Acute interventions for aggression and agitation in psychosis: study protocol for a systematic review and network meta-analysis

Edoardo G Ostinelli,1,2 Armando D’Agostino,1 Farhad Shokraneh1, Georgia Salanti,3 Toshi A Furukawa4

ABSTRACT
Introduction Individuals with psychosis may access emergency services due to aggression and agitation. When the de-escalation technique fails to achieve tranquillisation, several pharmacological options are available. However, evidence on which intervention to prefer in terms of efficacy and tolerability to achieve resolution of the acute episode (ie, rapid tranquillisation) of aggression and agitation is currently fragmentary.

Methods and analysis We will include all randomised controlled trials comparing drugs or drug combinations or placebo for aggression or agitation episodes in adult individuals with psychosis. We will include individuals with psychosis (eg, schizophrenia and related disorders, bipolar disorder with psychotic symptoms, psychotic depression) but not substance or medication-induced psychosis or psychosis due to another medical condition. Our primary outcomes are the change in aggression or agitation scores within few hours since the administration of the intervention (efficacy outcome) and the proportion of participants who dropped out due to adverse effects (tolerability outcome). We will retrieve relevant studies from the register of studies of the Cochrane Schizophrenia Group. Also, we will run additional searches on CENTRAL, Embase and PubMed to retrieve potentially eligible studies focusing on other psychiatric diagnoses than those in the schizophrenia spectrum. We will conduct a random-effects network meta-analysis (NMA) for primary and secondary outcomes. In case of rare events of dichotomous outcomes, a common-effect Mantel-Haenszel NMA will be used instead. We will use the surface under the cumulative ranking curve and the mean ranks to rank all available treatments. Local and global methods of evaluation of inconsistency will be employed. Quality of evidence contributing to network estimates of the main outcomes will also be assessed with Confidence in Network Meta-Analysis.

Ethics and dissemination This study does not require ethical approval. We will disseminate our findings by publishing results in a peer-reviewed journal.

PROSPERO registration number CRD42019137945.

BACKGROUND Mental health-related behavioural emergencies are reported to account for around 6% of the visits to the emergency department in USA.1 In this context, aggression and agitation are clinically relevant issues in both psychiatric and emergency department settings.2 Verbal and physical aggression may result from psychomotor agitation, of which restlessness, excitability and dysphoria are prominent symptoms.3,4 Violent behaviour has been reported to occur in 18.5% of participants with psychosis included in a recently published systematic review.5 Consequences of this phenomenon may affect patients, their caregivers, healthcare personnel and community members, both in terms of harmful events towards self or others and stigma-driven victimisation of patients.6–9

Strengths and limitations of this study
- Our systematic review and random-effects network meta-analysis (NMA) will provide pragmatic answers to a critical question in the field of psychiatry: which drug or drugs combination, if any, is currently the best evidence-based choice in terms of efficacy and tolerability for rapid tranquillisation of aggression or agitation episodes in individuals with psychosis?
- The analysis will benefit from maximum statistical power by combining direct and indirect comparisons in an NMA, measuring the relative efficacy of these treatments.
- We will evaluate the contribution of drugs to their combination intervention by performing a component NMA as a sensitivity analysis.
- Limitations of individual studies will be judged with the revised Cochrane risk of bias tool (RoB2) and the credibility of the results for the main outcomes will be assessed using the Confidence in Network Meta-Analysis web application.
- Some studies might have evaluated more pragmatic outcomes without a continuous measurement of the level of aggression or agitation, thus not being able to contribute our primary efficacy outcome.
When facing acutely aggressive and agitated individuals, guidelines recommend clinicians to assess the feasibility of management with de-escalation techniques. Should this be deemed impractical or proven ineffective, a pharmacological approach is the mainstay for rapid tranquillisation. Several systematic reviews have evaluated the comparative efficacy and safety of pharmacological interventions from randomised controlled trials (RCTs) in schizophrenia spectrum illness. These reviews contributed currently available policy-making policies: for instance, NICE guidelines recommend the use of lorazepam or the combination of haloperidol and promethazine as pharmacological interventions for aggression and agitation in the context of schizophrenia spectrum illnesses.

