically important change in general mental state at end of treatment.
6. Functioning: specific - clinically important change in social functioning at end of treatment.
7. Adverse effects/events: death - any cause (at end of treatment).

If data are not available for these prespecified outcomes but are available for ones that are similar, we will present the closest outcome to the prespecified one in the table but take this into account when grading the finding.

Search methods for identification of studies

Electronic searches

Cochrane Schizophrenia Group's study-based register of trials

The Information Specialist will search the register using the following search strategy.
(*Early Intervention* AND *Special*) in Intervention Field of STUDY
In such study-based register, searching the major concept retrieves all the synonyms and relevant studies because all the studies have already been organised based on their interventions and linked to the relevant topics (Shokraneh 2017; Shokraneh 2018).
This register is compiled by systematic searches of major resources (AMED, BIOSIS, CENTRAL, CINAHL, ClinicalTrials.Gov, Embase, MEDLINE, PsycINFO, PubMed, WHO ICTRP) and their monthly updates, ProQuest Dissertations and Theses A&I and its quarterly update, Chinese databases (CBM, CNKI, and Wanfang) and their annual updates, handsearches, grey literature, and conference proceedings (see Cochrane Schrizophrenia website). There are no language, date, document type, or publication status limitations for inclusion of records into the register.

Searching other resources

1. Reference searching

We will inspect references of all included studies for further relevant studies.

2. Personal contact

We will contact the first author of each included study and known experts in the field for information regarding unpublished trials. We will note the outcome of this contact in the 'Included studies' or 'Studies awaiting classification' tables.

Data collection and analysis

Selection of studies

Review authors SP and AM will independently inspect citations from the searches and identify relevant abstracts; BL will independently re-inspect a random 20% sample of these abstracts to ensure reliability of selection. Where disputes arise, we will acquire the full report for more detailed scrutiny. SP and AM will then obtain and inspect full reports of the abstracts or reports meeting the review criteria. BL will reinspect a random 20% of these full reports in order to ensure reliability of selection. In case of disagreement, we will involve another member of the review team (AC) to reach a final decision. Where it is not possible to resolve disagreement by discussion, we will attempt to contact the authors of the study concerned for clarification.

Data extraction and management

1. Extraction

Review authors SP, AM, and FDC will independently extract data from all included studies. In addition, to ensure reliability, BL will independently extract data from a random sample of these studies, comprising 10% of the total. We will attempt to extract data presented only in graphs and figures whenever possible, but will include only if two review authors independently obtain the same result. If studies are multicentred, then where possible we will extract data relevant to each. We will discuss any disagreement and document our decisions. If necessary, we will attempt to contact authors through an open-ended request in order to obtain missing information or for clarification. AC will help clarify issues regarding any remaining problems and we will document these final decisions.

2. Management

2.1 Forms

We will extract data onto standard, predesigned, simple forms.

2.2 Scale-derived data

We will include continuous data from rating scales only if:

a) the psychometric properties of the measuring instrument have been described in a peer-reviewed journal (Marshall 2000);
b) the measuring instrument has not been written or modified by one of the trialists for that particular trial; and
c) the instrument should be a global assessment of an area of functioning and not subscores which are not, in themselves, validated or shown to be reliable. However there are exceptions; we will include subscores from mental state scales measuring positive and negative symptoms of schizophrenia.

Ideally the measuring instrument should either be i) a self-report or ii) completed by an independent rater or relative (not the therapist). We realise that this is not often reported clearly; in 'Description of studies' we will note if this is the case or not.

2.3 Endpoint versus change data

There are advantages of both endpoint and change data: change data can remove a component of between-person variability from the analysis; however, calculation of change needs two assessments (baseline and endpoint) that can be difficult to obtain in unstable and difficult-to-measure conditions such as schizophrenia. We have decided primarily to use endpoint data, and only use change data if the former are not available (Deeks 2011).

2.4 Skewed data

Continuous data on clinical and social outcomes are often not normally distributed. To avoid the pitfall of applying parametric tests to non-parametric data, we will apply the following standards to relevant continuous data before inclusion.

