Psilocybin for Treatment-Resistant Depression: Effectiveness and Value

Research Protocol

November 25, 2026

Prepared for

©Institute for Clinical and Economic Review, 2026

Table of Contents

1. Background, Objectives, and Research Questions 1
1.1. Background 1
1.2. Objectives 2
1.3. Research Questions 2
1.4. PICOTS Criteria 2
2. Evidence Review Methods 5
2.1. Search Methods and Data Sources 5
2.2. Eligibility Criteria 6
2.3. Study Selection Process 6
2.4. Data Extraction Strategy 7
2.5. Risk of Bias 7
2.6. Subgroup Credibility 8
2.7. Clinical Trial Diversity 8
2.8. Publication Bias 8
2.9. Evidence Synthesis 8
References 11
Appendix A. PRISMA Checklist A1
Appendix B. Search Strategies B1
Appendix C. Data Extraction Summary Table Shell C1

1. Background, Objectives, and Research Questions

1.1. Background

Major depressive disorder (MDD) is common; about one-fifth of Americans will experience MDD.[1,2] Appropriate initial treatments include first-line antidepressant medications, such as selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs), as well as psychotherapy.[3] Many Americans with MDD do not receive treatment at all, but among those who do, an estimated 2.8 million experience treatment-resistant depression (TRD).[1,2] Definitions for TRD vary,[4,5] but a common definition is inadequate response to two different antidepressant medications despite an adequate dose and duration of treatment. MDD accounts for annual United States (US) costs of more than $300 billion, about one-third of which are direct treatment costs and about two-thirds of which are workplace costs.[6] The annual national burden of TRD is estimated to account for 27%–41% of annual MDD costs.[2] For patients with MDD, the lifetime prevalence of suicide attempts is 31%, with 4% of women and 7% of men ultimately dying by suicide.[7,8] Patients with TRD have a greater than three-fold risk of death by suicide than patients with non-TRD MDD.[9]

The success of another course of antidepressant medications in TRD is much lower than in others with MDD; however, many patients with TRD are treated with additional courses.[10,11] Other appropriate interventions in TRD include augmentation of first-line antidepressant therapy with an antipsychotic medication.[12] In patients with persistent symptoms, recommended options include esketamine (S-ketamine), ketamine (used off-label for TRD), or repetitive transcranial magnetic stimulation (rTMS, often referred to as simply "TMS"). Electroconvulsive therapy may be used in more severe cases.[13]

Recently, interest has grown in the use of psychedelic medications for TRD. For example, psilocybin, the psychedelic agent found in "magic mushrooms,"[14,15] is being evaluated for TRD. Although psilocybin is still a Schedule I substance (illegal under federal law), in some states, such as Oregon and Colorado, there are existing programs that allow individuals to receive psilocybin sessions at licensed facilities. COMP360 is a synthetic preparation of psilocybin that has been developed by Compass Pathways. Compass Pathways completed its New Drug Application (NDA) submission to the U.S. Food and Drug Administration (FDA) on November 9, 2026 for COMP360 (synthetic psilocybin) to treat TRD,[16] and was also selected for the Commissioner's National Priority Voucher program, which potentially has a shortened one-to-two-month FDA review.[17]

1.2. Objectives

The scope of this project was previously available for public comment and has been revised upon further discussions and input from stakeholders. In accordance with the revised scope, this project will assess both the comparative clinical effectiveness and economic impacts of psilocybin (25 mg) for the treatment of TRD. The assessment aims to systematically evaluate the existing evidence, taking uncertainty into account. To that aim, the assessment is informed by two research components: a systematic review of the existing evidence and an economic evaluation. This document presents the protocol for the systematic review of existing evidence (i.e., the clinical review). See the model analysis plan (expected posting: January 22, 2027) for details on the proposed methodology and model structure that will be used for the economic evaluation.

1.3. Research Questions

To inform our review of the clinical evidence, we have developed the following research questions:

- What is the net health benefit of psilocybin (25 mg) versus esketamine in the population described below?
- What is the net health benefit of psilocybin (25 mg) versus intravenous ketamine in the population described below?
- What is the net health benefit of psilocybin (25 mg) versus Repetitive Transcranial Magnetic Stimulation (rTMS) in the population described below?

1.4. PICOTS Criteria

In line with the above research questions, the following specific criteria have been defined utilizing PICOTS (Population, Interventions, Comparators, Outcomes, Timing, Setting and Study Design) elements.

Population

The population of focus for the review is adults aged 18 years and older with major depressive disorder, without psychotic features, and for whom two or more prior antidepressants prescribed at adequate dose and duration, during the current episode, have not resulted in adequate therapeutic benefit. As data permit, we also plan to examine subgroups of patients suggested by the clinical literature, patients, and clinical experts. These include subgroups defined by:

- Age (e.g., 18-64 vs. ≥65 years of age)
- Number of prior antidepressant trials (e.g., two, three, four)

- Episode of depression (e.g., first or recurrent)
- Comorbidity with other mental health conditions (e.g., anxiety, personality disorder, bipolar, post-traumatic stress disorder [PTSD])
- Concomitant use of medication for mental health (e.g., antidepressants, antipsychotics)
- Prior exposure to psychedelics

Intervention

The intervention of interest is:

- Psilocybin (25 mg)

Comparators

Data permitting, we intend to compare with:

- Esketamine (SPRAVATO®, Johnson & Johnson)
- Intravenous (IV) ketamine, an anesthetic agent used off-label for TRD
- Repetitive Transcranial Magnetic Stimulation (rTMS)

Outcomes

The outcomes of interest are described in the list below.

