



OPEN ACCESS



Comparative oral monotherapy of psilocybin, lysergic acid diethylamide, 3,4-methylenedioxymethamphetamine, ayahuasca, and escitalopram for depressive symptoms: systematic review and Bayesian network meta-analysis

Tien-Wei Hsu,^{1,2,3} Chia-Kuang Tsai,⁴ Yu-Chen Kao,^{5,6} Trevor Thompson,⁷ Andre F Carvalho,⁸ Fu-Chi Yang,⁴ Ping-Tao Tseng,^{9,10,11,12} Chih-Wei Hsu,¹³ Chia-Ling Yu,¹⁴ Yu-Kang Tu,^{15,16} Chih-Sung Liang^{5,6}

For numbered affiliations see end of the article

Correspondence to: C-S Liang
lcsyfw@gmail.com
(ORCID 0000-0003-1138-5586)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2024;386:e078607
<http://dx.doi.org/10.1136/bmj-2023-078607>

Accepted: 20 June 2024

ABSTRACT

OBJECTIVE

To evaluate the comparative effectiveness and acceptability of oral monotherapy using psychedelics and escitalopram in patients with depressive symptoms, considering the potential for overestimated effectiveness due to unsuccessful blinding.

DESIGN

Systematic review and Bayesian network meta-analysis.

DATA SOURCES

Medline, Cochrane Central Register of Controlled Trials, Embase, PsycINFO, ClinicalTrial.gov, and World Health Organization's International Clinical Trials Registry Platform from database inception to 12 October 2023.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES

Randomised controlled trials on psychedelics or escitalopram in adults with depressive symptoms. Eligible randomised controlled trials of psychedelics (3,4-methylenedioxymethamphetamine (known as MDMA), lysergic acid diethylamide (known as LSD), psilocybin, or ayahuasca) required oral monotherapy with no concomitant use of antidepressants.

DATA EXTRACTION AND SYNTHESIS

The primary outcome was change in depression, measured by the 17-item Hamilton depression rating scale. The secondary outcomes were all cause discontinuation and severe adverse events. Severe adverse events were those resulting in any of a list of negative health outcomes including, death, admission to hospital, significant or persistent incapacity, congenital birth defect or abnormality, and suicide attempt. Data were pooled using a random effects model within a Bayesian framework. To avoid estimation bias, placebo responses were distinguished between psychedelic and antidepressant trials.

RESULTS

Placebo response in psychedelic trials was lower than that in antidepressant trials of escitalopram (mean difference -3.90 (95% credible interval -7.10 to -0.96)). Although most psychedelics were better than placebo in psychedelic trials, only high dose psilocybin was better than placebo in antidepressant trials of escitalopram (mean difference 6.45 (3.19 to 9.41)). However, the effect size (standardised mean difference) of high dose psilocybin decreased from large (0.88) to small (0.31) when the reference arm changed from placebo response in the psychedelic trials to antidepressant trials. The relative effect of high dose psilocybin was larger than escitalopram at 10 mg (4.66 (95% credible interval 1.36 to 7.74)) and 20 mg (4.69 (1.64 to 7.54)). None of the interventions was associated with higher all cause discontinuation or severe adverse events than the placebo.

CONCLUSIONS

Of the available psychedelic treatments for depressive symptoms, patients treated with high dose psilocybin showed better responses than those treated with placebo in the antidepressant trials, but the effect size was small.

SYSTEMATIC REVIEW REGISTRATION

PROSPERO, CRD42023469014.

Introduction

Common psychedelics belong to two classes: classic psychedelics, such as psilocybin, lysergic acid diethylamide (known as LSD), and ayahuasca; and entactogens, such as 3,4-methylenedioxymethamphetamine (MDMA).¹ Several randomised controlled trials have shown efficacy of psychedelics for people with clinical

WHAT IS ALREADY KNOWN ON THIS TOPIC

Psychedelic treatment resulted in significant efficacy in treating depressive symptoms and alleviating distress related to life threatening diagnoses and terminal illness

Meta-analyses have reported standardised mean difference of psychedelics for depression reduction ranging from 1.37 to 3.12, while antidepressant trials were approximately 0.3

No network meta-analysis has examined comparative efficacy between psychedelics and antidepressants for depressive symptoms, and effect sizes of psychedelics might be overestimated because of unsuccessful blinding and response expectancies

WHAT THIS STUDY ADDS

To avoid estimation bias, placebo responses in psychedelic and antidepressant trials were separated; placebo response in psychedelic trials was lower than that in antidepressant trials

Among all psychedelics studied, only high dose psilocybin was associated with greater effectiveness than placebo response in antidepressant trials (standardised mean difference 0.31)

Among all psychedelics, only high dose psilocybin was associated with greater effectiveness than escitalopram

depression.^{2 3} The proposed mechanism of its fast and persistent antidepressant effects is to promote structural and functional neuroplasticity through the activation of intracellular 5-HT_{2A} receptors in the cortical neurons.⁴ Additionally, the increased neuroplasticity was associated with psychedelic's high affinity directly binding to brain derived neurotrophic factor receptor TrkB, indicating a dissociation between the hallucinogenic and plasticity promoting effects of psychedelics.⁵ A meta-analysis published in 2023 reported that the standardised mean difference of psychedelics for depression reduction ranged from 1.37 to 3.12,² which are considered large effect sizes.⁶ Notably, the standardised mean difference of antidepressant trials is approximately 0.3 (a small effect size).^{7 8}

Although modern randomised controlled trials involving psychedelics usually use a double blinded design, the subjective effects of these substances can compromise blinding.⁹ Unsuccessful blinding may lead to differing placebo effects between the active and control groups, potentially introducing bias into the estimation of relative treatment effects.¹⁰ Concerns have arisen regarding the overestimated effect sizes of psychedelics due to the issues of blinding and response expectancy.⁹ Psychedelic treatment is usually administered with psychological support or psychotherapy, and thereby the isolated pharmacological effects of psychedelics remain to be determined.² Surprisingly, on 1 July 2023, Australia approved psilocybin for the treatment of depression¹¹; the first country to classify psychedelics as a medicine at a national level.

To date, only one double blind, head-to-head randomised controlled trial has directly compared a psychedelic drug (psilocybin) with an antidepressant drug (escitalopram) for patients with major depressive disorder.¹² This randomised controlled trial reported that psilocybin showed a better efficacy than escitalopram on the 17 item Hamilton depression rating scale (HAMD-17).

