Cochrane Library

Cochrane Database of Systematic Reviews

Placebo versus no treatment for people with schizophrenia (Protocol)

Takeuchi H, Shimomura Y, Kikuchi Y, Nomura N, Hird E, Wu H, Agid O, Leucht S

Takeuchi H, Shimomura Y, Kikuchi Y, Nomura N, Hird E, Wu H, Agid O, Leucht S.
Placebo versus no treatment for people with schizophrenia (Protocol).
Cochrane Database of Systematic Reviews 2023, Issue 1. Art. No.: CD015403.
DOI: 10.1002/14651858.CD015403.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Protocol]

Placebo versus no treatment for people with schizophrenia

Hiroyoshi Takeuchi[1], Yutaro Shimomura[1], Yuhei Kikuchi[1], Nobuyuki Nomura[2], Emily Hird[3], Hui Wu[2], Ofer Agid[4,5,6], Stefan Leucht[2]

[1]Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan. [2]Section for Evidence Based Medicine in Psychiatry and Psychotherapy, Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, Munich, Germany. [3]Institute of Psychiatry, Psychology and Neuroscience, King's College, London, UK. [4]Schizophrenia Program, Centre for Addiction and Mental Health, Toronto, Canada. [5]Department of Psychiatry, University of Toronto, Toronto, Canada. [6]Institute of Medical Science, University of Toronto, Toronto, Canada

Contact: Hiroyoshi Takeuchi, hirotak@dk9.so-net.ne.jp.

Editorial group: Cochrane Schizophrenia Group.
Publication status and date: New, published in Issue 1, 2023.

Citation: Takeuchi H, Shimomura Y, Kikuchi Y, Nomura N, Hird E, Wu H, Agid O, Leucht S. Placebo versus no treatment for people with schizophrenia (Protocol). *Cochrane Database of Systematic Reviews* 2023, Issue 1. Art. No.: CD015403. DOI: 10.1002/14651858.CD015403.

ABSTRACT

Objectives

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

To compare the effects of placebo versus no treatment in people with schizophrenia.

BACKGROUND

Description of the condition

Schizophrenia is a severe mental disorder presenting with positive symptoms such as hallucinations, delusions, and disorganisation, negative symptoms such as diminished expression and amotivation, and cognitive impairment. Its prevalence is 0.5% with a slightly higher rate in males than females (McGrath 2008). The most frequent age at onset is early 20s for males and late 20s for females (Dobbs 2010). It is a chronic and recurrent disorder; more than 60% of people with schizophrenia experience relapse (Morgan 2014) with a recovery rate of 13.5% (Jääskeläinen 2013). The course of illness consists of premorbid stage (no or few symptoms), prodromal stage (attenuated symptoms), syndromal stage (psychotic symptoms), and chronic or residual stage (psychotic symptoms, negative symptoms, cognitive symptoms, and functional disability) (Lieberman 2018). Schizophrenia is a disorder with high burden; the disability-adjusted life years (DALYs) is the third highest among mental disorders in 2019 (vizhub.healthdata.org/gbd-compare/).

Description of the intervention

Placebo is identical to the drug intervention formed as tablets, capsules, and injections in appearance but does not contain active compounds (i.e. pharmacological placebo). Placebo is used commonly in clinical trials investigating pharmacological interventions as a control to detect true intervention effects; indeed, placebo can improve acute and chronic clinical conditions (Krogsbøll 2009) despite no inclusion of active compounds.

In the 'no treatment' condition participants do not receive any pharmacological (including placebo) or psychological interventions. Placebo can be superior to no treatment for acute and chronic clinical conditions (Krogsbøll 2009).

Antipsychotics are the mainstay of treatment for schizophrenia not only to improve acute symptoms (Huhn 2019) but also to prevent relapse (Ceraso 2020). Nonetheless, placebo can improve acute symptoms to a lesser degree than antipsychotics (Agid 2013; Rutherford 2014) with a small-to-moderate effect size of approximately 0.3.

