Psychological interventions for positive symptoms in schizophrenia: protocol for a network meta-analysis of randomised controlled trials

Irene Bighelli,1 Georgia Salanti,2 Cornelia Reitmeir,1 Sofia Wallis,1 Corrado Barbui,3 Toshi A Furukawa,4 Stefan Leucht1

ABSTRACT

Introduction There is rising awareness that we need multidisciplinary approaches integrating psychological treatments for schizophrenia, but a comprehensive evidence based on their relative efficacy is lacking. We will conduct a network meta-analysis (NMA), integrating direct and indirect comparisons from randomised controlled trials (RCTs) to rank psychological treatments for schizophrenia according to their efficacy, acceptability and tolerability.

Methods and analysis We will include all RCTs comparing a psychological treatment aimed at positive symptoms of schizophrenia with another psychological intervention or with a no treatment condition (waiting-list and treatment as usual). We will include studies on adult patients with schizophrenia, excluding specific subpopulations (eg, first-episode patients or patients with psychiatric comorbidities). Primary outcome will be the change in positive symptoms on a published rating scale. Secondary outcomes will be acceptability (dropout), change in overall and negative symptoms of schizophrenia, response, relapse, adherence, depression, quality of life, functioning and adverse events. Published and unpublished studies will be sought through database searches, trial registries and websites. Study selection and data extraction will be conducted by at least two independent reviewers. We will conduct random-effects NMA to synthesize all available evidence for each outcome and obtain a comprehensive ranking of all treatments. NMA will be conducted in Stata and R within a frequentist analysis framework. The risk of bias in studies will be evaluated using the Cochrane Risk of Bias tool and the credibility of the evidence will be evaluated using an adaptation of the Grading of Recommendations Assessment, Development and Evaluation framework to NMA, recommended by the Cochrane guidance. Subgroup and sensitivity analyses will be conducted to assess the robustness of the findings.

Ethics and dissemination No ethical issues are foreseen. Results from this study will be published in peer-reviewed journals and presented at relevant conferences.

PROSPERO registration number CRD42017067795.

INTRODUCTION

Schizophrenia is a debilitating and often lifelong disorder that ranks among the top 20 causes of disability according to the World Health Report.1 Although pharmacological interventions have been the mainstay of treatment for schizophrenia, antipsychotics have a number of limitations (limited response, high incidence of disabling side effects and poor adherence to treatment)2 and are problematic in many situations (such as medical comorbidities, tolerability problems and pregnancy). Besides, there has been growing recognition of the importance of psychological processes in psychosis, both as contributors to onset and persistence, and in terms of the negative psychological impact of a diagnosis of schizophrenia on the individual’s well-being, psychosocial functioning and life opportunities. Psychological interventions
for psychosis and schizophrenia have been developed to address these aspects, and in accordance with guidelines from the National Institute for Health and Care Excellence in the UK and the Schizophrenia Patient Outcomes Research Team in the USA. Psychological treatments are widely regarded as a necessary intervention for schizophrenia.

A broad range of interventions that can be defined as ‘psychological’ have been studied in the treatment of schizophrenia. These interventions can be provided at different stages of the illness and address different aspects, like social and cognitive functioning, adherence to medication and symptoms of schizophrenia. Table 1 presents the panorama of existing systematic reviews of randomised controlled trials (RCTs) that have been conducted on the topic. These reviews have mainly included studies comparing the intervention under examination with so called no treatment conditions (wait-list and treatment as usual (TAU)). Other reviews included also active comparisons with other psychological treatments. An attempt to provide information on active comparisons was made by Turner and colleagues, who performed separate meta-analysis when there were at least five eligible RCTs comparing one intervention to another psychological intervention. However, all the available reviews applied pairwise meta-analysis as a method, being able to pool results only when a comparison of two treatments was considered in existing studies. The comparative efficacy and tolerability of the existing interventions have not been checked yet; as a result, it is still currently unclear which are the most efficacious, the most acceptable and the best tolerable psychological treatments for schizophrenia.

To overcome this gap in the current knowledge, a NMA would be necessary to consider both direct and indirect comparisons, and produce hierarchies of the effects of the various psychological treatments in the various efficacy and tolerability outcomes. Such hierarchies are essential for guidelines, which should ideally be able to indicate which treatment is likely to be the best, the second best and so on for a given outcome. Only the method of NMA can provide such hierarchies by combining all the randomised evidence. Our aim is to produce such a NMA of all psychological interventions for schizophrenia in multiple outcomes. We focus here on the interventions primarily aimed at treating positive symptoms in the acute phase of the illness.