- Patient-Important Outcomes
    - Changes in depressive symptoms (e.g., Montgomery–Åsberg Depression Rating Scale [MADRS])
    - Rate of response, remission, and relapse
    - Changes in functional outcomes (e.g., Sheehan Disability Scale [SDS])
    - Changes in other mental health symptoms (e.g., anxiety, positive and negative emotions)
    - Changes in health-related quality of life (e.g., EuroQol 5-Dimension 5-Level [EQ-5D-5L])
    - Changes in cognition
    - Impact on employment and education
    - Time burden on patients from treatment
    - Adverse events including but not limited to:
        - Serious adverse events including acute events requiring higher-level medical care

        - Suicidal ideation, behavior, and self-harm
        - Prolonged visual hallucinations (i.e., hallucinogen persisting perception disorder [HPPD])
        - Psychotic symptoms
        - Cardiovascular effects
- Other Outcomes
    - Adverse events including
        - Discontinuation due to adverse events
    - Effects associated with abuse potential (e.g., euphoria, hallucinations and other perceptual distortions, alterations in cognition, and changes in mood)

Timing

Evidence on intervention efficacy, safety, and effectiveness will be collected from studies of any duration.

Setting

Evidence from all relevant settings will be considered, including inpatient, outpatient/clinic, office, and home settings.

Study Design

Randomized controlled trials (RCTs) will be included. Non-randomized controlled trials and high-quality comparative observational studies will be considered.

2. Evidence Review Methods

2.1. Search Methods and Data Sources

Procedures for the systematic literature review assessing the evidence on psilocybin (25 mg) for treatment-resistant depression will follow established best practices.[18,19] The review will be reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[20] The PRISMA guidelines include a list of 27 checklist items, which are described further in Appendix A.

To identify relevant studies for psilocybin (25 mg), we will search MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials, and APA PsycInfo for relevant studies. Each search will be limited to English language studies of human subjects and will exclude articles indexed as guidelines, letters, editorials, narrative reviews, case reports, or news items.

To identify relevant studies for the comparators (i.e., esketamine, IV ketamine, and rTMS), we will begin our search by reviewing the clinical trials included in the systematic literature review and network meta-analysis (NMA) published by Saelens and colleagues (2025).[21] Prior to publishing this protocol, we conducted a targeted literature review to identify systematic literature reviews and network meta-analyses that have evaluated the intervention and comparators of interest in our population of interest. We identified and selected the Saelens et al. NMA because its eligibility criteria aligned with ours and because it was more comprehensive and methodologically rigorous than other published reviews. Using RCTs from the Saelens et al. NMA, we will review and include RCTs that compare all listed drugs versus placebo or to each other in order to develop the NMA network (see data analysis section below). Next, we will search MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials, and APA PsycInfo for records published since April 13th 2023 (i.e., date of the last search by Saelens et al.). We will also review other published SLRs and NMAs to supplement our list of included studies.[22,23]

We will also conduct systematic searches for observational studies of psilocybin and the comparators of interest to identify data on outcomes not available from clinical trials. As the 2019 ICER review on esketamine (Spravato) for TRD[24] previously systematically searched for and identified observational studies for the comparators, our comparator searches will be limited to records published since April 19th 2019 (i.e., date of the last search by ICER).

The search strategies will be developed using the Population, Intervention, Comparator, and Study Design elements described above. The proposed search strategies include a combination of indexing terms (MeSH terms in MEDLINE and EMTREE terms in EMBASE), as well as free-text terms, and are presented in Appendix B.

To supplement the database searches, we will perform a manual check of the reference lists of included trials and reviews and invite key stakeholders to share references germane to the scope of this project. We will also supplement our review of published studies with data from conference proceedings, regulatory documents, information submitted by manufacturers, and other grey literature when the evidence meets ICER standards (for more information, see ICER's grey literature policy).

2.2. Eligibility Criteria

To ensure the most comprehensive efficacy data is identified and included for psilocybin (25 mg), in cases where data on psilocybin are available for the MDD population, we will explore whether data exist for a subgroup of those with TRD by reaching out to authors.

Given difficulties detecting potential side effects that may be infrequent but clinically important for assessment of safety, studies that examine infrequent side effects for psilocybin even when the dose studied is greater or lesser dose than 25 mg may be considered.

To limit bias of small-study effects, we will exclude studies of the intervention and comparators with fewer than 10 participants in any arm. This criterion is in line with the criteria of a published NMA evaluating efficacy of treatments for TRD.[21]

2.3. Study Selection Process

Subsequent to the literature search and removal of duplicate citations using both online and local software tools, study selection will be accomplished through two levels of screening, at the abstract and full-text level.

Two reviewers - a human reviewer and an artificial intelligence (AI) reviewer - will independently conduct masked screening of the titles and abstracts of all identified publications using Loon Lens version 1.0 to provide a recommendation to include or exclude the publication. A third human reviewer will adjudicate any disagreement between the human and AI reviewers. No study will be excluded at abstract level screening due to insufficient information. For example, an abstract that does not report an outcome of interest would be accepted for further review in full text. Citations accepted during abstract-level screening will be retrieved in full text for review. Full-text screening will also be conducted by two reviewers. A human reviewer and an AI reviewer will independently screen the full texts for inclusion, and a third human reviewer will adjudicate. Reasons for exclusion will be categorized according to the PICOTS elements during both title/abstract and full-text review.

