Comparative efficacy and acceptability of psychosocial interventions for stimulant use disorders: a systematic review and network meta-analysis

Franco De Crescenzo, Marco Ciabattini, Gian Loreto D’Alò, Riccardo De Giorgi, Cinzia Del Giovane, Carolina Cassar, Luigi Janiri, Nicolas Clark, Michael Joshua Ostacher, Andrea Cipriani

PROTOCOL
**Introduction**

Stimulant use disorder is a type of substance use disorder defined by the continued use of cocaine, amphetamine-type substances (ATS including amphetamines, methamphetamines, and ‘ecstasy’), or other stimulants leading to clinically significant impairment or distress (APA 2013).

Globally, cocaine and ATS are the most commonly abused stimulants with an annual prevalence of 0.38% and 1.20% respectively in those aged 15-64 years (UNODC 2016). Patients dependent on stimulants suffer from a range of psychological and physical sequelae including psychosis and other mental illnesses, neurological disorders, cardiovascular dysfunctions, sexually-transmitted diseases, blood-borne viral infections such as HIV, hepatitis B, hepatitis C, and ultimately increased mortality (Degenhardt 2012). Moreover, crime, sexual abuse, and interpersonal violence contribute to increase the social burden of disease (Atkinson 2014).

An increasing number of people is seeking treatment for stimulants use disorders (UNODC 2016), yet currently there is no widely accepted treatment. Evidence supporting the use of pharmacological therapy is sparse (Castells 2010; Pani 2010; Ciccarone 2011; Dürsteler 2015; Minozzi 2015; Nuijten 2016), whereas several clinical guidelines endorse psychosocial interventions as first-line treatment (APA 2006; NICE 2012).

A previous pairwise meta-analysis was done to assess the efficacy of all types of psychosocial interventions (Minozzi 2016). However, this study did not compare directly between different interventions and therefore was unable to generate hierarchies among the available treatments. Hence, we will undertake a network meta-analysis to compare and rank the efficacy and acceptability of psychosocial interventions for the treatment of stimulant use disorders.

**Objective**

The objective of this systematic review and network meta-analysis is to compare the efficacy and acceptability of different psychosocial interventions for the treatment of stimulant use disorder.

**Methods**

**Types of studies**
**Inclusion criteria**

We will include only randomised controlled trials (RCTs) comparing any specific structured psychosocial interventions against a control intervention (another structured psychosocial intervention, or treatment as usual, or non-contingent rewards) for the treatment of stimulant use disorder. We will include trials with a study duration \( \geq \) one month.

**Exclusion criteria**

We will exclude studies comparing different intensities of the same psychosocial interventions, without a comparison group (e.g. trials comparing one weekly session of CBT versus two weekly sessions of CBT). Quasi-randomised controlled trials, in which treatment assignment is decided through methods such as alternate days of the week, will also be excluded.

**Types of participants**

**Inclusion criteria**

Participants aged 18 or older, with a diagnosis of stimulant abuse/dependence according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) III, III-R, IV, IV-TR or V or the International Classification of Diseases (ICD) 10 criteria, irrespective of pattern of use, gender, age or nationality. Stimulants are considered to be substances that activate, enhance or increase neural activity. Stimulants include cocaine, amphetamines, dextroamphetamine, diethylpropion, ephedrine, metamphetamine, MDMA, methylphenidate, pemoline, phenmetrazine, phendimetrazine, phenylpropanolamine, and phenylpropanolamine (UNODC 2016; Kiluk 2016). We will include trials on patients with additional diagnosis of abuse/dependence on other substances, such as alcohol or cannabis, or people with comorbid psychiatric disorder.

**Exclusion criteria**

We will exclude studies recruiting children or adolescents (i.e. less than 18 years old). We will not use studies using ICD-9 as it does not include operationalised criteria, having only disease names and no diagnostic criteria.
**Types of interventions**
We will include the following structured psychosocial interventions: cognitive behavioural therapies (CBT), community reinforcement approach (CRA), contingency management (CM), interpersonal psychotherapy (IPT), mindfulness based interventions (MBI), supportive-expressive psychodynamic therapy (SE), family therapy (FT), the twelve-step program (TS) or a combination of them. We will exclude in our study brief interventions in individuals with stimulant misuse not actively searching for treatment (such as motivational interviewing).
We assume that any patient that meets all inclusion criteria is likely, in principle, to be randomised to any of the interventions.