However, it is not known how placebo compares to no treatment for improvement of symptoms and prevention of relapse in people with schizophrenia.

How the intervention might work

Although placebo does not contain active compounds, placebo can improve acute and chronic clinical conditions (Krogsbøll 2009). The placebo effects are exerted by individual patient and clinician factors and interaction between the patient, clinician, and treatment environment (Finniss 2010). The former includes patient's and clinician's beliefs, expectations, desire for symptom change, and past experiences (Finniss 2010). The latter includes clinician–patient relationship factors such as communication, empathy, reassurance, bedside manner, and enthusiasm; and treatment environment factors such as location, type, and nature of treatment (e.g. method of drug delivery, use of technological devices, and therapeutic procedure) (Finniss 2010).

In meta-analyses examining placebo effects in people with schizophrenia (Agid 2013; Rutherford 2014), two factors influencing the placebo effects were identified: individual patient factors and study factors. For the former, a younger age, shorter illness duration, and greater baseline symptom severity were associated with a greater placebo response (Agid 2013). For the latter, a later publication year, shorter study duration, and larger study site/centre number were associated with a greater placebo response (Agid 2013).

Why it is important to do this review

Placebo can improve acute symptoms of schizophrenia to a lesser degree than antipsychotics (Agid 2013; Rutherford 2014). On the other hand, switching to placebo is associated with a higher risk of relapse (Ceraso 2020) and symptom exacerbation (Takeuchi 2017) than continuing antipsychotics, although some people with schizophrenia prefer the idea of discontinuing/reducing antipsychotics at some time (Crellin 2022). However, it is not known whether placebo is superior to no treatment for improvement of symptoms and prevention of relapse. Although there have been some randomised controlled trials (RCTs) comparing drug intervention, placebo, and no treatment in people with schizophrenia (Pearl 1956; Sibilio 1957; Whittaker 1963) and a recent meta-analysis comparing placebo and no treatment in people with psychiatric disorders (Faltinsen 2022), no meta-analysis has been conducted to compare the effects of placebo versus no treatment specifically in people with schizophrenia.

This review will synthesise the available evidence comparing the effects of placebo versus no treatment. By addressing the benefits and risks of placebo use in clinical practice and research, this review will contribute to the development of optimal trial design in people with schizophrenia.

OBJECTIVES

To compare the effects of placebo versus no treatment in people with schizophrenia.

METHODS

Criteria for considering studies for this review

Types of studies

We will include randomised controlled trials (RCTs) comparing placebo and no treatment with a minimum study duration of one week. We expect that these trials will also have a drug arm, but this is not required for inclusion in this review. We will include double-blind RCTs, where double-blind refers to the placebo and drug arms, but not to the no treatment arm, given that participants in the no treatment group cannot be masked because they are aware they are not receiving any intervention. We will exclude quasi-randomised trials, such as those where allocation is undertaken on surname. We will include randomised cross-over studies but use only data up to the point of first cross-over because of the likely carry-over effects of the treatments (Elbourne 2002).

Types of participants

We will include people with schizophrenia spectrum disorders (i.e. schizophrenia, schizophreniform disorder, and schizoaffective disorder), regardless of specific conditions (e.g. those with first episode, treatment resistant, and predominant/prominent

negative symptoms). We will include studies if they include 70% or more participants with schizophrenia spectrum disorders. There is no clear evidence that the schizophrenia spectrum disorders are caused by fundamentally different disease processes or require different treatment approaches (Carpenter 1994).

We will include both studies involving participants with acute exacerbation of schizophrenia (i.e. acute phase studies) and studies involving participants with stable condition of schizophrenia (i.e. maintenance phase studies). These two types of studies and participants will be analysed in two separate comparisons.

Types of interventions

Placebo

Placebo is identical to the intervention in appearance but does not contain active compounds (i.e. pharmacological placebo). We will limit placebo to pharmacological placebo and will not include other placebos such as psychological and physical placebos, which are included in a recent meta-analysis (Faltinsen 2022), because we are interested in placebos that are blinded to both participants/patients, raters/assessors, and researchers/clinicians (the person who delivers psychological/physical placebo cannot be masked due to the nature of the interventions).