Endpoint data from studies with fewer than 200 participants

a) When a scale starts from the finite number zero, we will subtract the lowest possible value from the mean, and divide this by the standard deviation (SD). If this value is lower than one, it strongly suggests that the data are skewed and we will exclude these data. If this ratio is higher than one but less than two, there is suggestion that the data are skewed: we will enter these data and test whether their inclusion or exclusion would change the results substantially. If such data change the results we will enter as 'other data'. Finally, if the ratio is larger than two we will include these data, because it is less likely that they are skewed (Altman 1996).
b) If a scale starts from a positive value (such as the Positive and Negative Syndrome Scale (PANSS), which can have values from 30 to 210 (Kay 1986)), we will modify the calculation described above to take the scale starting point into account. In these cases

skewed data are present if 2 SD > (S − S min), where S is the mean score and 'S min' is the minimum score.
Please note: we will enter all relevant data from studies of more than 200 participants in the analysis irrespective of the above rules, because skewed data pose less of a problem in large studies. We will also enter all relevant change data, as when continuous data are presented on a scale that includes a possibility of negative values (such as change data), it is difficult to tell whether or not data are skewed.

2.5 Common measurement

To facilitate comparison between trials we aim, where relevant, to convert variables that can be reported in different metrics, such as days in hospital (mean days per year, per week or per month) to a common metric (e.g. mean days per month).

2.6 Conversion of continuous to binary

Where possible, we will make efforts to convert outcome measures to dichotomous data. This can be done by identifying cut-off points on rating scales and dividing participants accordingly into 'clinically improved' or 'not clinically improved'. It is generally assumed that if there is a 50% reduction in a scale-derived score such as the Brief Psychiatric Rating Scale (BPRS) (Overall 1962), or the PANSS (Kay 1986), this could be considered as a clinically significant response (Leucht 2005a; Leucht 2005b). If data based on these thresholds are not available, we will use the primary cut-off presented by the original authors.

2.7 Direction of graphs

Where possible, we will enter data in such a way that the area to the left of the line of no effect indicates a favourable outcome for EIP. Where keeping to this makes it impossible to avoid outcome titles with clumsy double-negatives (e.g. 'not unimproved') we will report data where the left of the line indicates an unfavourable outcome and note this in the relevant graphs.

Assessment of risk of bias in included studies

Review authors SP and AM will work independently to assess risk of bias by using criteria described in the *Cochrane Handbook for Systematic Reviews of Interventions* to assess trial quality (Higgins 2011a). This set of criteria is based on evidence of associations between potential overestimation of effect and the level of risk of bias of the article that may be due to aspects of sequence generation, allocation concealment, blinding, incomplete outcome data, and selective reporting, or the way in which these 'domains' are reported.
If the raters disagree, we will make the final rating by consensus. Where inadequate details of randomisation and other characteristics of trials are provided, we will attempt to contact authors of the studies in order to obtain further information. We will report non-concurrence in quality assessment, but if disputes arise regarding the category to which a trial is to be allocated, we will resolve this by discussion.
We will note the level of risk of bias in both the text of the review, 'Risk of bias' figures, and the 'Summary of finding' table(s).

Measures of treatment effect

1. Binary data

For binary outcomes we will calculate a standard estimation of the risk ratio (RR) and its 95% confidence interval (CI), as it has been shown that RR is more intuitive than odds ratios (ORs) (Boissel 1999); and that ORs tend to be interpreted as RR by clinicians (Deeks 2000). Although the number needed to treat for an additional beneficial outcome (NNTB) and the number needed to treat for an additional harmful outcome (NNTH), with their CIs, are intuitively attractive to clinicians, they are problematic to calculate and interpret in meta-analyses (Hutton 2009). For binary data presented in the 'Summary of findings' table(s) we will, where possible, calculate illustrative comparative risks.

2. Continuous data

For continuous outcomes we will estimate the MD between groups.

Unit of analysis issues

1. Cluster trials

Studies increasingly employ 'cluster randomisation' (such as randomisation by clinician or practice), but analysis and pooling of clustered data poses problems. Authors often fail to account for intraclass correlation in clustered studies, leading to a unit of analysis error whereby P values are spuriously low, CIs unduly narrow and statistical significance overestimated (Divine 1992). This causes type I errors (Bland 1997; Gulliford 1999).
Where clustering has been incorporated into the analysis of primary studies, we will present these data as if from a non-cluster randomised study, but adjust for the clustering effect.
Where clustering is not accounted for in primary studies, we will present data in a table, with a (*) symbol to indicate the presence of a probable unit of analysis error. We will seek to contact first authors of studies to obtain intraclass correlation coefficients for their clustered data and to adjust for this by using accepted methods (Gulliford 1999).
We have sought statistical advice and have been advised that the binary data from cluster trials presented in a report should be divided by a 'design effect'. This is calculated using the mean number of participants per cluster (m) and the intraclass correlation

coefficient (ICC): thus design effect = 1 + (m − 1) * ICC (Donner 2002). If the ICC is not reported we will assume it to be 0.1 (Ukoumunne 1999).