Objectives
To estimate relative treatment effects and obtain a hierarchy for the psychological treatments in patients with schizophrenia, in terms of:
1. efficacy on positive symptoms.
2. acceptability.
3. other efficacy measures, such as overall symptoms, negative symptoms, response, relapse, adherence, depression, quality of life and functioning.
4. tolerability.

METHODS AND ANALYSIS
Criteria for considering studies for this review
Methods for this systematic review have been developed according to the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P) checklist, and the PRISMA extension statement for reporting of systematic reviews incorporating NMA of healthcare interventions. This systematic review and NMA is registered in the PROSPERO database; the record in PROSPERO will be updated with any amendment made to the protocol.

Types of studies
We will include all RCTs in which participants with schizophrenia received a psychological intervention as defined below (see Types of interventions section). Studies whose sequence generation was at high risk of bias (e.g., randomisation by the date of birth or day of the week) will be excluded. We will accept open and blinded RCTs; this choice is particularly relevant in trials on psychological interventions, in which in best case only the assessor of outcome can be blind, but not the therapist. Open RCTs will be excluded in a sensitivity analysis. We will include both trials in which psychological interventions were compared with a control condition and trials in which they were compared with another intervention. There will be no language restriction in order to avoid the problem of ‘language bias’. In case we retrieve references in languages in which we are not fluent, study authors will be contacted to check inclusion criteria and eventually ask for study data. As an exception, we will not search Chinese databases, since serious concerns have been raised on the trustworthiness of Chinese trials found in these databases. Chinese studies found in Western databases will be considered for inclusion. In the case of cross-over studies, we will use only the first cross-over phase in order to avoid the problem of carry-over effects which are very likely in schizophrenia and with psychological treatments. We will exclude cluster RCTs.

Types of participants
Our aim is to collect information on the efficacy of psychological treatments on patients with positive symptoms. In order to select this population, we operationalised the inclusion criteria as follows. We will include adults, however defined by study authors, with a diagnosis of schizophrenia or related disorders (such as schizoaffective disorders); there is no clear evidence that the latter schizophrenia-like psychoses are caused by fundamentally different disease processes or require different treatment approaches. We will include trials irrespective of the diagnostic criteria used. Here we will follow the strategy of the Cochrane Schizophrenia Group to include not only studies that used specific diagnostic criteria such as International Classification of Diseases, 10th revision or Diagnostic and Statistical Manual of Mental Disorders, 5th edition, because these criteria are not meticulously used in clinical routine.
## Table 1  Existing reviews about psychological treatments for schizophrenia