**Comparator**
Any of the psychosocial interventions listed above will be eligible as comparators; treatment as usual (TAU) will be considered a control. Non-contingent rewards will be pooled as a separate control condition.

**Grouping of interventions**
We will merge the interventions "a priori", before carrying out the statistical analyses, through a consensus process within the review group. Studies comparing different modalities of the same type of psychotherapy (e.g. face-to-face or telephone), different treatment conditions (e.g. CBT or CBT plus sessions for carers), or different intervention formats (e.g. group or individual), will be considered as the same node in the network meta-analysis.

**Context**
Studies conducted in both inpatient and outpatient settings will be included.

**Types of outcome measures**

**Primary Outcomes**
1. Efficacy will be measured as the number of individuals abstinent. We will prioritize outcomes assessed by urinalysis; if missing, we will use self-reported outcomes (Kiluk 2016).
We will prioritise as primary efficacy outcome the continuous abstinence at 12 weeks (which is the number of participants achieving a sustained period of time of abstinence after 12 weeks of treatment), assessed by urinalysis or self-report;

- If the continuous abstinence at 12 weeks is not reported, we will extract data on the continuous abstinence during treatment (which is the number of participants achieving a sustained period of time of continuous abstinence during treatment), assessed by urinalysis or self-report;

- If the continuous abstinence is not reported at all, we will extract the point abstinence at 12 weeks, (which is the number of participants abstinent after 12 weeks of treatment), assessed by urinalysis or self-report.

2. Acceptability will be measured as the number of patients who dropped out from the study due to any cause after 12 weeks of treatment.

**Secondary Outcomes**

3. As secondary outcomes we will assess the number of individuals abstinent and dropouts as described above, for the long-term at 26 weeks.

4. Longest duration of abstinence, that is the average longest period of time (measured in weeks) without drug assumption achieved by study participants, assessed by urinalysis or self-report; We will prioritize outcomes assessed by urinalysis; if missing, we will use self-reported outcomes.

5. Abstinence at follow-up, measured as the number of participants abstinent at the longest follow up after study completion, assessed by urinalysis or self-report. We will prioritize outcomes assessed by urinalysis; if missing, we will use self-reported outcomes.

**Timing of Outcome Assessment**

We will assess as primary outcomes efficacy (number of individuals abstinent) and acceptability (dropouts) for the short-term at 12 weeks. If 12-week data are not available, we will use data ranging between 4 and 20 weeks (we will give preference to the time-point closest to 12 weeks).

Regarding secondary outcomes, we will assess efficacy (number of individuals abstinent) and acceptability (dropouts) for the long-term at 26 weeks. If 26-week data are not available,
we will use data ranging between 21 and 40 weeks (we will give preference to the time-point closest to 26 weeks).
The longest duration of abstinence as well will be measured for the short-term at 12 weeks (if 12-week data are not available, we will use data ranging between 4 and 20 weeks, giving preference to the time-point closest to 12 weeks) and the long-term at 26 weeks (if 26-week data are not available, we will use data ranging between 21 and 40 weeks, giving preference to the time-point closest to 26 weeks).
The abstinence at follow-up will be measured at the last follow-up visit after study completion. If participants received further treatments after the initial trial (e.g., continuous treatment or booster sessions), they will not be included in the follow-up analysis.