If cluster studies have been appropriately analysed and taken intraclass correlation coefficients and relevant data documented in the report into account, synthesis with other studies will be possible using the generic inverse variance technique.

2. Cross-over trials

A major concern of cross-over trials is the carry-over effect. This occurs if an effect (e.g. pharmacological, physiological or psychological) of the treatment in the first phase is carried over to the second phase. As a consequence, participants can differ significantly from their initial state at entry to the second phase, despite a washout phase. For the same reason cross-over trials are not appropriate if the condition of interest is unstable (Elbourne 2002). As both carry-over and unstable conditions are very likely in severe mental illness, we will only use data from the first phase of cross-over studies.

3. Studies with multiple treatment groups

Where a study involves more than two treatment arms, if relevant, we will present the additional treatment arms in comparisons. If data are binary, we will simply add these and combine within the two-by-two table. If data are continuous, we will combine data following the formula in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b). Where additional treatment arms are not relevant, we will not reproduce these data.

Dealing with missing data

1. Overall loss of credibility

At some degree of loss of follow-up, data must lose credibility (Xia 2009). We choose that, for any particular outcome, should more than 50% of data be unaccounted for we will not reproduce these data or use them within analyses. If, however, more than 50% of those in one arm of a study are lost, but the total loss is less than 50%, we will address this within the 'Summary of findings' table(s) by downgrading quality. Finally, we will also downgrade quality within the 'Summary of findings' table(s) should the loss be 25% to 50% in total.

2. Binary

In the case where attrition for a binary outcome is between 0% and 50% and where these data are not clearly described, we will present data on a 'once-randomised-always-analyse' basis (an intention-to-treat analysis (ITT)). Those leaving the study early are all assumed to have the same rates of negative outcome as those who completed. We will use the rate of those who stay in the study - in that particular arm of the trial - and apply this also to those who did not. We will undertake a sensitivity analysis testing how prone the primary outcomes are to change when data only from people who complete the study to that point are compared to the ITT analysis using the above assumptions.

3. Continuous

3.1 Attrition

We will use data where attrition for a continuous outcome is between 0% and 50%, and data only from people who complete the study to that point are reported.

3.2 Standard deviations

If SDs are not reported, we will try to obtain the missing values from the authors. If these are not available, where there are missing measures of variance for continuous data, but an exact standard error (SE) and CIs available for group means, and either P value or t value available for differences in mean, we can calculate SDs according to the rules described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b). When only the SE is reported, SDs are calculated by the formula SD = SE * $\sqrt{}$(n). The *Cochrane Handbook for Systematic Reviews of Interventions* presents detailed formulae for estimating SDs from P, t or F values, CIs, ranges or other statistics (Higgins 2011b). If these formulae do not apply, we will calculate the SDs according to a validated imputation method which is based on the SDs of the other included studies (Furukawa 2006). Although some of these imputation strategies can introduce error, the alternative would be to exclude a given study's outcome and thus to lose information. Nevertheless, we will examine the validity of the imputations in a sensitivity analysis that excludes imputed values.

3.3 Assumptions about participants who left the trials early or were lost to follow-up

Various methods are available to account for participants who left the trials early or were lost to follow-up. Some trials just present the results of study completers; others use the method of last observation carried forward (LOCF); while more recently, methods such as multiple imputation or mixed-effects models for repeated measurements (MMRM) have become more of a standard. While the latter methods seem to be somewhat better than LOCF (Leon 2006), we feel that the high percentage of participants leaving the studies early and differences between groups in their reasons for doing so is often the core problem in randomised schizophrenia trials. We will therefore not exclude studies based on the statistical

approach used. However, by preference we will use the more sophisticated approaches, i.e. we will prefer to use MMRM or multiple imputation to LOCF, and we will only present completer analyses if some kind of ITT data are not available at all. Moreover, we will address this issue in the item 'Incomplete outcome data' of the 'Risk of bias' tool.