We aimed to assess the comparative effectiveness and acceptability of oral monotherapy with psychedelics and escitalopram in patients experiencing depressive symptoms. Given that unsuccessful blinding can potentially lead to a reduced placebo response in psychedelic trials, we distinguished between the placebo responses in psychedelic and antidepressant trials. We also investigated the differences in patient responses between people who received extremely low dose psychedelics as a placebo and those who received a placebo in the form of a fake pill, such as niacin, in psilocybin trials.^{13 14} Our study allowed for a relative effect assessment of psychedelics compared with placebo responses observed in antidepressant trials.

Methods

The study protocol was registered with PROSPERO (CRD42023469014). We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension statement for reporting

systematic reviews incorporating network meta-analysis (NMA) (appendix 1).¹⁵

Data sources and searches

A comprehensive search of the Medline, Cochrane Central Register of Controlled Trials (CENTRAL), Embase, PsycINFO, ClinicalTrials.gov, and World Health Organization's International Clinical Trials Registry Platform databases were performed without language restrictions from database inception to 12 October 2023. We also searched the grey literature and reviewed reference lists of the included studies and related systematic reviews.^{2 3}

Study selection

Eligible studies were randomised controlled trials with parallel group or crossover designs. We included: (i) adults (≥ 18 years) with clinically diagnosed depression (eg, major depressive disorder, bipolar disorder, or other psychiatric disorders with comorbid clinical depression) or life threatening diagnoses and terminal illness with depressive symptoms; and (ii) adults with assessment of treatment response (preapplication/postapplication) using standard, validated, and internationally recognised instruments, such as HAMD-17. The outcome of interest was the change in depressive symptoms at the end of treatment compared with the controls, and we only extracted data from the first phase of crossover randomised controlled trials to avoid carry-over effects. Eligible psychedelic randomised controlled trials (including psilocybin, lysergic acid diethylamide, MDMA, and ayahuasca without dosage limit) required oral monotherapy without the concomitant use of antidepressants. For escitalopram, we included only fixed dose randomised controlled trials that compared at least two arms with different doses of oral form escitalopram (maximum dose of 20 mg/day) with placebo because psychedelic therapies usually use a fixed dose study design. We also included randomised controlled trials that evaluated psychedelic monotherapy compared with escitalopram monotherapy. We excluded follow-up studies and studies with healthy volunteers. We also excluded conference abstracts, editorials, reviews, meta-analyses, case reports, and case series, as well as publications reporting duplicate data. We did not consider ketamine because this drug is usually administered parenterally and is not a classic psychedelic.¹⁶ Screening and selection of the studies were performed independently by two authors. Discrepancies in study inclusion were resolved by deliberation among the reviewer pairs or with input from a third author. Appendix 2 shows the complete search strategies, and appendix 3 presents the reasons for exclusion.

Definition of outcomes, data extraction, and risk of bias assessment

The primary outcome was change in depressive symptoms from baseline (continuous outcome), as measured by a validated rating scale, such as HAMD-

17. When multiple measurement tools were used, they were selected in the following order: the HAMD-17, Montgomery-Åsberg depression rating scale, and Beck depression inventory (second edition). To improve interpretability, all extracted depression scores were converted to corresponding HAMD-17 scores using a validated method.¹⁷ We used a conservative correlation coefficient of 0.5 or other statistics (eg, *t* statistics) to calculate the standard deviation of change from baseline when unreported.¹⁸ The secondary outcomes were all cause discontinuation and severe adverse events (categorical outcomes). Severe adverse events were classified as those resulting in any of a list of negative health outcomes including, death, admission to hospital, significant or persistent incapacity, congenital birth defect or abnormality, and suicide attempt. Outcome data were extracted from original intention-to-treat or last observation carrying forward analysis, as well as from estimates of mixed-effect models for repeated measures.

Two authors independently extracted and reviewed the data, each being reviewed by another author. WebPlot Digitizer (<https://automeris.io/WebPlotDigitizer/>) was used to extract numerical data from the figures. Two authors independently used the Cochrane randomised trial risk of bias tool (version 2.0) to assess the risk of bias in the included trials, and discrepancies were resolved by consensus.¹⁹

Data synthesis

To estimate the relative effect between two interventions, we computed mean difference on the basis of change values (with 95% credible interval) for continuous outcomes (change in depressive symptoms) and odds ratios for categorical outcomes (all cause discontinuation and severe adverse event). To assess the clinical significance of the relative effect, we evaluated whether the mean difference exceeded the minimal important difference, which is estimated to be 3 points for HAMD-17.²⁰ We defined high, low, and extremely low doses of the included psychedelics as follows: (i) psilocybin: high dose (≥ 20 mg), extremely low dose (1–3 mg), low dose (other range); and (ii) MDMA: high dose (≥ 100 mg), extremely low dose (≤ 40 mg), low dose (other range). Escitalopram was divided into escitalopram 10 mg and escitalopram ≥ 20 mg. In previous clinical trials, a dose of 1 mg of psilocybin or a dose range of 1–3 mg/70 kg were used as an active control because these doses were believed not to produce significant psychedelic effects.^{21–22} A dose of 5 mg/70 kg can produce noticeable psychedelic effects.²² In many two arm psilocybin trials, the psilocybin dose in the active group typically falls within the range of 20–30 mg.^{12 21 23 24} In a three arm trial, 25 mg was defined as high dose, and 10 mg was considered a moderate dose.²¹ Another clinical trial also defined 0.215 mg/kg of psilocybin as a moderate dose for the active group.²⁵ Therefore, we used 20 mg and 3 mg as the boundaries for grouping psilocybin doses; when the dosage was calculated per kilogram in the study, we converted it to per 70 kg. For MDMA, in two trials

with three arms, 125 mg was defined as high dose, and 30–40 mg was defined as active control.^{26 27} Thus, we used 100 mg and 40 mg as the boundaries for grouping MDMA doses.