However, evidence supporting these recommendations is still fragmentary and a comprehensive ranking of all treatments evaluated in RCTs is still lacking. Also, several interventions lack of head-to-head comparisons with recommended and available treatment options. For example, none of the identified RCTs compared the intramuscular administration of the second-generation antipsychotic aripiprazole with either lorazepam or the haloperidol and promethazine combination.

Network meta-analysis (NMA) is a statistical technique that provides estimates through direct and indirect comparisons and allows to assess the relative ranking of different treatments, ultimately resulting in increased precision. Objective of this systematic review and NMA is to compare available pharmacological interventions in terms of efficacy and tolerability in the acute treatment of aggression and agitation episode (ie, rapid tranquillisation) in adults with psychosis presenting to an acute care setting. In order to ensure our findings to be pragmatic and to provide clinically meaningful recommendations, we will focus on the transdiagnostic concept of psychosis when not induced by substance or medication use or other medical conditions.

**METHODS**

We submitted this study protocol to PROSPERO (CRD42019137945). Methods for this systematic review and NMA have been developed following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guideline for systematic review and meta-analysis protocols (PRISMA-P, online supplementary file 1), the extension statement for NMA (PRISMA-NMA, online supplementary file 2) and proposed additional considerations for protocols of systematic reviews including NMA.

**Criteria for considering studies for this review**

**Types of studies**

RCTs reported as double-blind will be included. Single-blind and open studies will be included but the effect of their inclusion will be evaluated in a sensitivity analysis. Quasi-randomised studies (eg, allocating by using alternate days of the week) will be excluded. Cross-over and cluster randomised trials will be included. Studies where sequence generation was at high risk of bias or where allocation was clearly not concealed will not be included.

**Types of participants**

Individuals aged 18 years or older, of both sexes, showing psychotic symptoms and an aggression or agitation episode deemed to be in need of pharmacological tranquilisation. We will also include studies in which baseline mean scores of an aggression or agitation rating scale suggest the need of pharmacological tranquilisation (eg, Positive and Negative Syndrome Scale Excited Component (PANSS-EC)≥14, Modified Overt Aggression Scale (MOAS)≥4). The impact of including these studies will be analysed (refer to the Exploring heterogeneity, inconsistency and sensitivity analyses section). Studies will be considered eligible only if psychotic symptoms are thought to be secondary to an established psychiatric diagnosis (eg, schizophrenia, schizoaffective disorder, delusional disorder, brief psychotic disorder, bipolar mania or depression with psychotic features, psychotic depression) in at least 50% of the randomised population. Psychotic disorders induced by substances or medications or due to another medical condition (eg, dementia, Parkinson’s disease, autoimmune disorders, neurodevelopmental disorders) will not be considered eligible. Peri-partum psychosis will not be considered eligible. A concurrent secondary diagnosis of another psychiatric disorder will not be considered as exclusion criteria. Treatment settings may be the outpatient emergency departments or the hospital wards.