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Existing reviews</th>
<th>RCT*</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance and commitment therapy</td>
<td>Ongoing Cochrane review</td>
<td></td>
<td>TAU, pharmacological intervention and another psychosocial intervention</td>
</tr>
<tr>
<td>Adherence interventions</td>
<td>Gray et al</td>
<td>6</td>
<td>TAU and didactic health education</td>
</tr>
<tr>
<td>Active comparisons (befriending, CBT, cognitive remediation, psychoeducation, social skills training, supportive counselling)</td>
<td>Turner et al</td>
<td>48</td>
<td>Befriending, CBT, cognitive remediation, psychoeducation, social skills training, supportive counselling, family intervention, art therapy, body psychotherapy, occupational therapy and problem-solving therapy</td>
</tr>
<tr>
<td>Art therapy</td>
<td>Ruddy and Milnes</td>
<td>2</td>
<td>Standard care</td>
</tr>
<tr>
<td>Assertive community treatment</td>
<td>Marshall and Lockwood</td>
<td></td>
<td>TAU, hospital-based rehabilitation and case management</td>
</tr>
<tr>
<td>Befriending</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bibliotherapy</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body-oriented psychological therapy</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case management</td>
<td>Dieterich et al</td>
<td></td>
<td>Assertive community treatment, assertive outreach model, case management model and standard community care</td>
</tr>
<tr>
<td>CBT</td>
<td>Zimmermann et al (positive symptoms)</td>
<td>15</td>
<td>Waiting-list, TAU or another therapeutic treatment</td>
</tr>
<tr>
<td></td>
<td>Jones et al</td>
<td>20</td>
<td>Active (psychoeducation, family intervention, supportive psychotherapy, supportive counselling, cognitive remediation) and non-active control treatments (recreation and support, social activities, befriending and non-specific counselling)</td>
</tr>
<tr>
<td></td>
<td>Jauhar et al</td>
<td>52</td>
<td>Waiting-list, TAU or an intervention designed to control for the non-specific effects of psychotherapy (recreation and support, group support, befriending, supportive counselling/therapy, social activity therapy and goal-focused supportive contact) or active treatments (cognitive remediation and psychoeducation)</td>
</tr>
<tr>
<td></td>
<td>Van der Gaag et al (individually tailored)</td>
<td>18</td>
<td>Any control condition was accepted</td>
</tr>
<tr>
<td></td>
<td>Hazell et al (low intensity)</td>
<td>8</td>
<td>TAU and supportive psychotherapy</td>
</tr>
<tr>
<td></td>
<td>Kennedy and Xyrichis (auditory hallucinations)</td>
<td>2</td>
<td>Non-specialised therapy (focused on supportive interactions and social integration)</td>
</tr>
<tr>
<td>Cognitive remediation</td>
<td>Celia et al</td>
<td>45</td>
<td>TAU, active control (eg, computer games) another active treatment (eg, CBT)</td>
</tr>
<tr>
<td>Dance therapy</td>
<td>Ren and Xia</td>
<td>1</td>
<td>Standard care plus supportive counselling</td>
</tr>
<tr>
<td>Family interventions</td>
<td>Pitschel-Walz et al</td>
<td>25</td>
<td>TAU, patient intervention, other family interventions</td>
</tr>
<tr>
<td></td>
<td>Pharoah et al</td>
<td>25</td>
<td>TAU, discussion groups, psychoeducation, supportive psychotherapy, psychosocial support</td>
</tr>
<tr>
<td>Group psychotherapeutic treatments</td>
<td>Orfanos et al</td>
<td>34</td>
<td>TAU and other groups (active discussion group, support group, counselling group, occupational therapy group or problem-solving discussion group)</td>
</tr>
<tr>
<td>IPT</td>
<td>Roder et al</td>
<td>16</td>
<td>TAU, placebo-attention condition and other active treatments</td>
</tr>
<tr>
<td>Psychological and psychosocial interventions for negative symptoms in psychosis</td>
<td>Lutgens et al</td>
<td>95</td>
<td>TAU and active comparator (including psychoeducation, supportive therapy and cognitive remediation)</td>
</tr>
<tr>
<td>Metacognitive training</td>
<td>Eichner and Berna</td>
<td>12</td>
<td>TAU, wait-list control, supportive therapy, newspaper discussion group, CogPack (=cognitive remediation)</td>
</tr>
<tr>
<td>Mindfulness</td>
<td>Aust and Bradshaw</td>
<td>11†</td>
<td>Active control intervention (eg, befriending and progressive muscle relaxation) and TAU</td>
</tr>
<tr>
<td>Music therapy</td>
<td>Geretsegger et al</td>
<td>18</td>
<td>Placebo defined as an alternative therapy designed to control for effects of the therapist’s attention; TAU or no treatment</td>
</tr>
</tbody>
</table>

Continued
either. This decision should increase generalisability and representativeness.

Studies including participants with other diagnoses part of the psychosis spectrum will be included only if participants with a diagnosis of schizophrenia, schizophréniform or schizoaffective disorders were >80% of the participants considered. We will include studies recruiting patients with positive symptoms, either delusions, hallucinations or both, or in the phase of acute exacerbation of positive symptoms, however defined by inclusion criteria of the trial.

We will exclude studies focused on specific subpopulations of patients, such as (A) studies recruiting patients in which negative symptoms are predominant, according to authors’ definition, (B) studies on patients with concomitant psychiatric disorders or substance abuse, (C) studies recruiting patients with comorbid medical illnesses, (D) trials enrolling stable patients (relapse prevention studies), (E) studies on first-episode patients and (F) trials on patients who show prodromal signs of psychosis (also defined as ‘at risk for psychosis’).