We conducted random effects network meta-analysis and meta-analysis within a Bayesian framework.^{28 29} Previous meta-analyses considered all control groups as a common comparator; however, concerns have been raised regarding the overestimated effect sizes of psychedelics because of unsuccessful blinding and poor placebo response.⁹ Therefore, we treated the three treatments as distinct interventions: the placebo response observed in psychedelic trials, the placebo response observed in antidepressant escitalopram trials, and extremely low dose psychedelics (ie, psilocybin and MDMA). We calculated the relative effects of all interventions compared with these three groups, indicating the following three conditions: (1) the treatment response of placebo response in the psychedelic trials is assumed to be lower than that of placebo response in antidepressant trials because of unsuccessful blinding.⁹ As such, the relative effects compared with placebo response in the psychedelic trials represented potential overestimated effect sizes. (2) the placebo response in antidepressant trials is assumed to be the placebo response in antidepressant trials with adequate blinding, therefore, the relative effects compared with placebo response in antidepressant trials represents effect sizes in trials with adequate blinding. (3) Psychedelic drugs are usually administered with psychotherapy¹³ or psychological support,¹⁴ the relative effects of psychedelics compared with extremely low dose psychedelics might eliminate the concomitant effects from psychotherapeutic support, approximating so-called pure pharmacological effects.

In network meta-analysis, the validity of indirect comparison relies on transitivity assumption.³⁰ We assessed the transitivity assumption by comparing the distribution of potential effect modifiers across treatment comparisons. In addition, we assessed whether the efficacy of escitalopram is similar in placebo controlled randomised controlled trials (escitalopram *v* placebo response in antidepressant trials) and in the head-to-head randomised controlled trial (psilocybin *v* escitalopram) using network meta-analysis.¹² Furthermore, we assessed the efficacy of the different placebo responses (placebo response in the psychedelic trials *v* placebo response in antidepressant trials) as additional proof of transitivity. If the placebo response in antidepressant trials was better than that in the psychedelic trials, the transitivity assumption did not hold when grouping placebo response in antidepressant trials and placebo response in the psychedelic trials together. Finally, for the primary outcome (change in depressive symptoms), network meta-regression analyses were conducted to evaluate the impact of potential effect modifiers, including proportion of men and women in the study, mean age, baseline depression severity, disorder type, and follow-up assessment period. We assumed a common effect

on all treatment comparisons for each of the effect modifiers. In other words, all interactions between the treatment comparisons and the effect modifier were constrained to be identical.

We also conducted the following sensitivity analyses: analysing studies of patients with major depressive disorder; excluding studies with a high risk of bias; adjusting for baseline depression severity; and using correlation coefficient of zero (most conservative) to calculate the standard deviation of change from baseline when unreported.

Publication bias was assessed by visual inspection of a comparison adjusted funnel plots. The first funnel plot used placebo response in the psychedelic trials as the comparator. The second funnel plot used placebo response in antidepressant trials as the comparator. The third funnel plot used both placebo response in the psychedelic trials and placebo response in antidepressant trials as comparators simultaneously. Additionally, we conducted the Egger test, Begg test, and Thompson test to examine the asymmetry of the third funnel plot. A previous meta-analysis reported that the standardised mean difference of psychedelics for depression reduction ranged from 1.37 to 3.12.² Therefore, we also transformed the effect size of mean difference to standardised mean difference (Hedges' *g*) for the primary outcome. The global inconsistency of the network meta-analysis was examined by fitting an unrelated main effects model. Local inconsistency of the network meta-analysis was examined using node splitting methods.³¹ Four Markov chains were implemented. 50 000 iterations occurred per chain and the first 20 000 iterations for each chain were discarded as a warm-up. Convergence was assessed by visual inspection of the trace plots of the key parameters for each analysis. The prior settings and convergence results are shown in appendix 4. All statistical analyses were done using R version 4.3.1. The network meta-analysis and pairwise meta-analysis within a Bayesian framework were fitted using the Bayesian statistical software called Stan within the R packages *multinma*²⁸ and *brms*,²⁹ respectively. The frequentist random effects network meta-analysis, funnel plots, and tests for funnel plot asymmetry were conducted using the R package *netmeta*. Reasons for protocol changes are in appendix 5.

Assessment certainty of evidence for the primary outcome

The certainty of evidence produced by the network meta-analysis was evaluated using GRADE (Grading of recommendations, assessment, development and evaluation).³²⁻³⁴ We used a minimally contextualised framework with the value of 3 (minimal important difference) as our decision threshold. The certainty of evidence refers to our certainty that the intervention had, relative to minimal intervention, any clinically minimal important difference. The optimal information size was calculated using a validated method.³²⁻³⁴

Patient and public involvement

Both patients and the public are interested in research on novel depression treatments and their efficacy compared with existing antidepressants. However, due to a scarcity of available funding for recruitment and researcher training, patients and members of the public were not directly involved in the planning or writing of this manuscript. We did speak to patients about the study, and we asked a member of the public to read our manuscript after submission.

Results

Characteristics of included study

After searching the database and excluding duplicated records, we identified 3104 unique potential studies. We then screened the titles and abstracts of these studies for eligibility and excluded 3062 of them, in which 42 studies remained. Twenty six studies were excluded after an assessment of the full text for various reasons (appendix 3). We identified three additional studies through a manual search resulting in total 19 eligible studies (efigure 1). Details of the characteristics of the included studies are shown in table 1. Protocols of psychological support or psychotherapy with psychedelic treatment are shown in table 2. Overall, 811 people (mean age of 42.49 years, 54.2% (440/811) were women) were included in psychedelic trials (15 trials), and 1968 participants (mean age of 39.35 years, 62.5% (1230/1968) were women) were included in escitalopram trials (five trials).

Risk of bias of the included studies

No psychedelic study (0/15) had a high overall risk of bias (efigure 2A and efigure 3A). The percentages of studies with high, some concerns, or low risk of bias in the 15 psychedelic trials were as follows: 0% (k=15), 33% (k=5), and 67% (k=10) for randomisation; 0% (k=0), 33% (k=5), and 67% (k=10) for deviations from intended interventions; 0% (k=0), 13% (k=2), and 87% (k=13) for missing outcome data; 0% (k=0), 33% (k=5), and 67% (k=10) for measurements of outcomes; 0% (k=0), 67% (k=1), and 93% (k=14) for selection of reported results. No non-psychedelic studies (0/5) were rated as high risk of bias (efigure 2B and efigure 3B). The percentages of studies with high, some concerns, and low risk of bias in the five non-psychedelic trials were as follows: 0% (k=0), 80% (k=4), and 20% (k=1) for randomisation; 0% (k=0), 100% (k=5), and 0% (k=0) for deviations from intended interventions; 0% (k=0), 80% (k=4), and 20% (k=1) for missing outcome data; 0% (k=0), 80% (k=4), and 20% (k=1) for measurements of outcomes; 0% (k=0), 20% (k=1), and 80% (k=4) for selection of reported results.