2.4. Data Extraction Strategy

Data will be extracted into Microsoft Excel. The basic design and elements of the extraction forms will follow those used for other ICER reports. Elements include a description of patient populations, sample size, duration of follow-up, funding source, study design features, interventions (agent, dosage, frequency, schedules), concomitant therapy allowed and used (agent, dosage, frequency, schedules), outcome assessments, and results.

The data extraction will be performed using Claude, an AI assistant created by Anthropic, taking the following steps:

1. A reviewer will create the extraction shell (datasheet template) and extract at least one study in the desired format.

2. A reviewer will create a prompt for Claude to extract. The prompt will state the purpose of the data extraction (i.e., for evidence synthesis in HTA), outline the extraction task, and provide all relevant documents, including the shell datasheet plus an example of extracted data, this protocol, and study materials such as the full-text manuscript and supplement. Claude will be instructed to extract data only from the documents provided and to validate its own extraction.

3. Claude will extract the data from the study documents provided.

4. A reviewer with expertise in data interpretation and research design, and sufficient knowledge of the disease area of interest will validate 100% of the AI-extracted data.

5. A random proportion of data will be validated by another reviewer for additional quality assurance.

2.5. Risk of Bias

We will examine the risk of bias for each randomized trial in this review using criteria published in the Cochrane risk-of-bias tool for randomized trials (RoB 2).[25,26] Risk of bias will be assessed by study outcome for each of the following domains: bias arising from randomization process, deviation from intended interventions, missing outcome data, measurement of the outcome, selection of the reported results, plus an overall judgment. Two reviewers will independently assess these domains. Any disagreements will be resolved through discussion or by consulting a third reviewer.

2.6. Subgroup Credibility

We will examine the credibility of subgroup analyses (i.e., effect modification analyses) determined to be clinically important for the review using criteria published in the Instrument for the Credibility of Effect Modification ANalyses (ICEMAN) (Version 1.1).[27] The credibility of each effect modifier will be assessed by outcome of interest, time point of interest, and effect measure. One reviewer will assess the credibility of the subgroup analyses and a second reviewer will validate the ratings.

2.7. Clinical Trial Diversity

We will evaluate the demographic diversity for each clinical trial in this review using the ICER-developed Clinical trial Diversity Rating (CDR) Tool.[28] Three demographic categories will be evaluated: race and ethnicity, sex, and age. Representation for each demographic category will be evaluated relative to the disease prevalence, using the metric "Participant to Disease-prevalence Representation Ratio" (PDRR). Each clinical trial will receive a rating of Good, Fair, or Poor for the three demographic categories. One reviewer will use the extracted and validated data from the trials to calculate the ratings. A second reviewer will validate the ratings.

2.8. Publication Bias

Given the emerging nature of the evidence base for new treatments, we will scan the ClinicalTrials.gov site to identify studies completed more than two years ago. Search terms include "treatment-resistant depression," "TRD," and "psilocybin." We will select studies which would have met our inclusion criteria, and for which no findings have been published. We will provide qualitative analysis of the objectives and methods of these studies to ascertain whether there may be a biased representation of study results in the published literature.

2.9. Evidence Synthesis

The purpose of the evidence synthesis is to estimate the clinical effectiveness of the interventions being compared. The analysis will be based on the data from all relevant studies identified from the systematic review. This section contains a summary of the evidence base and synthesis of outcome results.

Summary of Evidence Base

The studies will be summarized in the text and in evidence tables of the Evidence Report. This summary is key to understanding the evidence base pertaining to the topic. An evidence table shell is presented in Appendix C. Relevant data include those listed in the data extraction section. Any key differences between the studies in terms of the study design, patient characteristics,

interventions (including dosing and frequency), outcomes (including definitions and methods of assessments), and risk of bias will be noted in the text of the report.

Synthesis of Results

The results of the studies will be synthesized for each outcome and described narratively in the report. Analyses to be conducted will reflect the nature and quality of the evidence base (see below). Key considerations for interpreting the results will be specified and described in the Evidence Report.

We will assess the feasibility of conducting an NMA under a Bayesian framework. An NMA simultaneously combines both the direct evidence (i.e., estimates obtained from head-to-head comparisons) and indirect evidence (i.e., estimates obtained from common comparator(s)).[29,30] For continuous outcomes (e.g., change from baseline in MADRS score), treatment effects will be estimated as mean differences using a generalized linear model with a normal likelihood and identity link. If trials report Hamilton Depression Rating Scale-17 (HAM-D), instead of MADRS, we will use the mapping formula to convert HAM-D scores to MADRS scores.[21,31] For binary outcomes (e.g., response and remission rates), the NMA model will use a binomial likelihood and logit link. Similar to continuous outcome, if trials report HAM-D, instead of MADRS, following published analyses from Zimmerman and colleagues, we will assume the response rate is the same (≥50% reduction from baseline) and ≤10 on MADRS is assumed to be equal to ≤7 on HAM-D.[31] For all analyses, we will run random-effects models, under the assumption that the amount of between-study variance (i.e., heterogeneity) is constant across all treatment comparisons. Fixed-effect models will also be run, and model fit will be compared using the DIC and residual deviance.

We will use noninformative prior distributions for all model parameters. Each model will use three MCMC chains, with the first 40,000 iterations discarded as "burn-in" and inferences based on an additional 40,000 iterations. Convergence will be assessed with the Gelman-Rubin statistic and visually using trace plots. If the chains do not converge, an additional 10,000 iterations will be run, sequentially, until convergence is achieved.

Furthermore, for any network containing closed loops of evidence, we will empirically compare the direct and indirect estimates using a node-splitting approach to assess if the NMA consistency assumption is violated.[32] If there is evidence of inconsistency, the results will be presented separately for the direct and indirect evidence. If there is no evidence of inconsistency, we will present the pooled results.