Assessment of heterogeneity

1. Clinical heterogeneity

We will consider all included studies initially, without seeing comparison data, to judge clinical heterogeneity. We will simply inspect all studies for participants who are clearly outliers or situations that we had not predicted would arise and, where found, discuss such situations or participant groups.

2. Methodological heterogeneity

We will consider all included studies initially, without seeing comparison data, to judge methodological heterogeneity. We will simply inspect all studies for clearly outlying methods which we had not predicted would arise and discuss any such methodological outliers.

3. Statistical heterogeneity

3.1 Visual inspection

We will inspect graphs visually to investigate the possibility of statistical heterogeneity.

3.2 Employing the I^2 statistic

We will investigate heterogeneity between studies by considering the I^2 statistic alongside the Chi^2 P value. The I^2 statistic provides an estimate of the percentage of inconsistency thought to be due to chance (Higgins 2003). The importance of the observed value of I^2 depends on the magnitude and direction of effects as well as the strength of evidence for heterogeneity (e.g. P value from Chi^2 test, or a confidence interval for I^2). We will interpret an I^2 estimate greater than or equal to 50% and accompanied by a statistically significant Chi^2 statistic as evidence of substantial heterogeneity (Deeks 2011). When substantial levels of heterogeneity are found in the primary outcome, we will explore reasons for heterogeneity (Subgroup analysis and investigation of heterogeneity).

Assessment of reporting biases

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results (Egger 1997). These are described in section 10.1 of the *Cochrane Handbook for Systemic reviews of Interventions* (Sterne 2011).

1. Protocol versus full study

We will try to locate protocols of included RCTs. If the protocol is available, we will compare outcomes in the protocol and in the published report. If the protocol is not available, we will compare outcomes listed in the methods section of the trial report with actually reported results.

2. Funnel plot

We are aware that funnel plots may be useful in investigating reporting biases but are of limited power to detect small-study effects. We will not use funnel plots for outcomes where there are 10 or fewer studies, or where all studies are of similar size. In other cases, where funnel plots are possible, we will seek statistical advice in their interpretation.

Data synthesis

We understand that there is no closed argument for preference for use of fixed-effect or random-effects models. The random-effects method incorporates an assumption that the different studies are estimating different, yet related, intervention effects. This often seems to be true to us and the random-effects model takes into account differences between studies, even if there is no statistically significant heterogeneity. There is, however, a disadvantage to the random-effects model: it puts added weight onto small studies, which often are the most biased ones. Depending on the direction of effect, these studies can either inflate or deflate the effect size. We choose to use a random-effects model for analyses.

Subgroup analysis and investigation of heterogeneity

1. Subgroup analyses

1.1 Standard EIP treatment duration

Where the duration of EIP treatment differs by more than six months from the standard two-year duration of EIP care, we will include only standard duration EIP trials in a subgroup analysis.

2. Investigation of heterogeneity

We will report if inconsistency is high. Firstly, we will investigate whether data have been entered correctly. Secondly, if data are correct, we will inspect the graph visually and remove outlying studies

successively to see if homogeneity is restored. For this review we have decided that should this occur with data contributing to the summary finding of no more than 10% of the total weighting, we will present data. If not, we will not pool these data and will discuss any issues. We know of no supporting research for this 10% cut-off, but are investigating use of prediction intervals as an alternative to this unsatisfactory state.

When unanticipated clinical or methodological heterogeneity is obvious we will simply state hypotheses regarding these for future reviews or versions of this review. We do not anticipate undertaking analyses relating to these.

Sensitivity analysis

We will carry out sensitivity analyses for primary outcomes only. If there are substantial differences in the direction or precision of effect estimates in any of the sensitivity analyses listed below, we will not add data from the lower-quality studies to the results of the higher-quality trials, but will present these data within a subcategory. If their inclusion does not result in a substantive difference, they will remain in the analyses.

1. Implication of randomisation

If trials are described in some way as to imply randomisation, we will compare data from the implied trials with trials that are randomised.

2. Assumptions for lost binary data

Where assumptions have to be made regarding people lost to follow-up (see Dealing with missing data) we will compare the findings when we use our assumption compared with completer data only. If there is a substantial difference, we will report results and discuss them but continue to employ our assumption.

3. Assumptions for lost continuous data

Where assumptions have to be made regarding missing SDs (see Dealing with missing data), we will compare the findings when we use our assumption compared with data that are not imputed. If there is a substantial difference, we will report results and discuss them but continue to employ our assumption.