**Types of interventions**

We aim to synthesise available data for first-generation antipsychotics, second-generation antipsychotics, benzodiazepines, antihistaminic drugs and valproic acid or their combinations. We will obtain information about the interventions of interest from head-to-head and placebo-controlled trials (ie, pill or pharmacological placebo). Inclusion of placebo-controlled trials may contribute via indirect evidence to the treatment comparisons of interest by increasing the precision and allowing additional connections in the network. Pharmacological interventions not described above will be considered for post-hoc inclusion in the network within the context of the joint randomisability assumption by evaluation of their effect as rapid tranquiliser. Interventions (drugs or drug combinations) may be administered via any route of administration consistent with the rapid tranquilisation aim (ie, long-acting injectable antipsychotics are not eligible since they aim at maintenance treatment). Any dosing strategy (ie, flexible or fixed) will be allowed. We will allow multiple administrations and we will consider them when evaluating the eligibility of dosages and fairness of comparisons for the primary efficacy outcome. Multiple administrations, when not delivered following a
fixed schedule, will contribute to the ‘individuals needing additional interventions’ outcome. We will include only study arms in which individuals have been randomised to drugs within the recommended dose range for aggressive or agitated individuals with psychosis. Should these not be available in the current literature, we will consider licensed dose ranges for psychiatric or, if absent, medical conditions. If a study included arms with both approved and unapproved doses, we will include only those arms that employed eligible doses. Should a comparison include different dosages of the same drug, we will consider them as a single node as long as the dosages are within the eligible range. Therefore, studies comparing the same drug at different dosages without another intervention will not contribute the network. For studies with a fixed dosing strategy, we will consider the planned dosage as the administered one. If a study employs a flexible dosing strategy, we will abide by the following hierarchy: (1) mean dose within 48 hours since the first administration; (2) mean dose until the end of the study; (3) average between the minimum and maximum doses allowed. Fairness of dose comparisons in head-to-head trials will be evaluated in a sensitivity analysis. Drug monotherapies and drug combinations within eligible dose ranges will be treated initially as different nodes in the network.

Outcome measures

Our primary outcomes are:
1. Efficacy (as continuous outcome)—aggression or agitation scores as close to 2 hours as possible since administration of the intervention (range: >0, ≤6 hours). Endpoint scores will be preferred to change scores, if both are reported. We will employ validated rating scales or subscales commonly used to evaluate aggression or agitation including but not limited to Behavioral Activity Rating Scale, Overt Agitation Severity Scale, MOAS or PANSS-EC. If more than one scale has been administered, we will prioritise scales measuring aggression over agitation and scales over subscales. It is likely that across included trials different rating scales have been used to measure the same outcome. To overcome this, we plan to harmonise continuous outcomes to a common metric (ie, standardised mean difference (SMD)). We acknowledge that during a trial additional drugs may be administered as rescue drug to give relief from adverse events. Should this occur, we will use the closest time point to 2 hours after baseline at which additional drugs have not yet been administered. Should this be unavailable, we will include data only if administration of rescue drugs was not unbalanced between the groups (eg, mean value of daily administered milligrams) and we will evaluate the impact of their inclusion in a sensitivity analysis (refer to the Exploring heterogeneity, inconsistency and sensitivity analyses section). If both aggression and agitation data are collected within the same trial at the preferred time points, we will choose the aggression one since it is more representative of clinical urgency. If only agitation outcomes have been measured or reported, we will use this as a proxy of aggression outcome.
2. Tolerability (as dichotomous outcome)—dropouts due to adverse events as close as to 24 hours since the administration of the intervention (range: >0, ≤7 days).

Our secondary outcomes include:
1. Efficacy (as continuous outcome)—aggression or agitation scores as close to 24 hours as possible since the administration of the intervention (range: >6, ≤48 hours).
2. Efficacy (as dichotomous outcome)—proportion of individuals needing one or more additional interventions (ie, additional administration of any drugs to achieve tranquillisation, needing restraints or seclusion) up to 6 hours since the administration of the intervention (range: >0, ≤6 hours). When the number of individuals in need of additional treatment is reported specifically for every adjunctive care option, summing these data may lead to double counting issues (ie, individuals who needed two or more different additional interventions). Hence, we will abide by the following order: (1) additional administration of any drugs, (2) needing restraints and (3) needing seclusion.
3. Efficacy (as dichotomous outcome)—proportion of individuals needing an additional intervention (ie, additional administration of any drugs to achieve tranquillisation, needing restraints or seclusion) up to 48 hours since the administration of the intervention (range: >0, ≤48 hours). Outcomes will be prioritised as stated above.
4. Tolerability (as dichotomous outcome)—proportion of individuals with at least one serious side effect as close to 24 hours as possible since the administration of the intervention (range: >0, ≤7 days). We define ‘serious’ side effects as stated by the original investigators. For the efficacy outcomes, should two available measurements be equidistant from the predefined time point we will prioritise earlier time points within the above stated range. For the tolerability outcomes, should the same conditions apply we will prefer later estimates over the earlier if both are eligible.