All NMAs will be conducted using RStudio. Results for all pairwise comparisons will be presented in tables in terms of point estimates and 95% credible intervals (CrIs). We will also present results for psilocybin versus esketamine, IV ketamine, and rTMS graphically.

All data analysis will be validated by an independent member of the research team. The validator will review and confirm the data analysis methods, data format, and analysis code. The validator will re-run the analysis, validate the results, and confirm the reported data. If the results match within reasonable limits (e.g., small differences may occur due to random sampling variability or model convergence, package versions, etc.), the results are considered validated.

References

1. Hasin DS, Sarvet AL, Meyers JL, et al. Epidemiology of Adult DSM-5 Major Depressive Disorder and Its Specifiers in the United States. *JAMA Psychiatry*. Apr 1 2018;75(4):336-346. doi:10.1001/jamapsychiatry.2017.4602
2. Zhdanava M, Pilon D, Ghelerter I, et al. The Prevalence and National Burden of Treatment-Resistant Depression and Major Depressive Disorder in the United States. *J Clin Psychiatry*. Mar 16 2021;82(2)doi:10.4088/JCP.20m13699
3. Qaseem A, Owens DK, Etxeandia-Ikobaltzeta I, et al. Nonpharmacologic and Pharmacologic Treatments of Adults in the Acute Phase of Major Depressive Disorder: A Living Clinical Guideline From the American College of Physicians. *Ann Intern Med*. Feb 2023;176(2):239-252. doi:10.7326/m22-2056
4. Gabriel FC, Stein AT, Melo DO, et al. Guidelines' recommendations for the treatment-resistant depression: A systematic review of their quality. *PLoS One*. 2023;18(2):e0281501. doi:10.1371/journal.pone.0281501
5. Gaynes BN, Lux L, Gartlehner G, et al. Defining treatment-resistant depression. *Depress Anxiety*. Feb 2020;37(2):134-145. doi:10.1002/da.22968
6. Greenberg PE, Fournier AA, Sisitsky T, et al. The Economic Burden of Adults with Major Depressive Disorder in the United States (2010 and 2018). *Pharmacoeconomics*. Jun 2021;39(6):653-665. doi:10.1007/s40273-021-01019-4
7. Dong M, Zeng L-N, Lu L, et al. Prevalence of suicide attempt in individuals with major depressive disorder: a meta-analysis of observational surveys. *Psychological Medicine*. 2019;49(10):1691-1704. doi:10.1017/S0033291718002301
8. Nordentoft M, Mortensen PB, Pedersen CB. Absolute risk of suicide after first hospital contact in mental disorder. *Arch Gen Psychiatry*. Oct 2011;68(10):1058-64. doi:10.1001/archgenpsychiatry.2011.113
9. Kern DM, Canuso CM, Daly E, et al. Suicide-specific mortality among patients with treatment-resistant major depressive disorder, major depressive disorder with prior suicidal ideation or suicide attempts, or major depressive disorder alone. *Brain Behav*. Aug 2023;13(8):e3171. doi:10.1002/brb3.3171
10. Brown S, Rittenbach K, Cheung S, McKean G, MacMaster FP, Clement F. Current and Common Definitions of Treatment-Resistant Depression: Findings from a Systematic Review and Qualitative Interviews. *The Canadian Journal of Psychiatry*. 2019;64(6):380-387. doi:10.1177/0706743719828965
11. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. *American Journal of Psychiatry*. 2006;163(11):1905-1917. doi:10.1176/ajp.2006.163.11.1905 %M 17074942
12. American Psychological Association. *Clinical practice guideline for the treatment of depression across three age cohorts.* 2019. https://www.apa.org/depression-guideline
13. McQuaid JR, Buelt A, Capaldi V, et al. The Management of Major Depressive Disorder: Synopsis of the 2022 U.S. Department of Veterans Affairs and U.S. Department of Defense Clinical Practice Guideline. *Ann Intern Med*. Oct 2022;175(10):1440-1451. doi:10.7326/m22-1603
14. Wasson RG. Seeking the magic mushroom. Life1957. p. 100-120.
15. Reiff CM, Richman EE, Nemeroff CB, et al. Psychedelics and Psychedelic-Assisted Psychotherapy. *Am J Psychiatry*. May 1 2020;177(5):391-410. doi:10.1176/appi.ajp.2019.19010035