4. Risk of bias

We will analyse the effects of excluding trials that are at high risk of bias across one or more of the domains (see Assessment of risk of bias in included studies).

5. Imputed values

We will also undertake a sensitivity analysis to assess the effects of including data from trials where we use imputed values for ICC in calculating the design effect in cluster randomised trials.

6. Fixed- and random-effects

We will synthesise data using a random-effects model; however, we will also synthesise data for the primary outcome using a fixed-effect model to evaluate whether this alters the significance of the results.

ACKNOWLEDGEMENTS

Cochrane Schizophrenia's editorial base at The University of Nottingham, Nottingham, UK, produces and maintains standard text for use in the Methods section of their reviews. We have used this text as the basis of what appears here and adapted it as required.

REFERENCES

Additional references

Altman 1996
Altman DG, Bland JM. Detecting skewness from summary information. *BMJ* 1996;**313**(7066):1200.

APA 1980
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: Third Edition*. Washington, DC: American Psychiatric Association, 1980.

APA 1987
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: Thrid Edition Revised*. Washington, DC: American Psychiatric Association, 1987.

APA 1994
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition*. Washington, DC: American Psychiatric Association, 1994.

APA 2000
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition Text Revision*. Washington, DC: American Psychiatric Association, 2000.

APA 2013
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: Fifth Edition*. Arlington, VA: American Psychiatric Association, 2013.

Birchwood 1998
Birchwood M, Todd P, Jackson C. Early intervention in

psychosis: the critical-period hypothesis. *British Journal of Psychiatry* 1998;**172**(S33):53–9.

Birchwood 2014
Birchwood M, Lester H, McCarthy L, Jones P, Fowler D, Amos T, et al. The UK national evaluation of the development and impact of Early Intervention Services (the National EDEN studies): study rationale, design and baseline characteristics. *Early Intervention in Psychiatry* 2014;**8**(1):59–67.

Bland 1997
Bland JM. Statistics notes. Trials randomised in clusters. *BMJ* 1997;**315**(7108):600.

Boissel 1999
Boissel JP, Cucherat M, Li W, Chatellier G, Gueyffier F, Buyse M, et al. The problem of therapeutic efficacy indices. 3. Comparison of the indices and their use [Aperçu sur la problématique des indices d'efficacité thérapeutique, 3: comparaison des indices et utilisation. Groupe d'Etude des Indices D'efficacite]. *Therapie* 1999;**54**(4):405–11. [PUBMED: 10667106]

Chong 2016
Chong H, Teoh S, Wu D, Kotirum S, Chiou C, Chaiyakunapruk N. Global economic burden of schizophrenia: a systematic review. Neuropsychiatric Disease and Treatment 2016; Vol. 12:357–73.

Correll 2018
Correll C, Galling B, Pawar A, Krivko A, Bonetto C, Ruggeri M, et al. Comparison of early intervention services vs treatment as usual for early-phase psychosis: a systematic review, meta-analysis, and meta-regression. *JAMA Psychiatry* 2018;**75**(6):555–65.

Deeks 2000
Deeks J. Issues in the selection for meta-analyses of binary data. Proceedings of the 8th International Cochrane Colloquium; 2000 Oct 25-28; Cape Town. Cape Town: The Cochrane Collaboration, 2000.

Deeks 2011
Deeks JJ, Higgins JP, Altman DG, editor(s). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Divine 1992
Divine GW, Brown JT, Frazier LM. The unit of analysis error in studies about physicians' patient care behavior. *Journal of General Internal Medicine* 1992;**7**(6):623–9.

Donner 2002
Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine* 2002;**21**(19): 2971–80.

Egger 1997
Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**(7109):629–34.

Elbourne 2002
Elbourne D, Altman DG, Higgins JP, Curtina F, Worthingtond HV, Vaile A. Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* 2002;**31**(1):140–9.

Furukawa 2006
Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology* 2006;**59**(1):7–10.

Fusar-Poli 2017
Fusar-Poli P, McGorry PD, Kane JM. Improving outcomes of first-episode psychosis: an overview. *World Psychiatry* 2017;**16**(3):251–65.

Gulliford 1999
Gulliford MC. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. *American Journal of Epidemiology* 1999;**149** (9):876–83.

Harrison 2001
Harrison G, Hopper K, Craig T, Laska E, Siegel C, Wanderling J, et al. Recovery from psychotic illness: a 15- and 25-year international follow-up study. *British Journal of Psychiatry* 2001;**178**(6):506–17.