**Search methods for identification of studies**

We will carry out a comprehensive computer literature search of Cochrane Drugs and Alcohol Group Register of Trials, Medline, Embase, CINAHL, ISI Web of Science, PsycINFO using a validated search strategy (Minozzi et al. 2016). We will search the databases using MeSH and free-text terms relating to psychological interventions and stimulant abuse/dependence as shown in Appendix 1. We will combine the PubMed search with the Cochrane Highly Sensitive Search Strategy for identifying RCTs in MEDLINE: sensitivity-maximising version (2008 revision; Lefebvre 2011). We will also use the following sources: reference list of retrieved articles, conference proceedings. These will include the Society for the Study of Addiction, International Harm Reduction Association and American Association for the Treatment of Opioid Dependence. When needed, we will also establish a contact with investigators and relevant trial authors, in order to obtain information about unpublished or incomplete trials. All searches will also include non-English language literature.

**Data collection and analysis**

**Selection of studies**

Two authors (FDC, MC) will independently screen the abstracts of all publications that are obtained by the search strategy. Two authors (FDC, MC) will independently assess the full text of potentially relevant studies for inclusion. All the authors will discuss any conflict regarding eligibility of a study until an agreement is obtained. Three authors (FDC, GLD, MC) will extract data regarding: number and characteristics of participants, setting, type of
experimental and control intervention, length of follow-up, efficacy and acceptability outcomes as previously reported. A data extraction form will be designed to ensure consistency of collection and appraisal for each study. Conflicts of opinion regarding eligibility of a study will be discussed with a fourth review author (AC), having retrieved the full paper and consulted the original study authors if necessary, until consensus is reached. Methodological experts will be consulted if necessary.

**Assessment of risk of bias in included studies**

Two review authors [FDC, MC] will independently assess risk of bias for each study using the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins and Green 2011). We will resolve any disagreements by discussion or by involving another author [GLD]. The following domains will be assessed: random sequence generation, allocation concealment, blinding of participants and outcome assessment, incomplete outcome data, and selective outcome reporting. We will judge each potential source of bias as high, low or unclear and provide a supporting quotation from the study report together with a justification for our judgment in the 'Risk of bias' table. We will report the risk of bias judgements across different studies for each of the domains listed. Where information on risk of bias relates to unpublished data or correspondence with a trialist, we will note this in the 'Risk of bias' table. A judgement of high risk of bias in one or more domain will be considered as ‘high risk’ study, a judgement of low risk of bias in 5 or more domains will be considered as ‘low risk’ study, while all remaining situations will be considered as ‘moderate risk’ study. When considering treatment effects, we will take into account the risk of bias for the studies that contribute to that outcome.

**Dealing with missing data**

Study authors will be contacted when there are missing or unclear data. If missing dichotomous outcome data will be still present, they will be managed according to the intention to treat (ITT) principle, and we will assume that patients who dropped out after randomisation had a negative outcome. Missing continuous outcome data will be analysed on an endpoint basis, including only participants with a final assessment. When p-values, t-values, confidence intervals or standard errors are reported in articles, SDs will be calculated from their values (Furukawa 2006).
Assessment of clinical and methodological heterogeneity within treatment comparisons
The studies synthesized in each pairwise comparison should be similar enough in terms of patient characteristics, setting, outcome definitions etc. in order to obtain interpretable and useful results (Higgins and Green 2011). To evaluate the presence of clinical and methodological heterogeneity we will generate descriptive statistics for trial and study population characteristics across all eligible trials. We will assess the presence of clinical heterogeneity within each pairwise comparison by comparing these characteristics (Higgins and Green 2011).

Assessment of transitivity across treatment comparisons
The assumption of transitivity (i.e. one can compare indirectly treatments B and C via treatment A) underlies network meta-analysis and needs careful evaluation. In case transitivity is not plausible in a network of trials, the indirect and mixed treatment effect estimates are not valid. In order to infer about the assumption of transitivity (Salanti 2012), we will assess whether the included interventions are similar when they are evaluated in RCTs with different designs; for example, whether interventions are administered the same way in studies comparing active treatments to usual care and in those comparing active treatments to other active treatments.

Strategy for data synthesis

Methods for direct treatment comparisons
We will compare every treatment intervention carrying out pairwise meta-analyses using a random effects model by using the statistical software STATA.