Among other reasons, we exclude first-episode patients because they were found to have significantly higher response rates to treatments compared with chronic patients.18 19

**Types of interventions**

Any psychological intervention that occurs through interaction between therapist and patient, either face-to-face individually or in group, with the primary aim to reduce positive symptoms. Interventions with an explicit primary aim different from positive symptoms (eg, functioning, cognition, adherence to medication and knowledge of the illness) will be excluded. The identified treatments will be classified after identification of eligible studies. Psychological treatments will be compared with each other and to any non-pharmacological control condition considered in the included studies. Comparators will include the so called ‘treatment as usual’, waiting-list and inactive treatments. The effect of ‘non-active’ comparators will be analysed in a sensitivity analysis.20

Patients also receiving treatment as usual, including pharmacological interventions, will be included. If psychological treatments that we do not include among the interventions (eg, psychoeducation and supportive therapy) are used as control condition in the studies, they will be included as nodes in order to strengthen the network, but will not be part of our decision set.

**Outcome measures**

Outcomes will be measured at study endpoint, as defined in each study.

**Primary outcome**

Change in positive symptoms of schizophrenia, examined accordingly to the respective subscale of the Positive and Negative Syndrome Scale (PANSS), the Brief Psychiatric Rating Scale (BPRS) or the Scales for Assessment of Positive Symptoms or any other published scale.

As not all studies will have used the same scale, we will extract data according to the following hierarchy: mean change of the PANSS positive symptoms subscale from baseline to endpoint, if not available mean change of the BPRS positive symptoms subscale or if again not available the mean values at endpoint of the PANSS/BPRS positive symptoms subscale. The results of other rating scales will only be used if the instrument has been published in a peer-reviewed journal, because it has been shown that non-validated schizophrenia scales exaggerate differences.21

**Secondary outcomes**

Given the focus on treatments for positive symptoms, the results of this review will be informative for the treatment of positive symptoms. They will also describe how these interventions can have an effect on a number of other outcomes. With this aim, the following secondary outcomes will be assessed:
1. Acceptability, defined as the percentage of patients leaving the study early (dropout) for any reason. All-cause discontinuation due to any reason combines efficacy, tolerability and other factors, and can therefore be considered as a measure of ‘acceptability of treatment’ or of overall ‘effectiveness’.

2. Change in overall symptoms, measured by rating scales such as the PANSS or the BPRS, or any other published scale (e.g., the Manchester Scale) for the assessment of overall schizophrenic symptomatology. The results of other rating scales will only be used if the instrument has been published in a peer-reviewed journal.

3. Change in negative symptoms, measured by the respective subscale of the PANSS, or the ‘Scales for Assessment of Negative Symptoms’ or any other published scale.

4. Response, measured by the percentage of responders defined by reduction on the PANSS, BPRS or Clinical Global Impression (CGI) scores, accepting the criteria used by study authors.

5. Percentage of patients with relapse, by definitions operationalised by rating scales, and, if not available, number of rehospitalisations due to psychopathology. We will not include data from studies that used non-operationalised relapse criteria (e.g., clinical judgement).

6. Adherence, measured by any published rating scale (e.g., ‘Adherence Therapy Patients Satisfaction Questionnaire’ and ‘Adherence Rating Scale’).

7. Depression, measured by the Calgary Depression Scale for Schizophrenia, the Hamilton Depression Rating Scale, the Montgomery Asberg Depression Scale or other published symptom scales.

8. Quality of life, measured by any published rating scale (e.g., ‘Heinrichs quality of life scale’, Quality of Life Scale).

9. Functioning, measured by rating scales such as the Global Assessment of Functioning or the Psychosocial Performance Scale, or any other published rating scale.

10. Tolerability, measured as the percentage of patients experiencing adverse events. Adverse events associated with psychological treatments are not covered as comprehensively as in trials on pharmacological treatments. However, there is a raising awareness of the importance of considering possible harms associated with psychological interventions. Therefore we will collect any available information in clinical studies about this outcome, using a classification proposed by Linden and colleagues: (A) emergence of new symptoms; (B) deterioration of existing symptoms; (C) lack of improvement or deterioration of illness; (D) prolongation of treatment; (E) patient’s non-compliance; (F) strains in the patient–therapist relationship; (G) very good patient–therapist relationship, therapy dependency; (H) strains or changes in family relations; (I) strains or changes in work relations; (J) any change in the life circumstances of the patient; (K) stigmatisation. Suicide attempts and any other possible adverse event related to psychological treatment will also be considered.

11. Mortality. Psychosocial treatments may actually reduce or, by contrast, increase overall mortality, in particular connected to suicidality. To test this, we will examine this outcome in terms of (A) death for any reason, (B) death due to natural causes and (C) due to suicide.