16. Compass Pathways Completes COMP360 Psilocybin New Drug Application to U.S. FDA for Treatment-Resistant Depression. Compass Pathways plc; 2026/11/09, 2026. Accessed 2026/11/24. https://ir.compasspathways.com/News--Events-/news/news-details/2026/Compass-Pathways-Completes-COMP360-Psilocybin-New-Drug-Application-to-U-S--FDA-for-Treatment-Resistant-Depression/default.aspx
17. Compass Pathways Announces FDA Granted NDA Rolling Review Request and Awarded Commissioner's National Priority Voucher. Press release. Compass Pathways plc; April 24, 2026, 2026. Accessed October 15, 2026. https://ir.compasspathways.com/News--Events-/news/news-details/2026/Compass-Pathways-Announces-FDA-Granted-NDA-Rolling-Review-Request-and-Awarded-Commissioners-National-Priority-Voucher/default.aspx
18. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med*. Mar 1 1997;126(5):376-80.
19. Higgins J, Thomas, J, Chandler, J, Cumpston, M, Li, T, Page, MJ, Welch, VA. Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). https://training.cochrane.org/handbook/current
20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. Mar 29 2021;372:n71. doi:10.1136/bmj.n71
21. Saelens J, Gramser A, Watzal V, Zarate CA, Lanzenberger R, Kraus C. Relative effectiveness of antidepressant treatments in treatment-resistant depression: a systematic review and network meta-analysis of randomized controlled trials. *Neuropsychopharmacology*. 2025/05/01 2025;50(6):913-919. doi:10.1038/s41386-024-02044-5
22. Terao I, Tsuge T, Endo K, Kodama W. Comparative efficacy, tolerability and acceptability of intravenous racemic ketamine with intranasal esketamine, aripiprazole and lithium as augmentative treatments for treatment-resistant unipolar depression: A systematic review and network meta-analysis. *J Affect Disord*. Feb 1 2024;346:49-56. doi:10.1016/j.jad.2023.11.023
23. Li H, Cui L, Li J, Liu Y, Chen Y. Comparative efficacy and acceptability of neuromodulation procedures in the treatment of treatment-resistant depression: a network meta-analysis of randomized controlled trials. *Journal of Affective Disorders*. 2021/05/15/ 2021;287:115-124. doi:https://doi.org/10.1016/j.jad.2021.03.019
24. Atlas SJ, Agboola F, Fazioli K, et al. *Esketamine for the Treatment of Treatment-Resistant Depression: Effectiveness and Value. Final Evidence Report*. 2019. 2019/06/20. https://icer.org/wp-content/uploads/2020/10/ICER_TRD_Final_Evidence_Report_062019.pdf
25. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions. 2nd Edition.* Chichester (UK): John Wiley & Sons; 2019.
26. 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
27. Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *Cmaj*. 2020;192(32):E901-E906.
28. Agboola F, Wright A. A Framework for Evaluating the Diversity of Clinical Trials. *Journal of Clinical Epidemiology*. 2024:111299.
29. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. Oct 30 2004;23(20):3105-24. doi:10.1002/sim.1875
30. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ*. Oct 15 2005;331(7521):897-900. doi:10.1136/bmj.331.7521.897
31. Zimmerman M, Posternak MA, Chelminski I. Derivation of a definition of remission on the Montgomery–Asberg depression rating scale corresponding to the definition of remission on the

Hamilton rating scale for depression. *Journal of Psychiatric Research*. 2004/11/01/ 2004;38(6):577-582. doi:https://doi.org/10.1016/j.jpsychires.2004.03.007
32. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med*. Mar 30 2010;29(7-8):932-44. doi:10.1002/sim.3767

Appendix A. PRISMA Checklist

The checklist below is drawn from Page et al. 2021.[20]

Section and Topic	Item #	Checklist Item
TITLE		
Title	1	Identify the report as a systematic review.
ABSTRACT		
Abstract	2	See the PRISMA 2020 for Abstracts checklist.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of existing knowledge.
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.
METHODS		
Eligibility Criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.
Information Sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.
Search Strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.
Selection Process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.
Data Collection Process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.
Data Items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g., for all measures, time points, analyses), and if not, the methods used to decide which results to collect.
	10b	List and define all other variables for which data were sought (e.g., participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.
Study Risk of Bias Assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.
Effect Measures	12	Specify for each outcome the effect measure(s) (e.g., risk ratio, mean difference) used in the synthesis or presentation of results.
Synthesis Methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g., tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.

Section and Topic	Item #	Checklist Item
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.
Reporting Bias Assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).
Certainty Assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.
RESULTS		
Study Selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.
Study Characteristics	17	Cite each included study and present its characteristics.
Risk of Bias in Studies	18	Present assessments of risk of bias for each included study.
Results of Individual Studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.
Results of Syntheses	20a	For each synthesis, briefly summarize the characteristics and risk of bias among contributing studies.
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g., confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.
	20c	Present results of all investigations of possible causes of heterogeneity among study results.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.
Reporting Biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.
Certainty of Evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.
DISCUSSION		
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.
	23b	Discuss any limitations of the evidence included in the review.
	23c	Discuss any limitations of the review processes used.
	23d	Discuss implications of the results for practice, policy, and future research.
OTHER INFORMATION		
Registration and Protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.
	24c	Describe and explain any amendments to information provided at registration or in the protocol.

Section and Topic	Item #	Checklist Item
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.
Competing Interests	26	Declare any competing interests of review authors.
Availability of Data, Code, and Other Materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.

Appendix B. Search Strategies

Table B1. Search Strategy: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE and Versions(R) 1946 to Present, Cochrane Central Register of Controlled Trials, and APA PsycInfo 1967 - 2026 for psilocybin.

1	exp Depressive Disorder, Major/ or exp Depressive Disorder, Treatment-Resistant/
2	((major adj2 depress*) or "depressive disorder*" or "depressive episode*" or "unipolar depress*" or "clinical depression" or ((treatment or therapy) adj2 (resistant or refractory) adj3 depress*) or "refractory depression" or MDD or TRD).ti,ab,kf,kw.
3	1 or 2
4	Psilocybin/
5	(psilocyb* or psilocin* or "magic mushroom*" or "4-PO-DMT" or phosphoryloxy* or comp360 or "comp 360").ti,ab,kf,kw.
6	4 or 5
7	3 and 6
8	exp clinical trials as topic/ or random allocation/ or double-blind method/ or single-blind method/ or cross-over studies/ or placebos/ or control groups/ or ("randomized controlled trial" or "controlled clinical trial" or "clinical trial" or "clinical trial, phase ii" or "clinical trial, phase iii" or "clinical trial, phase iv" or "pragmatic clinical trial" or "equivalence trial").pt.
9	(random* or placebo* or sham or double blind* or single blind* or controlled trial* or clinical trial* or crossover or cross over or active control* or active comparator*).ti,ab. or (control* adj2 (clinical or group* or trial* or study or studies or design* or arm*)).ti,ab.
10	exp cohort studies/ or exp case-control studies/ or ("observational study" or "comparative study").pt. or (cohort* or "case control*" or observational or retrospective* or prospective* or longitudinal or "real world").ti,ab.
11	8 or 9 or 10
12	7 and 11
13	exp animals/ not humans.sh.
14	12 not 13
15	limit 14 to english language
16	15 not (case reports or comment or editorial or news or newspaper article or "published erratum" or review or "systematic review" or "meta-analysis" or "scoping review" or guideline or practice guideline or consensus development conference or historical article or biography or autobiography or bibliography or interview or lecture or address or personal narrative or portrait or patient education handout or video-audio media or webcast or legal case or legislation or periodical index or retraction of publication or expression of concern).pt.
17	remove duplicates from 16