Search strategy

We will search the Cochrane Schizophrenia Group register for studies relevant to aggression and agitation in schizophrenia spectrum illnesses. In this 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. This allows rapid and accurate searches that reduce waste in the next steps of systematic reviewing. This register is compiled by systematic searches of major resources (AMED, 1985–present; BIOSIS, 1969–present; CENTRAL, not applicable; CINAHL, 1982–present; ClinicalTrials. Gov, 2000–present; Embase, 1974–present; ISRCTN, 2000–present; MEDLINE, 1946–present; PsycINFO,
Participants: number, proportion of first-episode

Methods: date of the study, study design, randomi-

There is no language, date, document type, or
publication status limitations for inclusion of records
into the register. An additional search on CENTRAL (not
applicable), Embase (1974-present) and PubMed (1946-
present) databases will be performed with keywords rele-
vant to psychotic depression, episodes of bipolar mania
or depression with psychosis, delusional disorder and
psychotic symptoms, combined with condition-specific
(e.g., aggression, agitation) and study type-specific (e.g.,
random, RCT) terms. We plan no restrictions in terms of
language, country or publication period for this search.
We will also inspect similar systematic or non-systematic
reviews as additional sources of potentially eligible trials.
Search queries, dates and number of identified records
will be published in the supplementary material and
depicted in a PRISMA flowchart.

Studies selection
Two members of the team will independently review any
identified record. In case of disagreement, the record will
be included to assess its eligibility at the full-text phase.
Two members of the team will independently review the
full text of retrieved potentially eligible studies. We will
develop a checklist of eligibility criteria to ease the system-
atic assessment of the retrieved full texts. We will resolve
any disagreement by discussion with a third member of
the team. Identified records referring to the same trial
(i.e., secondary publication) will be collated. We will clas-
sify trials according to the Cochrane Handbook defini-
tions in four mutually exclusive categories (i.e., included,
excluded, awaiting assessment, ongoing). We will list the
reasons for exclusion and any other relevant comment in
the supplementary material. We will contact original
investigators of the trials to retrieve any additional infor-
mation deemed useful. In case of no response, we will send
weekly reminders up to 3 weeks (i.e., three reminders).
We will report in the supplementary material any attempt
to contact original investigators and the extent of their
contribution (e.g., information and/or data provided, no
answer).

Data extraction
We will design and use a structured data extraction form
to ensure consistency of extracted information and data.
We will extract the following data:

- Methods: date of the study, study design, randomi-
sation (i.e., individual or cluster), number of study
  centres.
- Participants: number, proportion of first-episode
  patients, gender, age, underlying treatments, inclu-
sion and exclusion criteria, co-morbid disorders.
- Interventions: dose, dose range, dosing strategy,
frequency of administration, route of administration,
use of rescue medication.
- Outcomes: definition, measures, timing of assessment.
- Others: sponsorship or funding for trial (e.g., industry,
  academic), notable conflict of interests of trial orig-
inal investigators.

Data estimates retrievable by graphs will be extracted
(e.g., mean, SD) using Plot Digitizer, an electronic ruler.42
If both the SD and the SE are missing but p values or CIs
are available, we will calculate SDs following the Cochrane
Handbook guidelines.43 Should the original publication
provide continuous data relevant to an outcome hereby
specified as dichotomous, we will impute response rates
from means and SD by employing a validated imputation
method.44

Missing outcome data
To overcome the issue of missing outcome data, we aim at
replicating the methods previously employed in a recent
NMA on depression.45 For continuous outcomes, when
data for a specific outcome is available for both inten-
tion-to-treat and per-protocol populations, preference
will be given to the former. We will prefer data based on
mixed-effect models of repeated measurements, multiple
imputation or any other appropriate imputation methods
over last-observation-carried-forward data.46 We will
exclude participants with missing outcome data from the
analysis. For the dichotomous efficacy outcome, we will
assume that participants with missing outcome data failed
to achieve an adequate response. Dropouts from assess-
ment will be assumed not to have changed from baseline
for the given outcome: not in need of additional interven-
tions in terms of efficacy outcomes, and not having devel-
oped an adverse effect in terms of tolerability outcomes.
We will evaluate the appropriateness of the imputation
method in the Risk of bias judgement section.