**Search strategy**

**Electronic searches**

The following sources will be searched without restrictions for language or publication period: Embase, MEDLINE, PsycINFO and PubMed. The search terms that will be used for PubMed are provided as online supplementary material. We will also search the following international databases:

1. WHO International Clinical Trials Registry Platform,
2. BIOSIS,
3. Cochrane Collaboration Controlled Trials Register,

**Reference lists and other sources**

References of all selected studies will be inspected for other published reports and citations of unpublished studies. We will also inspect previous reviews conducted on psychological treatments for schizophrenia to check if some studies meet our inclusion criteria as well. In addition, we will contact the first author of each included study published in the last 30 years for missing information about their studies.

**Identification and selection of studies**

Studies identified through electronic and manual searches will be listed with citation, titles and abstracts, in Citavi; duplicates will be excluded. The eligibility for inclusion process will be conducted in two separate stages:

1. Two authors will independently inspect title and abstracts identified in the literature searches and exclude those not pertinent. Disagreement will be resolved by discussion and where doubt still remains, we will acquire the full article for further inspection and the article will proceed to the next stage.

2. Once the full articles are obtained, two reviewers will independently assess them for eligibility. Disagreements will be resolved by discussion and, if needed, a third senior author will be involved. When required, further information will be obtained from study authors.

**Data extraction**

Two authors will independently extract data from all selected trials. When disagreement arises we will resolve it by discussion and, if needed, involving a third senior author. Where this is not sufficient we will contact the study authors.

The following data will be collected from each included study:

Study citation, year(s) of study, registration number to trials registries, year of publication, location, setting, number of centres, sample size, diagnostic criteria and funding/sponsor (industry or academic).

Methodology, including study design (type of RCT), number of arms and risk of bias (see Risk of bias assessment section).

Characteristics of study participants, including gender, age, details on diagnosis, number randomised to each arm, sociodemographic characteristics, whether psychological treatments naive at baseline or with previous experience with the experimental intervention).

Characteristics of intervention, including number and frequency of sessions, therapy setting, expertise of therapist and researcher allegiance at study arm level.

Outcome measures, including information on whether an intention-to-treat approach has been used and how it was defined.

The two reviewers will independently input data into an Access database, especially created for this study. The software will automatically detect any inconsistencies and they will be resolved by discussion.

Measurement of treatment effect

Relative treatment effects

Continuous outcomes: for continuous outcomes we will use the standardised mean difference (SMD), because we expect that the studies use different rating scales of overall schizophrenia symptomatology.

Dichotomous outcomes: the effect size for dichotomous outcomes will be the risk ratio (RR) and its 95% CIs.

Relative treatment ranking

We will estimate the probability for each intervention to be ranked at each possible place, given the relative effect sizes as estimated in NMA. As described in Salanti et al.,25 we will obtain a hierarchy of the competing interventions using the surface under the cumulative ranking curve (SUCRA) and mean ranks. SUCRA values will be expressed as percentage, showing the relative probability of an intervention to be among the best options.

Dealing with missing outcome data and missing statistics

For continuous outcomes we will extract data for all randomised patients, if possible, and we will give preference to data based on mixed-effect models of repeated measurements of multiple imputations over last-observation-carried-forward data.

We will use published SDs, where available. When SEs instead of SDs are presented, the former will be converted to SDs.25 If both are missing, we will estimate SDs from P values or CIs, as described in Section 7.7.3 of the Cochrane Handbook for Systematic Reviews.26 If none of these options are available, we will contact the original authors. When no information can be obtained we will derive SDs from those of the other studies using a validated imputation technique.25

For dichotomous outcomes, everyone allocated to the intervention will be counted whether they completed the follow-up or not. If the authors applied such a strategy, we will use their results. If the original authors presented only the results of the per-protocol or completer population, we will assume that those participants lost to follow-up would not have changed in a given outcome. In terms of efficacy, this means that they would be conservatively considered to have not responded to treatment or control. In terms of tolerability, it would mean that participants would not have developed a side-effect.