Higgins 2003
Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327** (7414):557–60.

Higgins 2011a
Higgins JP, Altman DG, Sterne JA, editor(s). Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2011b
Higgins JP, Deeks JJ, editor(s). Chapter 7: Selecting studies and collecting data. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Hutton 2009
Hutton JL. Number needed to treat and number needed to harm are not the best way to report and assess the results of randomised clinical trials. *British Journal of Haematology* 2009;**146**(1):27–30. [PUBMED: 19438480]

Kane 2016
Kane J, Robinson D, Schooler N, Mueser K, Penn D, Rosenheck R, et al. Comprehensive versus usual community care for First-Episode Psychosis: 2-Year outcomes from the NIMH RAISE Early Treatment Program. *American Journal of Psychiatry* 2016;**173**(4):362–72.

Kay 1986
Kay SR, Opler LA, Fiszbein A. *Positive and Negative Syndrome Scale (PANSS) Manual*. North Tonawanda (NY): Multi-Health Systems, 1986.

Kirkbride 2012
Kirkbride JB, Errazuriz A, Croudace TJ, Morgan C, Jackson D, Boydell J, et al. Incidence of schizophrenia and other psychoses in England, 1950-2009: a systematic review and meta-analyses. *PLoS One* 2012;**7**(3):1–21. DOI: https://doi.org/10.1371/journal.pone.0031660

Kirkbride 2017
Kirkbride JB, Hameed Y, Ankireddypalli G, Ioannidis K, Crane CM, Nasir M, et al. The epidemiology of First-Episode Psychosis in Early Intervention in Psychosis services: Findings from the Social Epidemiology of Psychoses in East Anglia [SEPEA] Study. *American Journal of Psychiatry* 2017;**174**:143–53. DOI: https://doi.org/10.1176/appi.ajp.2016.16010103

Lally 2017
Lally J, Ajnakina O, Stubbs B, Culllinane M, Murphy KC, Gaughran F, et al. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. *British Journal of Psychiatry* 2017;**211**(6):350–8. DOI: https://doi.org/10.1192/bjp.bp.117.201475

Leon 2006
Leon AC, Mallinckrodt CH, Chuang-Stein C, Archibald DG, Archer GE, Chartier K. Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology. *Biological Psychiatry* 2006;**59**(11): 1001–5. [PUBMED: 16905632]

Leucht 2005a
Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean?. *Schizophrenia Research* 2005;**79**(2-3):231–8. [PUBMED: 15982856]

Leucht 2005b
Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel R. Clinical implications of brief psychiatric rating scale scores. *British Journal of Psychiatry* 2005;**187**:366–71. [PUBMED: 16199797]

Marshall 2000
Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *British Journal of Psychiatry* 2000;**176**: 249–52.

Marshall 2011
Marshall M, Rathbone J. Early intervention for psychosis. *Cochrane Database of Systematic Reviews* 2011, Issue 6. DOI: 10.1002/14651858.CD004718.pub3

Mason 1995
Mason P, Harrison G, Glazebrook C, Medley I, Dalkin T, Croudace T. Characteristics of outcome in schizophrenia at 13 years. *British Journal of Psychiatry* 1995;**167**(5):596–603.

Meltzer 2002
Meltzer H, Singleton N, Lee A, Bebbington P, Brugha T, Jenkins R. *The Social and Economic Circumstances of Adults with Mental Disorders*. London: Stationery Office, 2002.

Morgan 2014
Morgan C, Lappin J, Heslin M, Donoghue K, Lomas B, Reuninghaus U, et al. Reappraising the long-term course and outcome of psychotic disorders: the AESOP-10 study. *Psychological Medicine* 2014;**44**(13):2713–26. DOI: https://doi.org/10.1017/S0033291714000282

Murray 1996
Murray C, Lopez AD. Evidence-based health policy - lessons from the Global Burden of Disease Study. *Science* 1996;**274**(5288):740–3.

NHS England 2015
NHS England, National Collaborating Centre for Mental Health, National Institute for Health and Care Excellence. *Implementing the Early Intervention in Psychosis Access and Waiting Time Standard (Guidance)*. 1. NHS England, 2015.

NICE 2014
National Institute for Health and Care Excellence. Psychosis and schizophrenia in adults: prevention and management (Clinical guideline). nice.org.uk/guidance/cg178. National Institute for Health and Care Excellence, 2014.