No treatment

In the no treatment condition, the participants do not receive any pharmacological (including placebo) or psychological interventions.

This includes both 1) no treatment: participants are not offered the trial active intervention but they are assessed a number of times during the trial, and they are not promised treatment after the end of the trial and 2) waiting list: participants are not offered the trial active intervention but they are assessed a number of times during the trial, and they are usually promised that the active intervention will be offered to them after the end of the trial.

Any usual care (e.g. pharmacological/psychological treatment) will be allowed as long as both intervention and control groups receive it.

Types of outcome measures

Acute phase studies will be analysed at endpoint only.

Maintenance phase studies will be classified into up to three months (short term); up to six months (medium term); and more than six months (long term).

Primary outcomes

1. Average endpoint or change score on overall mental state scale (acute phase studies).
2. Study-defined relapse at endpoint (maintenance phase studies).
3. Leaving the study early due to adverse events related to side effects.

Secondary outcomes

1. Mental state.
1.1 General.
1.1.1 Average endpoint or change score on overall mental state scale (e.g. the Positive and Negative Syndrome Scale (PANSS) and the Brief Psychiatric Rating Scale (BPRS)) (acute phase studies).
1.1. 2 Number of participants with clinically important change in overall mental state.

1.2 Specific.
1.2.1 Average endpoint or change score on specific mental state scale (e.g. positive symptoms and negative symptoms assessed by PANSS subscales).

2. Relapse.
2.1 Study-defined relapse at endpoint (maintenance phase studies).
2.2 Study-defined relapse at three months, four to six months, seven to 12 months, and more than one year (maintenance phase studies).

3. Leaving the study early.
3.1 Due to any reason.
3.2 Due to inefficacy.
3.3 Due to adverse events related to side effects.

4. Global state.
4.1 Average endpoint or change score on global state scale (e.g. Clinical Global Impression-Severity (CGI-S) and Global Assessment of Functioning (GAF)).
4.2 Number of participants with clinically important change in global state scale (e.g. CGI-S and GAF).

5. Functioning.
5.1 Average endpoint or change score on general functioning scale (e.g. Quality of Life Scale (QLS), Personal and Social Performance (PSP), and Social and Occupational Functioning Assessment Scale (SOFAS)).

6. Quality of life.
6.1 Average endpoint or change score on quality of life scale (e.g. EuroQol (EQ-5D)).

7. Hospitalisation.
7.1 Number of participants hospitalised.

8. Death.
8.1 Due to any reason.
8.2 Due to natural causes.
8.3 Due to suicide.

9. Adverse events.
9.1 Number of participants with at least one adverse event.

9.2 Specific: movement disorders.
9.2.1 Akathisia.
9.2.2 Dyskinesia.
9.2.3 Dystonia.
9.2.4 Parkinsonism.
9.2.5 Use of antiparkinson medication.

9.3 Specific: sedation.
9.4 Specific: weight gain.
9.5 Specific: sexual dysfunction.
9.6 Specific: anticholinergic side effects.

10. Treatment satisfaction.

10.1 Average endpoint or change score on participants' satisfaction scale (e.g. Drug Attitude Inventory (DAI-10)).
10.2 Number of participants satisfied with treatment.

Search methods for identification of studies

Electronic searches

The Information Specialist will search the register using the following search strategy:

Placebo and *no treatment* in Intervention Field of STUDY

In such a study-based register, searching the major concept retrieves all the synonyms and relevant studies. This is because the studies have already been organised, based on their interventions, and linked to the relevant topics (Shokraneh 2017). This allows rapid and accurate searches that reduce waste in the next steps of systematic reviewing (Shokraneh 2019).

Following the methods from Cochrane (Lefebvre 2021), the Information Specialist compiles this register from systematic searches of major resources and their monthly updates (unless otherwise specified).

- Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library.
- MEDLINE.
- Embase.
- Allied and Complementary Medicine (AMED).
- BIOSIS.
- Cumulative Index to Nursing and Allied Health Literature (CINAHL).
- PsycINFO.
- PubMed.
- US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov).
- World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) (www.who.int/ictrp).
- ProQuest Dissertations and Theses A&I and its quarterly update.

The register also includes handsearches and conference proceedings (see Group's website). It does not place any limitations on language, date, document type or publication status.

Searching other resources

Reference searching

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

Personal contact

We will contact the first author of each included study for information regarding unpublished trials and additional data. We will note the outcome of this contact in the 'Characteristics of included studies' or 'Characteristics of studies awaiting classification' tables.

Data collection and analysis

Selection of studies

At least two of three review authors (Hiroyoshi Takeuchi (HT), Yutaro Shimomura (YS), Yuhei Kikuchi (YK)) will independently inspect citations from the searches and identify relevant abstracts. Where disputes arise, we will acquire the full report for more detailed scrutiny. At least two of three review authors (HT, YS, YK) will then obtain and independently inspect full reports of the abstracts or reports meeting the review criteria. Where it is not possible to resolve disagreement by discussion, we will discuss with the senior author of the team to resolve it. If following discussion with the third author disagreement exists, we will attempt to contact the authors of the study concerned for clarification. All decisions will be documented.

Data extraction and management

Extraction

At least two of three review authors (HT, YS, YK) will independently extract data from all included studies. 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. We will discuss any disagreement. Where it is not possible to resolve disagreements by discussion, we will discuss with the senior author. All decisions will be documented. If necessary, we will attempt to contact authors through an open-ended request in order to obtain missing information or for clarification. HT and Stefan Leucht (SL) will help clarify issues regarding any remaining problems and we will document these final decisions.

Management

Forms

We will extract data onto standard, pre-designed, simple forms.

Scale-derived data

We will include continuous data from rating scales only if:

- the psychometric properties of the measuring instrument have been described in a peer-reviewed journal (Marshall 2000);
- the measuring instrument has not been written or modified by one of the trialists for that particular trial; and
- 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, we will include subscores of scales if these were validated or if these were pre-defined in a scale such as the positive symptom, negative symptom and general symptom scores of PANSS (Kay 1986).

Ideally the measuring instrument should either be i) a self-report or ii) completed by an independent rater or relative (not the therapist). We, however, realise that this is not often reported clearly.

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 change data, and only use endpoint data if the former are not available. If necessary,

we will combine endpoint and change data in the analysis. This procedure is possible when using mean differences (MDs) (Deeks 2020) and also when using standardised mean differences (SMDs). Although theoretically the combination of change and endpoint data when SMDs are used can be problematic, meta-epidemiological research has shown that on average no major over- or underestimations can be expected (da Costa 2013).

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.

For endpoint data from studies including fewer than 200 participants, we will calculate the observed mean minus the lowest possible value of the scale and divide this by the standard deviation (Higgins 2020).

For example, in a scale that has possible lowest values higher than 0 (such as PANSS), which can have values from 30 to 210 (Kay 1986), we will subtract the minimum score (in this case 30) from the observed mean, and then divide by the standard deviation. In a scale that has 0 as minimum possible score, we will divide the observed mean by the standard deviation.

For this calculation, we will check the original publication of the scales referenced in the studies, in order to understand if they can have a lowest possible score different from 0, and the adjustment described above is needed or not.

If the ratio obtained is lower than one, it strongly suggests that the data are skewed. If it is higher than one but less than two, there is a suggestion that the data are skewed; if the ratio is larger than two we will include these data, because it is less likely that they are skewed (Altman 1996).

Where there is suggestion of skewedness (ratio < 2), we will exclude the relevant studies in a sensitivity analysis to check if they have an impact on the results (see Sensitivity analysis for further details).
These skewed results will nevertheless be reported in 'other data tables'.
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 data are skewed.

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).