Network meta-analysis

In the network structure, all interventions were connected, with two main structures (fig 1). All psychedelics were compared with placebo response in the psychedelic trials, and escitalopram was compared with placebo response in antidepressant

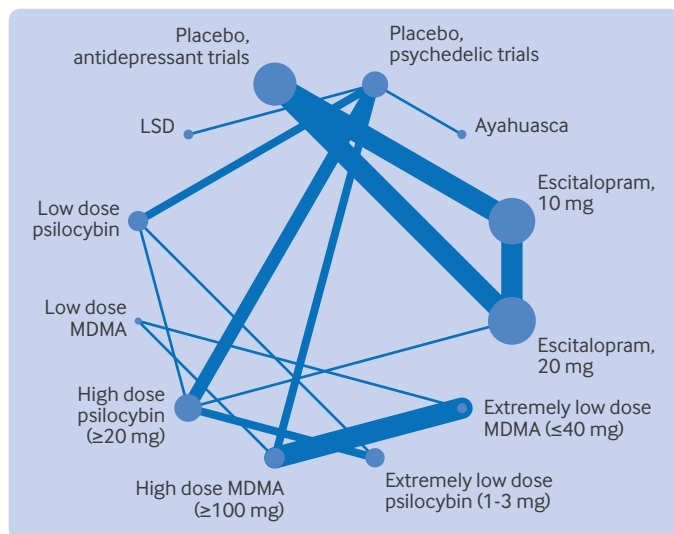


Fig 1 | Network structure. LSD=lysergic acid diethylamide; MDMA=3,4-methylenedioxymethamphetamine

trials. A head-to-head comparison of high dose psilocybin and 20 mg escitalopram connected the two main structures.¹²

In the main network meta-analysis, all interventions, except for extremely low dose and low dose MDMA, were associated with a larger mean difference exceeding the minimal important difference of 3 points on the HAMD-17 than with placebo response in the psychedelic trials (fig 2). Notably, placebo response in antidepressant trials (3.79 (95% credible interval 0.77 to 6.80)) and extremely low dose psilocybin (3.96 (0.61 to 7.17)) were better than placebo response in the psychedelic trials, with mean differences exceeding 3 and 95% credible intervals that did not cross zero. Additionally, in comparison with placebo response in antidepressant trials (fig 2), the relative effects of high dose psilocybin (6.52 (3.19 to 9.57)), escitalopram 10 mg (1.86 (0.21 to 3.50)), and escitalopram 20 mg (1.82 (0.16 to 3.43)) did not cross zero. Only high dose psilocybin resulted in a mean difference that was greater than 3. The standardised mean difference of high dose psilocybin decreased from large (0.88) to small (0.31) when the reference arm was changed from placebo response in the psychedelic trials to placebo response in antidepressant trials.

When compared with extremely low dose psilocybin (fig 2), only the relative effects of high dose psilocybin (6.35 (95% credible interval 3.41 to 9.21)) and placebo response in the psychedelic trials (-3.96 (-7.17 to -0.61)) showed a larger mean difference exceeding 3, without crossing zero. All relative effects between interventions are shown in efigure 4. Importantly, the mean differences of high dose psilocybin compared with escitalopram 10 mg (4.66 (1.36 to 7.74); standardised mean difference 0.22), escitalopram 20 mg (4.69 (1.64 to 7.54); 0.24), high dose MDMA (4.98 (1.23 to 8.67); 0.32), and low dose

psilocybin (4.36 (1.20 to 7.51); 0.32) all exceeded 3 and did not cross zero (efigure 4).

Transitivity assumption

The assessment of transitivity assumption is shown in efigure 5 and efigure 6. We compared the efficacy of escitalopram in the placebo controlled antidepressant trials⁸ with that in the head-to-head trial (psilocybin *v* escitalopram)¹² using network meta-analysis and pairwise meta-analysis. The results of the network meta-analysis showed that the relative effects between these two study designs (0.64 (95% credible interval -4.41 to 5.40), efigure 6A; 1.94 (-2.66 to 6.14), efigure 6B) included zero, and the mean differences did not exceed 3. Placebo response in antidepressant trials was better than placebo response in the psychedelic trials with a small effect size (3.79 (0.77 to 6.80), standardised mean difference 0.2), and the mean difference exceeded 3 (fig 2).

Sensitivity analyses

When including only patients with major depressive disorder, the relative effects of escitalopram 20 mg, escitalopram 10 mg, ayahuasca, and high dose psilocybin were better than placebo response in antidepressant trials, while placebo response in the psychedelic trials was worse than placebo response in antidepressant trials (fig 3). However, only the mean differences for high dose psilocybin (6.82 (95% credible interval 3.84 to 9.67)), ayahuasca (5.38 (0.02 to 10.61)), and placebo response in the psychedelic trials (-4.00 (-6.87 to -1.13)) exceeded 3. When compared with extremely low dose psilocybin (excluding the effects from concomitant psychotherapeutic support), only the 95% credible intervals of the relative effects of high dose psilocybin (4.36 (0.54 to 8.27); standardised mean difference 0.30) and placebo response in the psychedelic trials (-6.46 (-10.41 to -2.32), standardised mean difference -0.46) exceeded 3 and did not cross zero (fig 3). All of the relative effects between interventions are shown in efigure 7. Notably, the relative effects of high dose psilocybin compared with escitalopram 10 mg (4.96 (1.97 to 7.82)), escitalopram 20 mg (4.97 (2.19 to 7.64)), and low dose psilocybin (3.82 (0.61 to 7.04)) all exceeded 3 and did not cross zero (efigure 7).

The other three sensitivity analyses showed similar findings with the main analyses: exclusion of studies with high risk of bias (efigure 8); adjustment of baseline depression severity (efigure 9); and use of most conservative correlation coefficient of zero (efigure 10).

All cause discontinuation and severe adverse event

When referencing placebo in psychedelic trials, no interventions were associated with higher risks of all cause discontinuation rate nor severe adverse event rate (efigure 11).