Table B2. Search Strategy: EMBASE for psilocybin

1	depression'/exp OR 'major depression'/exp
2	(depress* OR 'major depress*' OR 'depressive disorder*' OR 'depressive episode*' OR 'unipolar depress*' OR 'clinical depression' OR 'treatment resistant depression' OR 'treatment-resistant depression' OR 'refractory depression' OR mdd OR trd):ti,ab,kw
3	#1 OR #2

4	psilocybin'/exp OR 'psilocin'/exp
5	(psilocyb* OR psilocin OR 'magic mushroom*' OR '4-phosphoryloxy-n,n-dimethyltryptamine' OR '4-po-dmt' OR comp360 OR 'comp 360'):ti,ab,kw
6	#4 OR #5
7	#3 AND #6
8	randomized controlled trial'/exp OR 'controlled clinical trial'/exp OR 'clinical trial'/exp OR 'controlled study'/exp OR 'crossover procedure'/exp OR 'double blind procedure'/exp OR 'placebo'/exp
9	random*:ti,ab OR placebo*:ti,ab OR sham*:ti,ab OR 'double blind*':ti,ab OR 'controlled trial*':ti,ab OR 'clinical trial*':ti,ab OR crossover:ti,ab OR 'cross over':ti,ab OR 'active control*':ti,ab OR 'active comparator*':ti,ab
10	quasi experimental study'/exp OR 'observational study'/exp OR 'cohort analysis'/exp OR 'case control study'/exp OR 'comparative study'/exp OR 'retrospective study'/exp OR 'prospective study'/exp OR 'longitudinal study'/exp
11	(nonrandomi* OR 'non randomi*' OR 'quasi experiment*' OR 'quasi random*' OR 'open label' OR cohort* OR 'case control*' OR observational OR retrospective* OR prospective* OR longitudinal OR 'real world' OR 'comparative stud*'):ti,ab
12	#8 OR #9 OR #10 OR #11
13	#7 AND #12
14	(animal/exp OR nonhuman/exp OR 'animal experiment'/exp) NOT human/exp
15	(mouse:ti,ab OR mice:ti,ab OR rat:ti,ab OR rats:ti,ab OR rodent*:ti,ab OR murine:ti,ab OR zebrafish:ti,ab OR 'in vitro':ti,ab OR preclinical:ti,ab OR 'pre-clinical':ti,ab OR ipsc:ti,ab OR docking:ti,ab OR 'cell line*':ti,ab)
16	#14 OR #15
17	#13 NOT #16
18	#17 NOT ('case report'/de OR 'practice guideline'/de OR 'chapter'/it OR 'conference review'/it OR 'editorial'/it OR 'erratum'/it OR 'letter'/it OR 'note'/it OR 'review'/it OR 'short survey'/it)
19	#18 NOT (survey:ti OR qualitative:ti OR interview*:ti OR 'focus group*':ti OR 'thematic analysis':ti OR attitude*:ti OR perception*:ti)
20	#19 AND [english]/lim
21	#20 AND [medline]/lim
22	#20 NOT #21

Table B3. Search Strategy: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE and Versions(R) 1946 to Present, Cochrane Central Register of Controlled Trials, and APA PsycInfo 1967 - 2026 for comparators. (Updated since 2023 [last date of search from network meta-analysis published by Saelens and colleagues[21]])

1	exp depressive disorder, treatment-resistant/
2	exp treatment resistant depression/
3	exp depressive disorder, major/ or exp major depression/ or depress*.ti,ab.
4	(treatment-resist* or treatment resist* or therapy-resist* or therapy resist* or refract* or resist* or intractable or nonrespon* or non-respon* or unrespon* or fail* or trd or difficult to treat or difficult-to-treat).ti,ab.
5	treatment failure/
6	4 or 5