Methods for indirect and mixed comparisons
We will also compare all the outcomes of included interventions directly and indirectly through the execution of a network meta-analysis (NMA), in order to obtain a compendium of the different treatments impact on the treated condition. We will perform NMA using a random-effects model within a frequentist setting assuming equal heterogeneity across all comparisons and we will account for correlations induced by multi-arm studies (Lu 2006; Salanti 2009). The models will enable us to estimate the probability for each intervention to
be the best for each outcome, given the relative effect sizes as estimated in NMA. We will perform NMA in STATA using the 'mvmeta' command and self programmed STATA routines available at http://www.mtm.uoi.gr (Chaimani 2014; White 2011; White 2012). We will present jointly the relative ranking of treatments for efficacy and acceptability. Results from NMA will be presented as summary relative effect sizes. For continuous outcomes we will use standardized mean difference (SMD) Hedges’ adjusted g, while for dichotomous outcomes, we will use the odds ratio (OR) for each possible pair of treatments. We will also estimate the ranking probabilities for all interventions of being at each possible rank. Then, we will obtain a treatment hierarchy using the surface under the cumulative ranking curve (SUCRA) and mean ranks. SUCRA can also be interpreted as the percentage of efficacy/safety of a treatment that would be ranked first without uncertainty.

Assessment of statistical heterogeneity

We will statistically assess the presence of heterogeneity for all direct pair wise comparisons using the $\tau^2$. The assessment of statistical heterogeneity in the entire network will be based on the magnitude of the heterogeneity variance parameter ($\tau^2$) estimated from the NMA models. We will compare the magnitude of the heterogeneity variance with the empirical distribution as derived by Turner (Turner 2012). We will also estimate a total $I^2$ value for heterogeneity in the network as described elsewhere (Jackson 2014).

Assessment of statistical inconsistency

Consistency in a network of treatments refers to the agreement between direct and indirect evidence on the same comparisons. Joint analysis can be misleading if the network is substantially inconsistent. Inconsistency can be present if the trials in the network have very different protocols and their inclusion/exclusion criteria are not comparable or may result as an uneven distribution of the effect modifiers across groups of trials that compare different treatments. We will first check for any erroneous data abstraction. Then, to evaluate the presence of inconsistency locally, we will use the loop-specific approach. This method evaluates the consistency assumption in each closed loop of the network separately as the difference between direct and indirect estimates for a specific comparison in the loop (inconsistency factor) (Veroniki 2013). The magnitude of the inconsistency factors and their 95% CIs can then be used to infer about the presence of inconsistency in each loop. We will assume a common heterogeneity estimate within each loop.
Investigation of heterogeneity and inconsistency

If we find important heterogeneity or/and inconsistency, we will explore the possible sources. If sufficient studies are available, we will perform meta-regression or subgroup analyses for the primary outcomes by using the following effect modifiers as possible sources of inconsistency and or heterogeneity: (i) year of publication; (ii) psychiatric comorbidities; (iii) setting; (iv) intensity of the psychosocial treatment; (v) risk of bias; (vi) sample size; (vii) sex; (viii) age; (ix) type of stimulant used; (x) substitution therapy.

GRADE quality assessment of the comparisons in the network

The main results of the review will be presented in 'Summary of findings' (SoF) tables, as recommended by the Cochrane Collaboration (Schünemann 2011). We will make the SoF tables for estimates from the network meta-analysis based on the methodology developed from the GRADE Working Group (Atkins 2004). For more details, see Salanti 2014 (Salanti 2014). We will include an overall grading of the evidence for the following major outcomes:

- Acceptability: Drop-outs due to any cause, measured after 12 weeks of treatment.

We will grade quality of evidence considering study limitations, indirectness, inconsistency, imprecision of effect estimates, and risk of publication bias. According to the software GRADEpro (https://gradepro.org), we will assign four levels of quality of evidence: high, moderate, low, very low.

Sensitivity analysis

In order to obtain data about specific subgroups of stimulant-addicted patients, we will conduct analysis for the primary outcomes of studies including participants addicted to cocaine only.

In order to assess the possible confounding effect of methadone therapy on the outcomes, we will conduct separate analyses for the primary outcomes including only studies with participants undergoing methadone therapy.
References