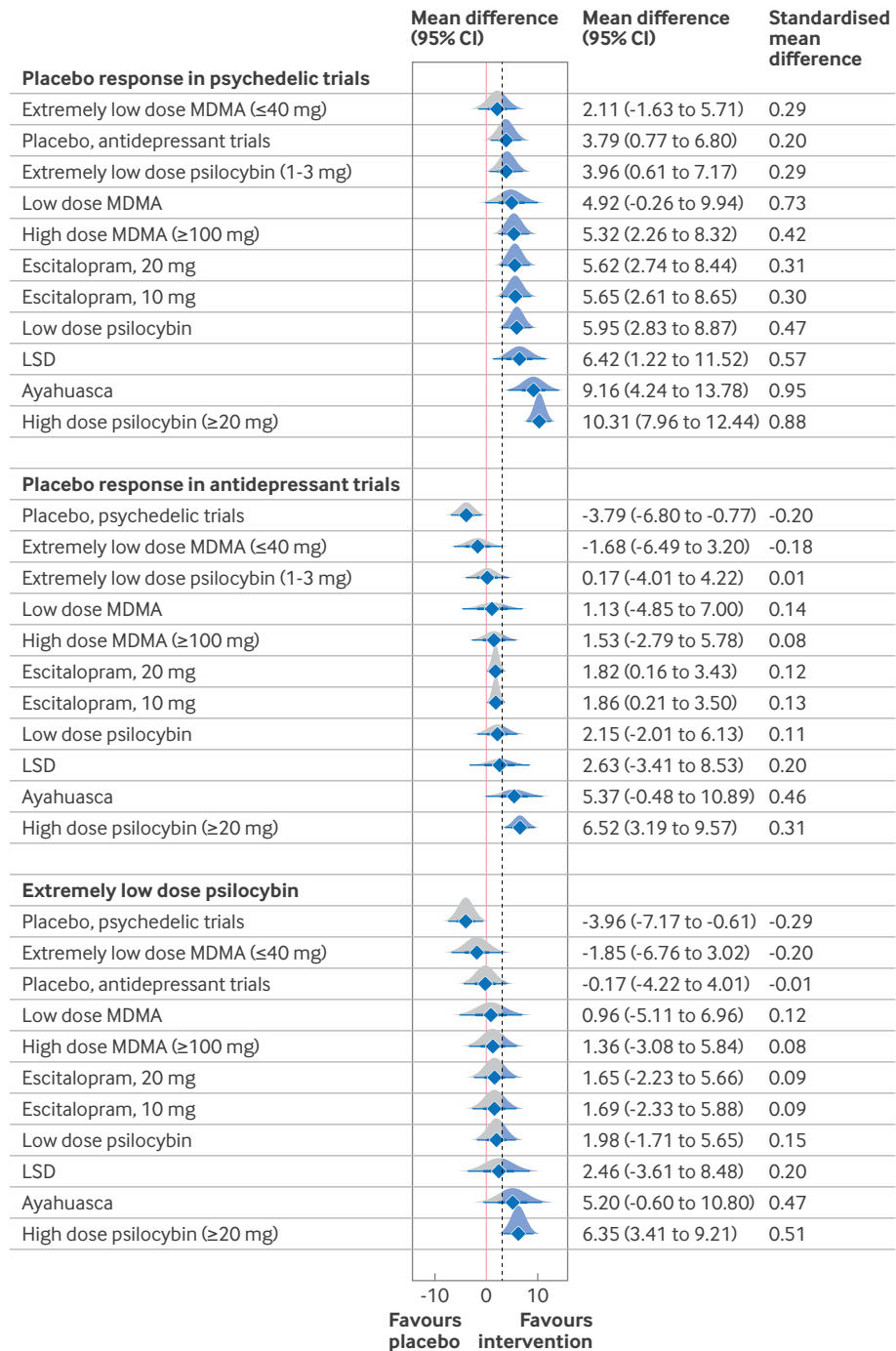


Fig 2 | Forest plots of network meta-analytical estimates v different reference arms by observed placebo response. The dotted line represents the minimal important difference of 3 whereas the red line indicates 0. LSD=lysergic acid diethylamide; MDMA=3,4-methylenedioxymethamphetamine

Network meta-regression and publication bias

In network meta-regression analyses, the 95% credible intervals of the relative effects of the baseline depressive severity, mean age, and percentage of women, crossed zero (etable 3). The results of the statistical tests (Egger, Begg, and Thompson-Sharp tests) for funnel plot asymmetry and visual inspection of funnel plots did not show publication bias (efigure 12). The results of GRADE

assessment are provided in the efigure 13. Most of the certainty of evidence for treatment comparisons was moderate or low.

Consistency assumptions

The back calculation methods for all the models (appendix 6) did not show any inconsistencies. The node splitting methods also did not show any inconsistencies (appendix 7).

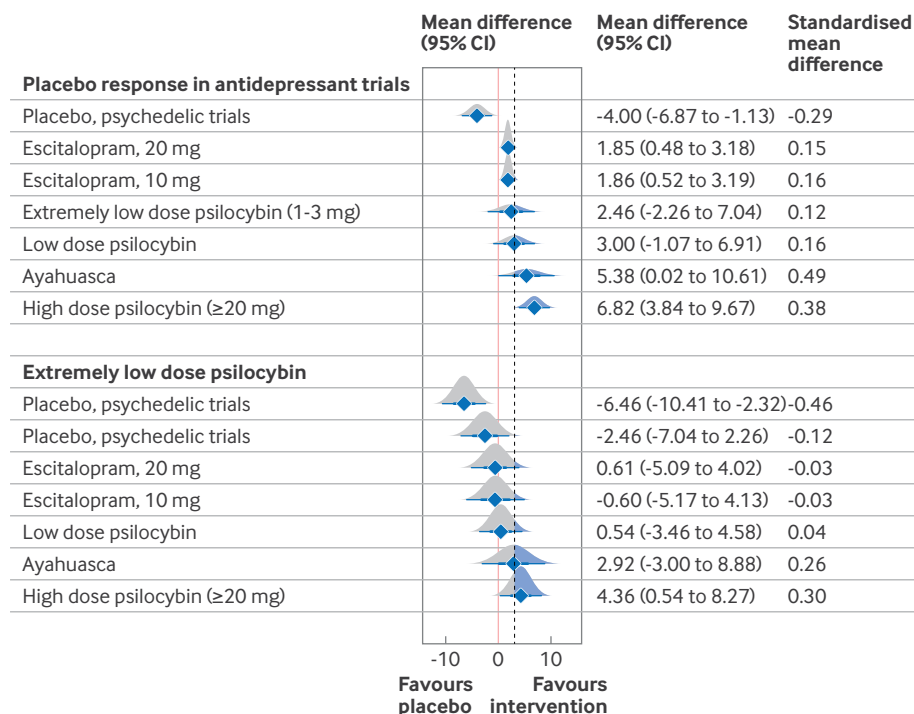


Fig 3 | Forest plots of network meta-analytical estimates when considering a population with major depressive disorder