7	3 and 6
8	1 or 2 or 7
9	(esketamine or S-ketamine or S ketamine or Ketanest or Ketanest S or spravato or JNJ 54135419 or JNJ-54135419).ti,ab.
10	ketamine/
11	(ketamin* or Ketaset or Ketalar or CI-581 or CI581 or CI 581 or Calipsol or Kalipsol or Calypsol).ti,ab.
12	transcranial magnetic stimulation/
13	(TMS or rTMS or transcranial magnetic* or theta burst* or iTBS or cTBS or dTMS).ti,ab.
14	9 or 10 or 11 or 12 or 13
15	8 and 14
16	exp clinical trials as topic/ or random allocation/ or double-blind method/ or single-blind method/ or cross-over studies/ or placebos/ or control groups/ or ("randomized controlled trial" or "controlled clinical trial" or "clinical trial" or "clinical trial, phase ii" or "clinical trial, phase iii" or "clinical trial, phase iv" or "pragmatic clinical trial" or "equivalence trial").pt.
17	(random* or placebo* or sham or double blind* or single blind* or controlled trial* or clinical trial* or crossover or cross over or active control* or active comparator*).ti,ab. or (control* adj2 (clinical or group* or trial* or study or studies or design* or arm*)).ti,ab.
18	exp cohort studies/ or exp case-control studies/ or ("observational study" or "comparative study").pt. or (cohort* or "case control*" or observational or retrospective* or prospective* or longitudinal or "real world").ti,ab.
19	16 or 17 or 18
20	15 and 19
21	limit 20 to yr="2023-Current"
22	exp animals/ not humans.sh.
23	21 not 22
24	limit 23 to english language
25	24 not (case reports or comment or editorial or news or newspaper article or "published erratum" or review or "systematic review" or "meta-analysis" or "scoping review" or guideline or practice guideline or consensus development conference or historical article or biography or autobiography or bibliography or interview or lecture or address or personal narrative or portrait or patient education handout or video-audio media or webcast or legal case or legislation or periodical index or retraction of publication or expression of concern).pt.
27	remove duplicates from 25

Table B4. Search Strategy: EMBASE for comparators (Updated since 2023 [last date of search from network meta-analysis published by Saelens and colleagues[21]])

1	treatment resistant depression'/exp
2	major depression'/exp OR depress*:ti,ab
3	treatment resist*':ti,ab OR 'treatment-resist*':ti,ab OR 'therapy resist*':ti,ab OR 'therapy-resist*':ti,ab OR refract*:ti,ab OR resist*:ti,ab OR intractable:ti,ab OR nonrespon*:ti,ab OR 'non-respon*':ti,ab OR unrespon*:ti,ab OR fail*:ti,ab OR trd:ti,ab OR 'difficult to treat':ti,ab OR 'difficult-to-treat':ti,ab
4	treatment failure'/exp
5	#3 OR #4
6	#2 AND #5

7	#1 OR #6
8	esketamine'/exp
9	esketamine:ti,ab OR 's ketamine':ti,ab OR 's-ketamine':ti,ab OR ketanest:ti,ab OR spravato:ti,ab OR 'jnj 54135419':ti,ab OR 'jnj-54135419':ti,ab
10	ketamine'/exp
11	ketamin*:ti,ab OR ketaset:ti,ab OR ketalar:ti,ab OR ci581:ti,ab OR 'ci 581':ti,ab OR calipsol:ti,ab OR kalipsol:ti,ab OR calypsol:ti,ab
12	transcranial magnetic stimulation'/exp
13	tms:ti,ab OR rtms:ti,ab OR 'transcranial magnetic*':ti,ab OR 'theta burst*':ti,ab OR itbs:ti,ab OR ctbs:ti,ab OR dtms:ti,ab
14	#8 OR #9 OR #10 OR #11 OR #12 OR #13
15	#7 AND #14
16	randomized controlled trial'/exp OR 'controlled clinical trial'/exp OR 'clinical trial'/exp OR 'controlled study'/exp OR 'crossover procedure'/exp OR 'double blind procedure'/exp OR 'placebo'/exp
17	random*:ti,ab OR placebo*:ti,ab OR sham*:ti,ab OR 'double blind*':ti,ab OR 'controlled trial*':ti,ab OR 'clinical trial*':ti,ab OR crossover:ti,ab OR 'cross over':ti,ab OR 'active control*':ti,ab OR 'active comparator*':ti,ab
18	quasi experimental study'/exp OR 'observational study'/exp OR 'cohort analysis'/exp OR 'case control study'/exp OR 'comparative study'/exp OR 'retrospective study'/exp OR 'prospective study'/exp OR 'longitudinal study'/exp
19	(nonrandomi* OR 'non randomi*' OR 'quasi experiment*' OR 'quasi random*' OR 'open label' OR cohort* OR 'case control*' OR observational OR retrospective* OR prospective* OR longitudinal OR 'real world' OR 'comparative stud*'):ti,ab
20	#16 OR #17 OR #18 OR #19
21	#15 AND #20
22	#21 AND [2023-2026]/py
23	('animal'/exp OR 'nonhuman'/exp OR 'animal experiment'/exp) NOT 'human'/exp
24	(mouse:ti,ab OR mice:ti,ab OR rat:ti,ab OR rats:ti,ab OR rodent*:ti,ab OR murine:ti,ab OR zebrafish:ti,ab OR 'in vitro':ti,ab OR preclinical:ti,ab OR 'pre-clinical':ti,ab OR ipsc:ti,ab OR docking:ti,ab OR 'cell line*':ti,ab)
25	#23 OR #24
26	#22 NOT #25
27	#26 AND [english]/lim
28	#27 AND [medline]/lim
29	#27 NOT #28
30	#29 NOT ('case report'/de OR 'human tissue'/de OR 'practice guideline'/de OR 'chapter'/it OR 'conference review'/it OR 'editorial'/it OR 'letter'/it OR 'note'/it OR 'review'/it OR 'short survey'/it)
31	#30 NOT (survey:ti OR qualitative:ti OR interview*:ti OR 'focus group*':ti OR 'thematic analysis':ti OR attitude*:ti OR perception*:ti)

Table B5. Search Strategy: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE and Versions(R) 1946 to Present, and APA PsycInfo 1967 - 2026 for comparators. (Updated since 2019 [last date of search from ICER 2019 report on Spravato for TRD][24])