Risk of bias assessment

Risk of bias will be assessed for each included study using the Cochrane Collaboration ‘risk of bias’ tool.26 27 The following domains will be considered:

1. Sequence generation: was the allocation sequence adequately generated?
2. Allocation concealment: was allocation adequately concealed?
3. Blinding of participants: was knowledge of the allocated treatment adequately prevented during the study?
4. Blinding of outcome assessors: were outcomes evaluated by blind raters? Were adequate measures taken to prevent them from discovering treatment allocation during the study?
5. Incomplete outcome data: were incomplete outcome data adequately addressed?
6. Selective reporting: are reports of the study free from suggestion of selective outcome reporting?
7. Researcher’s allegiance: do the researchers involved have a vested interest for the psychological treatment under investigation? We will additionally consider this point as possible source of bias, since it has been claimed to be relevant in trials on psychological interventions.28–30 An evaluation of high risk of bias will be given, for example, when the authors are founders of the therapy or have written a manual for that therapy.

A description of what was reported about the same domains in each study will be provided, and a judgement on the risk of bias will be made for each one of them, based on the following three categories: ‘high risk of bias’, ‘low risk of bias’ and ‘unclear risk of bias’ where information are not sufficient to make a judgement. Two independent review authors will assess the risk of bias in the selected studies. Any disagreement will be resolved through discussion. Where necessary, the authors of the studies will be contacted for further information. Studies will be classified as having low risk of bias if none of the domains above were rated as high risk of bias and three or less were rated as unclear risk; moderate if one was...
rated as high risk of bias or none was rated as high risk of bias, but four or more were rated as unclear risk and all other cases will be assumed to pertain to high risk of bias. We will not include studies in the data analyses whose sequence generation was at high risk of bias (e.g., randomisation by the date of birth or day of the week). Effects of high risk of bias in the other domains will be analysed by sensitivity analyses.

**Data analysis**

**Characteristics of the included studies**

We will produce descriptive statistics and study population characteristics across all eligible trials, describing the types of comparisons and other clinical or methodological variables, such as age, duration of illness, comedication, country, duration of study and number of sessions.

**Two-step procedure**

In a first step, we will perform series of conventional pairwise meta-analyses by combining studies that compared the same interventions, including the comparison between active treatments and the different control arms. In subgroups with very few RCTs available or if the requirements of NMA are not met, it can be that NMA will not be appropriate and, in this case, conventional pairwise meta-analysis will be the most straightforward approach. As heterogeneity is likely, a random-effects model will be used. In a second step, we will then perform a NMA within a frequentist framework.

**Assessment of heterogeneity**

The heterogeneity (variability in relative treatment effects within the same treatment comparison) will be measured with the tau-squared (the variance of the random-effects distribution). The heterogeneity variance will be assumed common across the various treatment comparisons (grouped by comparison type) and the empirical distributions will be used to characterise the amount of heterogeneity as low, moderate or high using the first and third quantiles. Potential reasons for heterogeneity will be explored by subgroup analysis (see below).

**Assessment of the transitivity assumption**

Joint analysis of treatments can be misleading if the network is substantially intransitive. We assume that patients who fulfill the inclusion criteria outlined in criteria for considering studies for this review section are equally likely to be randomised to any of the interventions that we plan to compare. We will need to investigate the distribution of clinical and methodological variables that can act as effect modifiers across treatment comparisons. We have maximised the chances of transitivity in our network with regard to clinical variables by limiting our samples to participants with schizophrenia and excluding specific subgroups like first-episode patients or patients with prevalent negative symptoms. Other clinical or methodological variables that may influence the efficacy of psychological interventions include administration mode and frequency of the treatment (like number of sessions and experience of the therapist), baseline severity (see ‘Investigation of heterogeneity and inconsistency’) and blinding, which will also be assessed in sensitivity analyses. We will investigate if these variables are similarly distributed across studies grouped by comparison. The comparability of studies comparing the intervention with treatment as usual or waiting-list conditions with those that provide head-to-head evidence will be examined carefully.

**Network meta-analysis**

NMA combines direct and indirect evidence for all relative treatment effects and can therefore provide estimates with maximum power and increased precision. If the collected studies appear to be sufficiently similar with respect to the distribution of effect modifiers (refer the Assessment of transitivity assumption section), we will conduct a random-effects NMA to synthesise all evidence for each outcome, and obtain a comprehensive ranking of all treatments. We will assume a single heterogeneity parameter for each network. We will present the summary SMDs or RRs for all pairwise comparisons in a league table. We will also estimate the prediction intervals to assess how much the common heterogeneity affects the relative effect with respect to the extra uncertainty anticipated in a future study. To rank the various treatments for each outcome, we will use the SUCRA and the mean ranks.