Discussion

Principal findings

This network meta-analysis investigated the comparative effectiveness between psychedelics and escitalopram for depressive symptoms. Firstly, we found that the placebo response observed in antidepressant trials was associated with greater effectiveness than that observed in psychedelic trials. Secondly, when compared with placebo responses in antidepressant trials, only escitalopram and high dose psilocybin were associated with greater effectiveness, and only high dose psilocybin exceeded minimal important difference of 3. Notably, the effect size of high dose psilocybin decreased from large to small. Thirdly, among the included psychedelics, only high dose psilocybin was more likely to be better than escitalopram 10 mg or 20 mg, exceeding the minimally important difference of 3. Fourthly, in patients with major depressive disorder, escitalopram, ayahuasca, and high dose psilocybin were associated with greater effectiveness than placebo responses in antidepressant trials; however, only high dose psilocybin was better than extremely low dose psilocybin, exceeding minimal important difference of 3. Taken together, our study findings suggest that among psychedelic treatments, high dose psilocybin is more likely to reach the minimal important difference for depressive symptoms in studies with adequate blinding design, while the effect size of psilocybin was similar to that of current antidepressant drugs, showing a mean standardised mean difference of 0.3.⁷

Comparison with other studies

In a randomised controlled trial, treatment response was defined as the response observed in the active arm; placebo response was defined as the response observed in the control (placebo) arm.¹⁰ Treatment response consists of non-specific effects, placebo effect, and true treatment effect; placebo response consisted of non-specific effects and placebo effect. Therefore, when the placebo effect is not the same for the active and control arms within a randomised controlled trial, the estimation of the true treatment effect is biased. For example, in a psychedelic trial, unsuccessful blinding may occur due to the profound subjective effects of psychedelics. This unblinding may lead to high placebo effect in the active arm and low placebo effect in the control arms, and the true treatment effect is overestimated.¹⁰ Without addressing unequal placebo effects within studies, the estimation of meta-analysis and network meta-analysis are biased.¹⁰ However, in most psychedelic trials, blinding was either reported as unsuccessful or not assessed at all. For example, two trials of lysergic acid diethylamide reported unsuccessful blinding,^{35 36} whereas the trial of ayahuasca only reported that five of 10 participants misclassified the placebo as ayahuasca.³⁷ In trials of MDMA, participants' accuracy in guessing which treatment arm they were in ranged from approximately 60-90%.^{26 27 38-40} In the case of most psilocybin trials, blinding was not assessed, with the exception of the study by Ross and colleagues in 2016.¹³ In that study, participants were asked to guess whether the psilocybin or an

active control was received, and the correct guessing rate was 97%. In our study, we established several network meta-analysis models addressing this issue, and we found that placebo response in the psychedelic trials was associated with less effectiveness than that in antidepressant trials. Therefore, the effect sizes of psychedelics compared with placebo response observed in psychedelic trials may be overestimated. All of the psychedelics' 95% credible intervals of the relative effects crossed zero when compared with the placebo response in antidepressant trials, except for high dose psilocybin.

The comparisons between psychedelics and escitalopram showed that high dose psilocybin was more likely to be better than escitalopram. Psilocybin was usually administered with psychotherapy or psychological support.^{13 14} Therefore, the greater effectiveness of psilocybin may be from not only pharmacological effects but also psychotherapeutic support. However, we also found that high doses of psilocybin was associated greater effectiveness than extremely low doses of psilocybin. This effect also indicates that the effectiveness of psilocybin cannot be attributed only to concomitant psychotherapy or psychological support.

In patients with major depressive disorder, ayahuasca, low dose psilocybin, high dose psilocybin, escitalopram 10 mg, and escitalopram 20 mg were associated with greater effectiveness than the placebo response in antidepressant trials. However, when compared with extremely low dose psilocybin, only high dose psilocybin was associated with better effectiveness; the standardised mean difference decreased from 0.38 (compared with placebo response in antidepressant trials) to 0.30 (compared with extremely low dose psilocybin). As such, the effectiveness of psilocybin should be considered with concomitant psychotherapeutic support in people with major depressive disorder. The effect size of high dose psilocybin was similar with antidepressant trials of patients with major depressive disorder showing a mean standardised mean difference of 0.3.^{7 8}

Strengths and limitations of this study

This study has several strengths. We conducted separate analyses for placebo response in antidepressant trials, placebo response in psychedelic trials, and an extremely low active dose of psychedelics, thereby mitigating the effect of placebo response variations across different studies. This approach allowed us to assess the efficacy of psychedelics more impartially and make relatively unbiased comparisons than if these groups were not separated. This study supported the transitivity assumption of the efficacy of escitalopram in placebo controlled antidepressant trials with that in psilocybin versus escitalopram head-to-head trial, thereby bridging the escitalopram trials and psychedelic trials. We also performed various sensitivity analyses to ensure the validation of our statistical results.

Nevertheless, our study has several limitations. Firstly, we extracted only the acute effects of the interventions. A comparison of the long term effects of psychedelics and escitalopram remains unclear. Secondly, participants in the randomised controlled trials on MDMA were predominantly diagnosed with post-traumatic stress disorder, whereas participants in the randomised controlled trials on escitalopram were patients with major depressive disorder. However, depressive symptoms in post-traumatic stress disorder could be relatively treatment resistant, requiring high doses of psychotropic drugs.⁴¹ Moreover, our study focused not only on major depressive disorder but also on the generalisability of psychedelic treatment for depressive symptoms. Thirdly, although all available studies were included, the sample size of the psychedelic randomised controlled trials was small ($k=15$). Fourthly, when using extremely low dose psychedelics as a reference group, the relative effect may also eliminate some pharmacological effects because our study found that extremely low dose psychedelics could not be considered a placebo. Fifthly, in network meta-analysis, direct evidence for one treatment comparison may serve as indirect evidence for other treatment comparisons,⁴² and biases in the direct evidence might affect estimates of other treatment comparisons. Because the absolute effect of escitalopram in the head-to-head trial (high dose psilocybin v escitalopram 20 mg)¹² was lower than those of placebo controlled trials, the relative effects of high dose psilocybin might be slightly overestimated when compared with other treatments in the current study. We addressed this issue by use of a Bayesian network meta-analysis, distinguishing between placebo response in psychedelic trials and placebo response in antidepressant trials. Specifically, we only considered that the 95% credible interval of the relative effect between two comparisons did not cross zero. Indeed, the relative effect of escitalopram 20 mg between these two study designs included zero. Finally, our network meta-analysis may not have sufficient statistical power to detect potential publication bias due to the scarcity of trials and participants.