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 BPRS (Overall 1962), or the PANSS (Kay 1986) which correspond to 'much improved' according to the clinical global impressions (CGI, Guy 1976) of raters, could be considered as a clinically significant response (Leucht 2005a; Leucht 2005b), in particular for acutely ill people. However, we assumed that most participants included in the studies would be chronic. For these even small improvements such as an at least 20% or 30% reduction of the BPRS or PANSS which correspond to 'minimally improved' on the CGI (Leucht 2005a; Leucht 2005b) may be meaningful. Therefore, these cut-offs were chosen as the primary ones. If data based on these thresholds are not available, we will use the primary cut-off presented by the original authors, because the exact cut-off is not so important in a meta-analysis using risk ratios or odds ratios as effect sizes (Furukawa 2010).

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 placebo. Where keeping to this makes it impossible to avoid outcome titles with clumsy double-negatives (e.g. 'not un-improved'), 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

At least two of three review authors (HT, YS, YK) 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 2017). 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, figures, and the summary of findings table(s).

Measures of treatment effect

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 (Boissel 1999); and that odds ratios 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.

Continuous data

For continuous outcomes we will estimate MD between groups, in particular when natural (such as days, kilograms, etc.) are used. We prefer not to calculate effect size measures (SMD). However, if scales of very considerable similarity are used, we will presume there is a small difference in measurement, and we will calculate SMD. It should be noted that SMD can be transformed to MD by using the formula MD = SMD x standard deviation of the scale of interest (Higgins 2020).

Unit of analysis issues

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 intra-class 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 intra-class 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 intra-class 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 intra-class correlation coefficients and relevant data documented in the report into account, synthesis with other studies will be possible using the generic inverse variance technique.

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 wash-out 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.

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 2020), as implemented in Review Manager (RevMan) calculator (Review Manager 2020). Where additional treatment arms are not relevant, we will not reproduce these data.

Dealing with missing data

Overall loss of credibility

Although at some degree of loss of follow-up, data lose credibility (Xia 2009), we will not exclude studies based on this.

However, if more than 50% of data are unaccounted for (lost to follow-up) we will exclude these studies in a Sensitivity analysis. If 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 down-rating certainty (and do not exclude the study in the sensitivity analysis). Finally, we will also downgrade certainty within the summary of findings table(s) should the loss be 25% to 50% in total.

Binary

We will present data on a 'once-randomised-always-analyse' basis (an intention-to-treat analysis (ITT)). We will undertake a Sensitivity analysis excluding studies using completer analyses. If the authors applied such a strategy, we would use their results. If the original authors presented only the results of the per-protocol or completer population, we assumed that those participants lost to follow-up would not have had the outcome of interest if they had stayed in the study.

Continuous

Standard deviations

If standard deviations (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 2020). When only the SE is reported, SDs are calculated by the formula SD = SE * √(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 2020). 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.

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

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.

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.

Statistical heterogeneity

Visual inspection

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

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 (Chapter 10, *Cochrane Handbook for Systematic Reviews of Interventions*) (Deeks 2020). 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).

Protocol versus full study

We will try to locate protocols of included randomised trials. 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. If details from Clinicaltrials.gov and WHO registry (ICTRP) are available, they will be included in the search results and these can be used to compare the differences between planned methods and published results.

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 ten 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. In a sensitivity analysis of the primary outcome we will apply the fixed model.

Subgroup analysis and investigation of heterogeneity

Subgroup analyses

Primary outcomes

Where possible we will perform subgroup analyses for:

- different illness phases, such as people with first episode schizophrenia spectrum disorders and chronically ill;
- different age groups (children and adolescents, adults, and elderly).

Investigation of heterogeneity

We will report if inconsistency is high. Firstly, we will investigate whether data have been entered correctly. If this is the case, the following strategies will be considered: a) pool the data despite the heterogeneity: an example where this strategy may be appropriate is the effects of all studies are in the same direction. In other words, the heterogeneity reflects the degree of an effect rather than its direction which is less problematic. Another example is when heterogeneity can be explained by appropriate subgroup analyses; b) exclude outlying studies: this strategy may apply, if reinspection of such studies reveal methodological or clinical differences that were previously overlooked; and c) not pool the studies. All decisions in this regard will be described and discussed.