**Assessment of inconsistency**

The strategic and conceptual evaluation of transitivity will be supplemented with a statistical evaluation of consistency, the agreement between direct and indirect evidence. We will employ local as well as global methods to evaluate consistency. Local methods detect ‘hot spots’ of inconsistency, evidence loops that are inconsistent or comparisons for which direct and indirect evidence disagree. We will employ a method that separates direct evidence from indirect evidence provided by the entire network and then evaluate the agreement of these two pieces of evidence. We will also evaluate consistency in the entire network by calculating the design-by-treatment interaction test and I-squared for network heterogeneity, inconsistency, and for both. Tests for inconsistency are known to have low power, and empirical evidence has suggested that 10% of evidence loops published in the medical literature are expected to be inconsistent. Therefore, interpretation of the statistical inference about inconsistency will be carried out with caution and possible sources of inconsistency will be explored even in the absence of evidence for inconsistency.

**Investigation of heterogeneity and inconsistency**

We expect small amounts of heterogeneity and inconsistency to be present given the variety of study settings we plan to include. The following potential effect modifiers of the primary outcome will be explored by subgroup analyses:
A. Number of sessions,
B. Study duration,
C. Setting: individual versus group,
D. Expertise of the therapist,
E. Baseline severity (PANSS or BPRS score at baseline),
F. Different types of patients, with a different clinical outcome concerning symptoms (if identified).

Sensitivity analyses
We will explore the following sensitivity analyses by excluding:
A. Studies in which the outcome assessor was not blind (open studies);
B. Studies that presented only complete analyses;
C. Studies characterised as pertaining to high risk of bias;
D. Studies with high risk of bias in researchers’ allegiance;
E. Studies focused on treatment resistant patients (study defined);
F. Studies with a non-active comparison group.

Publication bias
We will first examine funnel plots of pairwise NMAs if there are 10 or more studies included. We will also explore the association between study size and effect size with a comparison-adjusted funnel plot that has been adapted to NMA.41

Evaluating the quality of the evidence
The quality of evidence contributing to each network estimate will be evaluated using an adaptation of the Grading of Recommendations Assessment, Development and Evaluation framework specifically developed for NMA.37 We will characterise the credibility of a body of evidence based on the study limitations, imprecision, heterogeneity/inconsistency, indirectness and publication bias.

Statistical software
The analysis and presentation of results will be performed using the Stata packages network and network_graphs, the R package netmeta.

Acknowledgements
The authors acknowledge the collaborators Samantha Roberts in helping us to conduct the literature searches, and Maximilian Huhn, Johannes Schneider-Thoma, Marc Krause and Costanza Carmi for their help and suggestions.

Collaborators
Samantha Roberts; Maximilian Huhn, Johannes Schneider-Thoma; Marc Krause, Costanza Carmi.

Contributors
IB and SL designed this study, drafted and critically revised the protocol. IB will screen search results for inclusion, conduct data extraction and data analysis and draft the final manuscript. SL will assist with data extraction and analysis and revise the final manuscript. CR and SW will screen search results for inclusion and conduct data extraction. GS provided substantial methodological advice in planning the study and will assist with data analysis. CB and TAF contributed with clinical and methodological input in planning the study. All authors contributed to and have approved the final manuscript.

Funding
This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no 701717. This work was also supported by the German Research Foundation (DFG) and the Technical University of Munich within the Open Access Publishing Funding Programme.

Disclaimer
The funder had no role in developing the protocol.

Competing interests
SL in the past 3 years has received honoraria for consulting from Roche, TEVA, Otsuka, Lundbeck and LB Pharma; for lectures from Otsuka, Lundbeck, Janssen, ICON, Lilly, Sanofi Aventis, ADP Orphan, Roche and Servier; and for a publication from Roche. TAF has received lecture fees from Eli Lilly, Janssen, Meiji, Mitsubishi-Tanabe, MSD and Pfizer and consultancy fees from Takeda Science Foundation. He has received royalties from Igaku-Shoin and Nihon Bunka Kagakusha publishers. He has received research support from Mochida and Mitsubishi-Tanabe.

Patient consent
Not required.

Provenance and peer review
Not commissioned; externally peer reviewed.

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES
15. Woodhead M. 80% of China’s clinical trial data are fraudulent, investigation finds. BMJ 2016;355:i5396.