1	exp depressive disorder, treatment-resistant/
2	exp treatment resistant depression/
3	exp depressive disorder, major/ or exp major depression/ or depress*.ti,ab.
4	(treatment-resist* or treatment resist* or therapy-resist* or therapy resist* or refract* or resist* or intractable or nonrespon* or non-respon* or unrespon* or fail* or trd or difficult to treat or difficult-to-treat).ti,ab.
5	treatment failure/
6	4 or 5
7	3 and 6
8	1 or 2 or 7
9	(esketamine or S-ketamine or S ketamine or Ketanest or Ketanest S or spravato or JNJ 54135419 or JNJ-54135419).ti,ab.
10	ketamine/
11	(ketamin* or Ketaset or Ketalar or CI-581 or CI581 or CI 581 or Calipsol or Kalipsol or Calypsol).ti,ab.
12	transcranial magnetic stimulation/
13	(TMS or rTMS or transcranial magnetic* or theta burst* or iTBS or cTBS or dTMS).ti,ab.
14	9 or 10 or 11 or 12 or 13
15	8 and 14
16	exp cohort studies/ or exp case-control studies/ or ("observational study" or "comparative study").pt. or (cohort* or "case control*" or observational or retrospective* or prospective* or longitudinal or "real world").ti,ab.
17	15 and 16
18	limit 17 to yr="2019-Current"
19	exp animals/ not humans.sh.
20	18 not 19
21	limit 20 to english language
22	21 not (case reports or comment or editorial or news or newspaper article or "published erratum" or review or "systematic review" or "meta-analysis" or "scoping review" or guideline or practice guideline or consensus development conference or historical article or biography or autobiography or bibliography or interview or lecture or address or personal narrative or portrait or patient education handout or video-audio media or webcast or legal case or legislation or periodical index or retraction of publication or expression of concern).pt.
23	remove duplicates from 22

Table B6. Search Strategy: EMBASE for comparators (Updated since 2019 [last date of search from ICER 2019 report on Spravato for TRD][24])

1	treatment resistant depression'/exp
2	major depression'/exp OR depress*:ti,ab
3	treatment resist*':ti,ab OR 'treatment-resist*':ti,ab OR 'therapy resist*':ti,ab OR 'therapy-resist*':ti,ab OR refract*:ti,ab OR resist*:ti,ab OR intractable:ti,ab OR nonrespon*:ti,ab OR 'non-respon*':ti,ab OR unrespon*:ti,ab OR fail*:ti,ab OR trd:ti,ab OR 'difficult to treat':ti,ab OR 'difficult-to-treat':ti,ab
4	treatment failure'/exp
5	#3 OR #4
6	#2 AND #5
7	#1 OR #6
8	esketamine'/exp
9	esketamine:ti,ab OR 's ketamine':ti,ab OR 's-ketamine':ti,ab OR ketanest:ti,ab OR spravato:ti,ab OR 'jnj 54135419':ti,ab OR 'jnj-54135419':ti,ab
10	ketamine'/exp
11	ketamin*:ti,ab OR ketaset:ti,ab OR ketalar:ti,ab OR ci581:ti,ab OR 'ci 581':ti,ab OR calipsol:ti,ab OR kalipsol:ti,ab OR calypsol:ti,ab
12	transcranial magnetic stimulation'/exp
13	tms:ti,ab OR rtms:ti,ab OR 'transcranial magnetic*':ti,ab OR 'theta burst*':ti,ab OR itbs:ti,ab OR ctbs:ti,ab OR dtms:ti,ab
14	#8 OR #9 OR #10 OR #11 OR #12 OR #13
15	#7 AND #14
16	quasi experimental study'/exp OR 'observational study'/exp OR 'cohort analysis'/exp OR 'case control study'/exp OR 'comparative study'/exp OR 'retrospective study'/exp OR 'prospective study'/exp OR 'longitudinal study'/exp
17	(nonrandomi* OR 'non randomi*' OR 'quasi experiment*' OR 'quasi random*' OR 'open label' OR cohort* OR 'case control*' OR observational OR retrospective* OR prospective* OR longitudinal OR 'real world' OR 'comparative stud*'):ti,ab
18	#16 OR #17
19	#15 AND #18
20	('animal'/exp OR 'nonhuman'/exp OR 'animal experiment'/exp) NOT 'human'/exp
21	(mouse:ti,ab OR mice:ti,ab OR rat:ti,ab OR rats:ti,ab OR rodent*:ti,ab OR murine:ti,ab OR zebrafish:ti,ab OR 'in vitro':ti,ab OR preclinical:ti,ab OR 'pre-clinical':ti,ab OR ipsc:ti,ab OR docking:ti,ab OR 'cell line*':ti,ab)
22	#20 OR #21
23	#19 NOT #22
24	#23 AND [english]/lim
25	#24 AND [medline]/lim
26	#24 NOT #25
27	#26 NOT ('case report'/de OR 'human tissue'/de OR 'practice guideline'/de OR 'chapter'/it OR 'conference review'/it OR 'editorial'/it OR 'letter'/it OR 'note'/it OR 'review'/it OR 'short survey'/it)
28	#27 NOT (survey:ti OR qualitative:ti OR interview*:ti OR 'focus group*':ti OR 'thematic analysis':ti OR attitude*:ti OR perception*:ti)
29	#28 AND [2019-2026]/py

Appendix C. Data Extraction Summary Table Shell

Table C1. Study Design

Author & Year of Publication (Trial)	Study Design	Interventions (n) & Dosing Schedule	Inclusion & Exclusion Criteria	Patient Characteristics	Outcomes

Table Footnotes