Overall 1962
Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychological Reports* 1962;**10**:799–812.

RCPsych 2016
Brightey-Gibbons F, Hodge S, Palmer L. *Standards for Early Intervention in Psychosis Services*. Pilot. London: Royal College of Psychiatrists, 2016.

Ruggeri 2015
Ruggeri M, Bonetto C, Lasalvia A, Fioritti A, de Girolamo G, Santonastaso P, et al. Feasibility and effectiveness of a multi-element psychosocial intervention for first-episode psychosis: results from the cluster-randomized controlled GET UP PIANO trial in a catchment area of 10 million inhabitants. *Schizophrenia Bulletin* 2015;**41**(5):1192–203.

Schünemann 2011
Schünemann HJ, Oxman AD, Vist GE, Higgins JP, Deeks JJ, Glasziou P, et al. Chapter 12: Interpreting results and drawing conclusions. In Higgins JP, Green S editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Shokraneh 2017
Shokraneh F, Adams CE. Study-based registers of randomized controlled trials: Starting a systematic review with data extraction or meta-analysis. *BioImpacts* 2017;**7** (4):209–17. DOI: 10.15171/bi.2017.25

Shokraneh 2018
Shokraneh F, Adams CE. Gallstone, snake venom and witchcraft for schizophrenia: the challenges of classifying

[schizophrenia] trials. *Evidence-Based Medicine* 2018;**23** (Suppl 1):A18. DOI: 10.1136/bmjebm-2018-111024.36

Sterne 2011

Sterne JA, Egger M, Moher D, editor(s). Chapter 10: Addressing reporting biases. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Intervention. Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Ukoumunne 1999

Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PGJ. Methods for evaluating area-wide and organisation-based intervention in health and health care: a systematic review. *Health Technology Assessment* 1999;**3**(5): iii–92.

WHO 2004

World Health Organization. *ICD-10: international statistical classification of diseases and related health problems: tenth revision*. Geneva: World Health Organization, 2004.

WHO 2018

World Health Organization. *ICD-11: international statistical classification of diseases and related health problems: eleventh revision*. Geneva: World Health Organization, 2018.

Wiersma 1998

Wiersma D, Nienhuis F, Slooff C, Giel R. Natural course of schizophrenia disorders: a 15-year follow-up of a Dutch incidence cohort. *Schizophrenia Bulletin* 1998;**24**:75–85.

Xia 2009

Xia J, Adams CE, Bhagat N, Bhagat V, Bhoopathi P, El-Sayeh H, et al. Loss to outcomes stakeholder survey: the LOSS study. *Psychiatric Bulletin* 2009;**33**(7):254–7.

Yung 2008

Yung AR, Nelson B, Stanford C, Simmons MG, Cosgrave EM, Killackey E, et al. Validation of "prodromal" criteria to detect individuals at ultra high risk of psychosis: 2 year follow-up. *Schizophrenia Research* 2008;**105**(1 - 3):10–7.

* *Indicates the major publication for the study*

CONTRIBUTIONS OF AUTHORS

Stephen Puntis: developed and wrote the protocol.

Amedeo Minichino: developed and helped write the protocol.

Franco De Crescenzo: development of the protocol, checking final draft.

Andrea Cipriani: development of the protocol, advisor, checking final draft.

Belinda Lennox: development of the protocol, advisor, checking final draft.

DECLARATIONS OF INTEREST

Stephen Puntis: none.

Amedeo Minichino: none.

Franco De Crescenzo: none.

Andrea Cipriani: none.

Belinda Lennox: I work clinically in an early intervention in psychosis service, and am clinical lead for early intervention in psychosis for NHS England. I have no other declarations of interest.

SOURCES OF SUPPORT

Internal sources

- University of Oxford, Oxford, UK.

Employs review authors Stephen Puntis, Amedeo Minichino, Andrea Cipriani, and Belinda Lennox

- NIHR Post Doctoral Fellowship Scheme - Grant number PDF-2017-10-029, UK.

Funds review author Stephen Puntis

- MRC-Clarendon, UK.

Funds review author Amedeo Minichino for DPhil

- NIHR, UK.

Provides additional funding to review author Belinda Lennox as clinical director for NIHR CRN: Thames Valley and South Midlands

- Department of Epidemiology of the Regional Health Service, Italy.

Provides funding for review author Franco De Crescenzo for the development of National clinical guidelines

External sources

- None, Other.