Unit of analysis issues
We will extract data from cross-over studies only for the
first phase (i.e., before cross-over) to avoid including data
potentially affected by the carry-over effect. However,
given that we will extract data from relatively short trials,
we expect cross-over effect to be unlikely. In trials that
employ ‘cluster randomisation’ (e.g., by practice or clini-
cian), we will extract data adjusted for the clustering
effect. If these are not available, we will extract unad-
justed data and divide the sample size (for continuous
outcomes) or both sample size and number of events (for
dichotomous outcomes) by the design effect, defined as
follows47:

\[
\text{Design effect} = 1 + \left( m - 1 \right) \text{ICC}
\]

where \( m \) denotes the mean number of participants per
cluster and ICC is the intra-cluster correlation coefficient,
an estimate of the relative variability within and between
clusters.48 If the ICC is not reported, we will assume it to
be 0.1.49
Comparability of dosages

Some trials may have employed unfair dose comparisons, possibly biasing the resulting estimate. To overcome this, we will add a dichotomous variable to indicate whether the comparison was fair by narrowing the eligible range of doses (ie, excluding lower and upper 20% of doses of the recommended or licensed ranges). Should a combination include two or more drugs at different levels of the range of doses, we will rate the combination according to the drug at the higher one (ie, hierarchically as upper range, mid-range, lower range).

Risk of bias judgement

To evaluate the risk of bias, we will use the revised Cochrane risk of bias tool for randomised trials (RoB 2) (V.9, July 2019).50 Two independent raters will perform the evaluation following the full guidance document and using the provided Excel tool. In case of disagreement, consensus will be reached either by discussion or by consultation of a third member of the team. Studies contributing to least one primary outcome will be assessed in the following domains: bias arising from the randomisation process, bias arising from identification or recruitment of individual participants within clusters (only for cluster-randomised trials), bias due to deviations from intended interventions (effect of assignment to the interventions at baseline), bias due to missing outcome data, bias in measurement of the outcome, bias in selection of the reported result.

Should a study contribute to both primary outcomes, we will provide two outcome-specific separate risk of bias judgements.

Additional domains

Additionally, we will evaluate the influence of sponsorship and notable conflict of interest, defined as follow:

- Sponsorship (arm-level): (1) funded by a pharmaceutical company with an active interest or (2) one or more authors were employees of such company.
- Notable conflict of interest (arm-level): (1) funded by a pharmaceutical company with an active interest, (2) one or more authors were employees of a company with an active interest, (3) having received speaker fees or honoraria from a company with an active interest or (4) having disclosed any other notable conflict of interest.

These two domains will not contribute towards the overall risk of bias judgement but will be separately addressed (refer to the Exploring heterogeneity, inconsistency and sensitivity analyses section).

Overall risk of bias judgement

Studies will be classified according to the guidance manual for RoB 250:

- Overall ‘low risk of bias’ if all the considered domains are rated to be at low risk of bias.
- Overall ‘some concerns’ if (1) at least one of the considered domains is rated to raise concern and (2) none of the considered domains is rated to raise high risk of bias.
- Overall ‘high risk of bias’ if (1) at least one of the considered domains is rated to be at high risk of bias (irrespective of which domain is being assessed) or (2) the study is judged to raise concerns in multiple domains.