We will also perform meta-regressions to examine the associations between effect sizes of differences between placebo and no treatment, and age, illness duration, baseline symptom severity, publication year, study duration, and study site/centre number.

Sensitivity analysis

Where possible we will perform sensitivity analyses for the primary outcomes in order to explore the influence of the following factors on effect size. If there are substantial differences in the direction or

precision of effect estimates in any of the sensitivity analyses listed below, we will discuss them in the discussion section.

1. Implication of randomisation: we will exclude trials that are described as double-blind, but where randomisation is not explicitly mentioned.
2. Assumptions for missing data: we will exclude studies using completer analyses only (see Dealing with missing data).
3. Loss to follow-up: we will exclude studies where the overall loss of data was greater than 50%.
4. Risk of bias: we will analyse the effects of excluding trials that are at overall high risk of bias (see Assessment of risk of bias in included studies) for the meta-analysis of the primary outcomes.
5. Imputed values: we will also undertake a sensitivity analysis excluding trials where we use imputed values for ICC in calculating the design effect in cluster-randomised trials or where SDs were imputed.
6. Fixed- and random-effects: we will synthesise data using a random-effects model; however, we will also synthesise data for the primary outcomes using a fixed-effects model to evaluate whether this alters the significance of the results.
7. Skewed data: we will perform a sensitivity analysis excluding studies for which there is suggestion of skewedness (mean/SD ratio lower than 2 - see Data extraction and management). If this changes the results in comparison with the main analysis (from significantly favouring the intervention to significantly favouring the control, or viceversa), we will exclude these studies also from the main analysis, and present their data in 'Other data' tables.

Summary of findings and assessment of the certainty of the evidence

We will use the GRADE approach to interpret findings (Schünemann 2020); and will use GRADEpro GDT (GRADEpro GDT 2022) to export data from our review (Review Manager 2020) to create a summary of findings table(s). These tables provide outcome-specific information concerning the overall certainty of the 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. The overall risk of bias judgements will be used to feed into the GRADE assessment. We aim to select the following main outcomes for inclusion in the summary of findings table.

1. Average endpoint or change score on overall mental state scale (acute phase studies).
2. Study-defined relapse at endpoint (maintenance phase studies).
3. Leaving the study early due to adverse events related to side effects.
4. Leaving the study early due to any reason.
5. Number of participants hospitalised.
6. Average endpoint or change score on general functioning scale.
7. Average endpoint or change score on quality of life scale.

We will use short-term data for acute phase studies and endpoint data for maintenance phase studies.

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

ACKNOWLEDGEMENTS

The Cochrane Schizophrenia Group Editorial Base situated across the University of Melbourne, Australia, the Technical University of Munich, Germany, and the University of 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.

Editorial and peer-reviewer contributions

Cochrane Schizophrenia supported the authors in the development of this protocol.

The following people conducted the editorial process for this article.

- Sign-off Editor (editorial checks, final editorial decision): Irene Bighelli, Technical University of Munich, Germany.
- Handling Editor (selected peer reviewers, collated peer-reviewer comments, provided editorial guidance to authors, edited the article): Ajit Kumar, Latrobe Regional Hospital, Australia.
- Copy Editor (copy-editing and production): Luisa M Fernandez Mauleffinch, Copy Edit Support, Cochrane Central Production Service.
- Information Specialist (search strategy and search results): Anne Parkhill, University of Melbourne, Australia.
- Peer-reviewers[a] (clinical/content review, provided comments and recommended an editorial decision): Anna Ceraso, ASST Spedali Civili Brescia, Italy, Department of Mental Health and Addiction Services; Andreas S Lappas, Velindre University NHS Trust, United Kingdom, Department of Psychiatry, University of Thessaly, Medical School, Thessaly, Greece.