Implications and conclusions

Serotonergic psychedelics, especially high dose psilocybin, appeared to have the potential to treat depressive symptoms. However, study designs may have overestimated the efficacy of psychedelics. Our analysis suggested that the standardised mean difference of high dose psilocybin was similar to that of current antidepressant drugs, showing a small effect size. Improved blinding methods and standardised psychotherapies can help researchers to better estimate the efficacy of psychedelics for depressive symptoms and other psychiatric conditions.

AUTHOR AFFILIATIONS

¹Department of Psychiatry, E-DA Dachang Hospital, I-Shou University, Kaohsiung, Taiwan

²Department of Psychiatry, E-DA Hospital, I-Shou University, Kaohsiung, Taiwan

³Graduate Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

⁴Department of Neurology, Tri-Service General Hospital, National Defense Medical Centre, Taipei, Taiwan

⁵Department of Psychiatry, National Defense Medical Centre, Taipei, Taiwan

⁶Department of Psychiatry, Beitou Branch, Tri-Service General Hospital, Taipei, Taiwan

⁷Centre for Chronic Illness and Ageing, University of Greenwich, London, UK

⁸IMPACT (Innovation in Mental and Physical Health and Clinical Treatment) Strategic Research Centre, School of Medicine, Barwon Health, Deakin University, Geelong, VIC, Australia

⁹Institute of Biomedical Sciences, National Sun Yat-sen University, Kaohsiung, Taiwan

¹⁰Department of Psychology, College of Medical and Health Science, Asia University, Taichung, Taiwan

¹¹Prospect Clinic for Otorhinolaryngology and Neurology, Kaohsiung, Taiwan

¹²Institute of Precision Medicine, National Sun Yat-sen University, Kaohsiung, Taiwan

¹³Department of Psychiatry, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Kaohsiung, Taiwan

¹⁴Department of Pharmacy, Chang Gung Memorial Hospital Linkou, Taoyuan, Taiwan

¹⁵Institute of Health Data Analytics and Statistics, College of Public Health, National Taiwan University, Taipei, Taiwan

¹⁶Department of Dentistry, National Taiwan University Hospital, Taipei, Taiwan

Contributors: T-WH and C-KT contributed equally to this work and are joint first authors. Y-KT and C-SL contributed equally to this work and are joint last/corresponding authors. C-SL, T-WH, and Y-KT conceived and designed the study. T-WH, C-KT, C-WH, and P-TT selected the articles, extracted the data, and assess the risk of bias. C-LY did the systemic search. T-WH and C-SL wrote the first draught of the manuscript. TT, AFC, Y-CK, F-CY, and Y-KT interpreted the data and contributed to the writing of the final version of the manuscript. C-KT and T-WH have accessed and verified the data. C-SL and Y-KT were responsible for the decision to submit the manuscript. All authors confirmed that they had full access to all the data in the study and accept responsibility to submit for publication. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: The study was supported by grant from the National Science and Technology Council (NSTC 112-2314-B-016-036-MY2 and NSTC 112-2314-B-002-210-MY3). The funding source had no role in any process of our study.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: support from National Science and Technology Council for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required because this study is an analysis of aggregated identified clinical trial data.

Data sharing: The data that support the findings of this study are available from the corresponding author (C-SL) upon reasonable request.

Transparency: The lead author (C-SL) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned

Dissemination to participants and related patient and public communities: Dissemination of the work to the public and clinical community through social media and lectures is planned.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>.

- 1 Tupper KW, Wood E, Yensen R, Johnson MW. Psychedelic medicine: a re-emerging therapeutic paradigm. *CMAJ* 2015;187:1054-9. doi:10.1503/cmaj.141124
- 2 Ko K, Kopra El, Cleare AJ, Rucker JJ. Psychedelic therapy for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord* 2023;322:194-204. doi:10.1016/j.jad.2022.09.168
- 3 Muttoni S, Ardissino M, John C. Classical psychedelics for the treatment of depression and anxiety: A systematic review. *J Affect Disord* 2019;258:11-24. doi:10.1016/j.jad.2019.07.076
- 4 Vargas MV, Dunlap LE, Dong C, et al. Psychedelics promote neuroplasticity through the activation of intracellular 5-HT2A receptors. *Science* 2023;379:700-6. doi:10.1126/science.adf0435
- 5 Moliner R, Giryck M, Brunello CA, et al. Psychedelics promote plasticity by directly binding to BDNF receptor TrkB. *Nat Neurosci* 2023;26:1032-41. doi:10.1038/s41593-023-01316-5
- 6 Sawilowsky SS. New effect size rules of thumb. *J Mod Appl Stat Methods* 2009;8:26. doi:10.22237/jmasm/1257035100
- 7 Khan A, Brown WA. Antidepressants versus placebo in major depression: an overview. *World Psychiatry* 2015;14:294-300. doi:10.1002/wps.20241
- 8 Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018;391:1357-66. doi:10.1016/S0140-6736(17)32802-7
- 9 Muthukumaraswamy SD, Forsyth A, Lumley T. Blinding and expectancy confounds in psychedelic randomized controlled trials. *Expert Rev Clin Pharmacol* 2021;14:1133-52. doi:10.1080/17512433.2021.1933434
- 10 Nikolakopoulou A, Chaimani A, Furukawa TA, et al. When does the placebo effect have an impact on network meta-analysis results? *BMJ Evid Based Med* 2024;29:127-134. doi:10.1136/bmjebm-2022-112197
- 11 Nogrady B. Australia's approval of MDMA and psilocybin for PTSD and depression is premature, say critics. *BMJ* 2023;382:1599. doi:10.1136/bmj.p1599
- 12 Carhart-Harris R, Giribaldi B, Watts R, et al. Trial of psilocybin versus escitalopram for depression. *N Engl J Med* 2021;384:1402-11. doi:10.1056/NEJMoa2032994
- 13 Ross S, Bossis A, Guss J, et al. Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: a randomized controlled trial. *J Psychopharmacol* 2016;30:1165-80. doi:10.1177/0269881116675512
- 14 Raison CL, Sanacora G, Woolley J, et al. Single-dose psilocybin treatment for major depressive disorder: a randomized clinical trial. *JAMA* 2023;330:843-53. doi:10.1001/jama.2023.14530
- 15 Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777-84. doi:10.7326/M14-2385
- 16 Marcantoni WS, Akoumba BS, Wassef M, et al. A systematic review and meta-analysis of the efficacy of intravenous ketamine infusion for treatment resistant depression: January 2009 - January 2019. *J Affect Disord* 2020;277:831-41. doi:10.1016/j.jad.2020.09.007
- 17 Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH. Pooling health-related quality of life outcomes in meta-analysis: a tutorial and review of methods for enhancing interpretability. *Res Synth Methods* 2011;2:188-203. doi:10.1002/jrsm.46
- 18 Higgins JPT, Thomas J, Chandler J, et al. *Cochrane handbook for systematic reviews of interventions version 6.3*: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch, 2022.
- 19 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898. doi:10.1136/bmj.l4898
- 20 Hengartner MP, Plödel M. Estimates of the minimal important difference to evaluate the clinical significance of antidepressants in the acute treatment of moderate-to-severe depression. *BMJ Evid Based Med* 2022;27:69-73. doi:10.1136/bmjebm-2020-111600
- 21 Goodwin GM, Aaronson ST, Alvarez O, et al. Single-dose psilocybin for a treatment-resistant episode of major depression. *N Engl J Med* 2022;387:1637-48. doi:10.1056/NEJMoa2206443
- 22 Griffiths RR, Johnson MW, Carducci MA, et al. Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *J Psychopharmacol* 2016;30:1181-97. doi:10.1177/0269881116675513
- 23 Davis AK, Barrett FS, May DG, et al. Effects of psilocybin-assisted therapy on major depressive disorder: a randomized clinical trial. *JAMA Psychiatry* 2021;78:481-9. doi:10.1001/jamapsychiatry.2020.3285
- 24 Raison CL, Sanacora G, Woolley J, et al. Single-dose psilocybin treatment for major depressive disorder: a randomized clinical trial. *JAMA* 2023;330:843-53. doi:10.1001/jama.2023.14530