Statistical synthesis of study data

Characteristics of included studies and information flow in the network

We will provide descriptive statistics for the characteristics of trial and study population across eligible trials and also visualise relevant information by producing a network diagram. To understand which comparison contributes the most to the network estimate and how it affects the estimate, the percentage contribution of each direct meta-analysis will be described in a contribution matrix table.51

Pairwise meta-analyses

We will synthesise data of each pairwise comparison to obtain SMDs for continuous outcomes and ORs for dichotomous outcomes, and their 95% CIs. Assuming that different studies are measuring different but related treatment effects, we will use a random-effects model. We will compare the estimated distribution of the heterogeneity variance τ² to predictive distributions, as described elsewhere.52 53

Assessment of the transitivity assumption

The transitivity assumption is the basic premise of a NMA. Distribution of clinical and methodological variables likely to act as effect modifiers will be investigated.54 55 We assume that individuals who fulfil the inclusion criteria are, in principal, equally likely to be randomised to any of the interventions of interest (ie, jointly randomisable). When additional evidence of intransitivity is lacking and potential effect modifiers have similar distribution across the included studies, NMA is likely to give valid results. We acknowledge that aggression and agitation might be conceived as different behavioural aspects. However, we will assume that individuals with psychotic symptoms and high levels of either aggression or agitation will have a similar clinical presentation and will receive similar interventions. Aggression or agitation due to psychotic symptoms induced by substance or medication use or due to a medical condition may moderate efficacy and tolerability of rapid tranquillisation interventions. For instance, antipsychotics are usually preferred over benzodiazepines in individuals with alcohol intoxication because of the risk of respiratory depression, while for other substances interventions other than antipsychotics are often prioritised to avoid increasing cardiotoxicity risk.56 Despite this, clinical presentation of these conditions in the acute setting may be difficult to distinguish. To overcome this, we restricted eligible studies to those focusing (ie, ≥50% of total recruited sample) on individuals with psychosis...
not induced by external causes or medical conditions. We do not expect these last to be an effect modifier of efficacy or tolerability of rapid tranquilisation interventions. Therefore, inclusion of diagnoses other than schizophrenia should not threat the plausibility of our transitivity assumption.

Other clinical or methodological variables that may influence our primary outcomes are year of study, age, baseline severity of aggression or agitation, baseline severity of mental state, comorbid substance or alcohol abuse (secondary diagnosis). We will investigate the distribution of these variables across studies grouped by comparison.

Network meta-analyses
If undertaking NMA is deemed appropriate as described above, we will perform a joint synthesis of all trials per outcome. Network diagrams for all the included studies and for the subsets contributing each outcome will be produced. For dichotomous outcomes of rare events, we will perform a common-effect Mantel-Haenszel NMA. For all other outcomes, we will employ a random-effects NMA assuming a single heterogeneity parameter for all comparisons in each network. We will present summary estimates (ie, SMD, OR) for all pairwise comparisons in a league table. We will also provide the prediction intervals to assess how much the common heterogeneity affects the relative effect with respect to the additional uncertainty anticipated in a future study. We will calculate the probability for each drug to be ranked at each possible place and will present them in rankograms. To assess how likely it is for each drug to ‘beat’ most of its competitors in each outcome, we will calculate the SUITRA and the mean rank for each drug per outcome.

Assessment of inconsistency
The statistical manifestation of transitivity is termed consistency, which relies on agreement between direct and indirect estimates. Since consistency can be assessed only in the presence of both direct and indirect evidence (ie, closed loop), when these are absent—by definition—there cannot be inconsistency. Absence of statistical inconsistency never ensures the validity of transitivity, which itself remains an assumption. When closed loops are present in the network, inconsistency may result from variation of effect modifiers across treatment comparisons. To evaluate inconsistency, we will employ both local and global methods. We will employ the ‘separating indirect from direct evidence approach’, which separates direct evidence from indirect evidence to evaluate local consistency, and the design-by-treatment interaction test to evaluate consistency in the entire network.

Exploring heterogeneity, inconsistency and sensitivity analyses
We expect small amounts of heterogeneity and inconsistency to be present. We will perform the following subgroup analyses for our two primary outcomes: 1. Sponsorship/notable conflict of interest.