[a]Peer-reviewers are members of Cochrane Schizophrenia, and provided peer-review comments on this article, but they were not otherwise involved in the editorial process or decision-making for this article.

REFERENCES

Additional references

Agid 2013

Agid O, Siu CO, Potkin SG, Kapur S, Watsky E, Vanderburg D, et al. Meta-regression analysis of placebo response in antipsychotic trials, 1970-2010. *American Journal of Psychiatry* 2013;**170**:1335-44.

Altman 1996

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

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. [PMID: 10667106]

Carpenter 1994

Carpenter WT, Buchanan RW. Schizophrenia. *New England Journal of Medicine* 1994;**330**:681-90.

Ceraso 2020

Ceraso A, Lin JJ, Schneider-Thoma J, Siafis S, Tardy M, Komossa K, et al. Maintenance treatment with antipsychotic drugs for schizophrenia. *Cochrane Database of Systematic Reviews* 2020, Issue 8. Art. No: CD008016. [DOI: 10.1002/14651858.CD008016.pub3]

Crellin 2022

Crellin NE, Priebe S, Morant N, Lewis G, Freemantle N, Johnson S, et al. An analysis of views about supported reduction or discontinuation of antipsychotic treatment among people with schizophrenia and other psychotic disorders. *BMC Psychiatry* 2022;**22**(1):185. [PMID: 35291964]

da Costa 2013

da Costa BR, Nüesch E, Rutjes AW, Johnston BC, Reichenbach S, Trelle S, et al. Combining follow-up and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. *Journal of Clinical Epidemiology* 2013;**66**(8):847-55. [DOI: 10.1016/j.jclinepi.2013.03.009]

Deeks 2000

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

Deeks 2020

Deeks JJ, Higgins JP, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from training.cochrane.org/handbook.

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.

Dobbs 2010

Dobbs D. Schizophrenia: the making of a troubled mind. *Nature* 2010 ;**468**:154–6.

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.

Faltinsen 2022

Faltinsen E, Todorovac A, Staxen Bruun L, Hróbjartsson A, Gluud C, Kongerslev MT, et al. Control interventions in randomised trials among people with mental health disorders. *Cochrane Database of Systematic Reviews* 2022, Issue 4. Art. No: MR000050. [DOI: 10.1002/14651858.MR000050.pub2]

Finniss 2010

Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Biological, clinical, and ethical advances of placebo effects. *Lancet* 2010;**375**:686-95.

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.

Furukawa 2010

Furukawa TA, Akechi T, Wagenpfeil S, Leucht S. Relative indices of treatment effect may be constant across different definitions of response in schizophrenia trials. *Schizophrenia Research* 2011;**126**(1-3):212-9. [DOI: 10.1016/j.schres.2010.10.016]

GRADEpro GDT 2022 [Computer program]

McMaster University (developed by Evidence Prime) GRADEpro GDT. Version accessed 3 March 2022. Hamilton (ON): McMaster University (developed by Evidence Prime), 2022. Available at gradepro.org.

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.

Guy 1976

Guy U. ECDEU Assessment Manual for Psychopharmacology. Revised edition. Rockville, Maryland, USA: National Institute of Mental Health, 1976.

Higgins 2003

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

Higgins 2017

Higgins JP, Altman DG, Sterne JA, editor(s). Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.2 (updated June 2017). The Cochrane Collaboration, 2017. Available from training.cochrane.org/handbook.

Higgins 2020

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from training.cochrane.org/handbook.

Huhn 2019

Huhn M, Nikolakopoulou A, Schneider-Thoma J, Krause M, Samara M, Peter N, et al. Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis. *Lancet* 2019;**394**:939–51.

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. [PMID: 19438480]

Jääskeläinen 2013

Jääskeläinen E, Juola P, Hirvonen N, McGrath JJ, Saha S, Isohanni M, et al. A systematic review and meta-analysis of recovery in schizophrenia. *Schizophrenia Bulletin* 2013;**39**:1296–306.