- 25 von Rotz R, Schindowski EM, Jungwirth J, et al. Single-dose psilocybin-assisted therapy in major depressive disorder: a placebo-controlled, double-blind, randomised clinical trial. *EClinicalMedicine* 2022;56:101809. doi:10.1016/j.eclinm.2022.101809
- 26 Mithoefer MC, Mithoefer AT, Feduccia AA, et al. 3,4-methylenedioxyamphetamine (MDMA)-assisted psychotherapy for post-traumatic stress disorder in military veterans, firefighters, and police officers: a randomised, double-blind, dose-response, phase 2 clinical trial. *Lancet Psychiatry* 2018;5:486-97. doi:10.1016/S2215-0366(18)30135-4
- 27 O'Alora G M, Grigsby J, Poulter B, et al. 3,4-Methylenedioxyamphetamine-assisted psychotherapy for treatment of chronic posttraumatic stress disorder: A randomized phase 2 controlled trial. *J Psychopharmacol* 2018;32:1295-307. doi:10.1177/0269881118806297
- 28 Phillippo DM. multnma: Bayesian network meta-analysis of individual and aggregate data. 2020.
- 29 Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw* 2017;80:1-28. doi:10.18637/jss.v080.i01
- 30 Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014;9:e99682. doi:10.1371/journal.pone.0099682
- 31 Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-44. doi:10.1002/sim.3767
- 32 Brignardello-Petersen R, Bonner A, Alexander PE, et al, GRADE Working Group. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *J Clin Epidemiol* 2018;93:36-44. doi:10.1016/j.jclinepi.2017.10.005
- 33 Izcovich A, Chu DK, Mustafa RA, Guyatt G, Brignardello-Petersen R. A guide and pragmatic considerations for applying GRADE to network meta-analysis. *BMJ* 2023;381:e074495. doi:10.1136/bmj-2022-074495
- 34 Brignardello-Petersen R, Guyatt GH, Mustafa RA, et al. GRADE guidelines 33: Addressing imprecision in a network meta-analysis. *J Clin Epidemiol* 2021;139:49-56. doi:10.1016/j.jclinepi.2021.07.011
- 35 Gasser P, Holstein D, Michel Y, et al. Safety and efficacy of lysergic acid diethylamide-assisted psychotherapy for anxiety associated with life-threatening diseases. *J Nerv Ment Dis* 2014;202:513-20. doi:10.1097/NMD.0000000000000113
- 36 Holze F, Gasser P, Müller F, Dolder PC, Liechti ME. Lysergic acid diethylamide-assisted therapy in patients with anxiety with and without a life-threatening illness: a randomized, double-blind, placebo-controlled phase ii study. *Biol Psychiatry* 2023;93:215-23. doi:10.1016/j.biopsych.2022.08.025
- 37 Palhano-Fontes F, Barreto D, Onias H, et al. Rapid antidepressant effects of the psychedelic ayahuasca in treatment-resistant depression: a randomized placebo-controlled trial. *Psychol Med* 2019;49:655-63. doi:10.1017/S0033291718001356
- 38 Mitchell JM, Bogenschutz M, Lilienstein A, et al. MDMA-assisted therapy for severe PTSD: a randomized, double-blind, placebo-controlled phase 3 study. *Focus (Am Psychiatr Publ)* 2023;21:315-28. doi:10.1176/appi.focus.23021011
- 39 Mithoefer MC, Wagner MT, Mithoefer AT, Jerome L, Doblin R. The safety and efficacy of +/-3,4-methylenedioxyamphetamine-assisted psychotherapy in subjects with chronic, treatment-resistant posttraumatic stress disorder: the first randomized controlled pilot study. *J Psychopharmacol* 2011;25:439-52. doi:10.1177/0269881110378371
- 40 Wolfson PE, Andries J, Feduccia AA, et al. MDMA-assisted psychotherapy for treatment of anxiety and other psychological distress related to life-threatening illnesses: a randomized pilot study. *Sci Rep* 2020;10:20442. doi:10.1038/s41598-020-75706-1
- 41 Chiba H, Oe M, Uchimura N. Patients with posttraumatic stress disorder with comorbid major depressive disorder require a higher dose of psychotropic drugs. *Kurume Med J* 2016;62:23-8. doi:10.2739/kurumemedj.MS65010
- 42 Li H, Shih M-C, Song C-J, Tu YK. Bias propagation in network meta-analysis models. *Res Synth Methods* 2023;14:247-65. doi:10.1002/jrsm.1614

Web appendix: Extra material supplied by authors