2. Arms with a proportion equal to or greater than 50% of participants with comorbid substance-use disorders.
3. Route of administration. In case of combined treatment with different routes of administration, we will classify it according to the drug at the higher level of range of doses (ie, hierarchically as upper range, mid-range, lower range). Should it be the same, we will classify it as ‘intramuscular’.

We will also perform the following sensitivity analyses on the primary outcomes:
1. Blinding (ie, double-blind vs non double-blind).
2. Only studies with balanced dosages, by excluding arms with eligible but unfair dose comparison.
3. Only studies with an overall low risk of bias.
4. Contribution of combining drugs. If feasible, we will perform a component NMA assuming an additive contribution of single drugs to drug combinations (ie, drug A and drug B to estimate the combination AB), and no interaction between the two. In this sensitivity analysis, we will group together all the benzodiazepine interventions both for single drugs and drug combinations.
5. Administration of rescue drugs to evaluate the impact of including in the network studies in which rescue drugs were allowed but whose administration was not unbalanced.
6. Studies with specific inclusion criteria for agitation or aggression, by excluding studies with baseline mean levels suggesting the need for an intervention.

Model implementation
We will fit all models in R. We will use the meta package to perform pairwise meta-analyses. All the random-effects NMA will be fit using the netmeta package assuming a common $\tau^2$.

Assessment of the confidence in the evidence from NMA
To assess the confidence we can place in the relative treatment effects estimated in NMA, we will use the Confidence in Network Meta-Analysis (CINeMA) framework, implemented in the web application https://cinema.ispm.unibe.ch. CINeMA evaluates the credibility of the findings across the following domains: within-study bias, across-studies bias, indirectness, imprecision, heterogeneity and incoherence. Imprecision, heterogeneity and incoherence domains require to specify a clinically important size of effect on the scale of the considered effect measure. We will set the following margin of non-equivalence: (1) active versus active comparison, a SMD of 0.1 for continuous outcome (severity of aggression/agitation) and an OR of 2.0 for dichotomous outcome (dropouts due to adverse events). We will apply these criteria to both active versus active and active versus...
placebo comparisons, since we are mainly interested in relative effect of active treatments.

**Patient and public involvement**

No patient was involved in the protocol development nor will be involved in the review.

**ETHICS AND DISSEMINATION**

This study does not require ethical approval. We will disseminate our findings by publishing results in a peer-reviewed journal.

**Author affiliations**

1 Department of Health Sciences, Università degli Studi di Milano, Milan, Italy
2 Cochrane Schizophrenia Group, Division of Psychiatry and Applied Psychology, Institute of Mental Health, School of Medicine, University of Nottingham, Nottingham, UK
3 Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland
4 Department of Health Promotion and Human Behavior and Department of Clinical Epidemiology, Graduate School of Medicine and School of Public Health, Kyoto University, Kyoto, Japan

**Twitter** Farhad Shokraneh @Farhad_Cochrane and Toshi A Furukawa @Toshi_FRKW

**Contributors**

EGO, TAF and GS conceived the study, EGO, TAF, GS, FS and ADA provided substantial contribution to the design of the study during its development. EGO, TAF, GS, FS and ADA revised the study protocol critically for important intellectual content and contributed to its improvement. FS have planned the search in Cochrane Schizophrenia Group (CSCG) register and will conduct it. EGO have planned the search outside the CSCG register and will conduct it. GS supervised the statistical analysis planning. EGO, TAF, GS, FS and ADA gave the final approval to publish the current protocol. All the authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Funding**

The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests**

GS was invited to participate in two scientific meetings on the use of real-world evidence organised by Merck and Biogen. TAF reports personal fees from Meiji, Mitsubishi-Tanabe, MSD and Pfizer and a grant from Mitsubishi-Tanabe, outside the submitted work; TAF has a patent 2018-1 77 688 pending.

**Patent consent for publication**

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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**

Eduardo G Ostinelli http://orcid.org/0000-0002-6717-0832
Armando D’Agostino http://orcid.org/0000-0002-2126-799X
Farhad Shokraneh http://orcid.org/0000-0001-9687-8560

**REFERENCES**