Kay 1986

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

Krogsbøll 2009

Krogsbøll LT, Hróbjartsson A, Gøtzsche PC. Spontaneous improvement in randomised clinical trials: meta-analysis of three-armed trials comparing no treatment, placebo and active intervention. *BMC Medical Research Methodology* 2009;**9**:1.

Lefebvre 2021

Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M, et al. Chapter 4: Searching for and selecting studies. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook.

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. [PMID: 16905632]

Leucht 2005a

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. [PMID: 16199797]

Leucht 2005b

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. [PMID: 15982856]

Lieberman 2018

Lieberman JA, First MB. Psychotic disorders. *New England Journal of Medicine* 2018;**379**:270-80. [PMID: 30021088]

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.

McGrath 2008

McGrath J, Saha S, Chant D, Welham J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiologic Reviews* 2008;**30**:67–76.

Morgan 2014

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

Overall 1962

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

Pearl 1956

Pearl D, Kamp HV, Olsen AL, Greenberg PD, Armitage SG. The effects of reserpine on schizophrenic patients. *American Journal of Psychiatry* 1956 ;**112**(11):936. [PMID: 13313802]

Review Manager 2020 [Computer program]

The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.4. The Cochrane Collaboration, 2020. Available at revman.cochrane.org.

Rutherford 2014

Rutherford BR, Pott E, Tandler JM, Wall MM, Roose SP, Lieberman JA. Placebo response in antipsychotic clinical trials: a meta-analysis. *JAMA Psychiatry* 2014;**71**:1409-21.

Schünemann 2020

Schünemann HJ, Higgins JP, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from training.cochrane.org/handbook.

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 2019

Shokraneh F, Adams CE. Study-based registers reduce waste in systematic reviewing: discussion and case report. *Systematic Reviews* 2019;**8**:129. [DOI: 10.1186/s13643-019-1035-3]

Sibilio 1957

Sibilio JP, Andrew G, Dart D, Moore KB, Stehman VA. Treatment of chronic schizophrenia with promazine hydrochloride. *AMA Archives of Neurology and Psychiatry* 1957;**78**(4):419-24. [PMID: 13457517]

Takeuchi 2017

Takeuchi H, Kantor N, Sanches M, Fervaha G, Agid O, Remington G. One-year symptom trajectories in patients with stable schizophrenia maintained on antipsychotics versus placebo: meta-analysis. *British Journal of Psychiatry* 2017;**211**:137–43.

Ukoumunne 1999

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

Whittaker 1963

Whittaker CB, Hoy RM. Withdrawal of perphenazine in chronic schizophrenia. *British Journal of Psychiatry* 1963;**109**:422-7. [PMID: 14000409]

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.

CONTRIBUTIONS OF AUTHORS

- Hiroyoshi Takeuchi: protocol development.
- Yutaro Shimomura: protocol development.
- Yuhei Kikuchi: protocol development.
- Nobuyuki Nomura: protocol development.
- Emily Hird: protocol development.
- Hui Wu: protocol development.
- Ofer Agid: protocol development.
- Stefan Leucht: protocol development.

DECLARATIONS OF INTEREST

- Hiroyoshi Takeuchi (HT): HT is an Editor of Cochrane Schizophrenia. He was not involved in the editorial process of the present protocol.
- Yutaro Shimomura (YS): YS has no conflicts of interest.
- Yuhei Kikuchi (YK): YK has no conflicts of interest.
- Nobuyuki Nomura (NN): NN has no conflicts of interest.
- Emily Hird (EH): EK has no conflicts of interest.
- Hui Wu (HW): HW is an Editor of Cochrane Schizophrenia. She was not involved in the editorial process of the present protocol.
- Ofer Agid (OA): AO has no conflicts of interest.
- Stefan Leucht (SL): SL is an Editor of Cochrane Schizophrenia. He was not involved in the editorial process of the present protocol.

SOURCES OF SUPPORT

Internal sources

- National Institute for Health and Care Research, UK

  provided funding for Cochrane Schizophrenia Group

External sources

- None, Other

  None