JAMA Psychiatry | Original Investigation

Psychedelic Therapy vs Antidepressants for the Treatment of Depression Under Equal Unblinding Conditions
A Systematic Review and Meta-Analysis

Zachary J. Williams, MD, PhD; Hannah Barnett, MSc; Balázs Szigeti, PhD

IMPORTANCE Psychedelic-assisted therapy (PAT) trials have high levels of functional unblinding, which biases results when comparing PAT with blinded interventions. Because PAT is effectively always open label, treatment results should be compared with those of open-label traditional antidepressants (TADs), so potential benefits associated with patients knowing their treatment is equal between the interventions.

OBJECTIVE To investigate the comparative effectiveness of PAT vs open-label traditional antidepressants (TADs; such as selective serotonin and norepinephrine reuptake inhibitors) for the treatment of major depression.

DATA SOURCES PubMed was systematically searched in March 2024 for trials of PAT and open-label TADs for the treatment of major depression without comorbidity in adults without psychosis in the outpatient setting. Extraction was supplemented with data from a review and meta-analysis of antidepressant drugs to assess the open-label vs blinded TAD difference.

DATA EXTRACTION AND SYNTHESIS Depression scores were extracted by 2 independent reviewers; estimates were pooled with both bayesian and frequentist mixed-effects models. Reporting follows the PRISMA guideline.

MAIN OUTCOMES AND MEASURE Following predefined hypotheses, the mean within-arm effect from baseline to primary end point (ie, patient improvement between PAT and open-label TAD trials on the 17-item Hamilton Depression Rating Scale) was compared. To assess the potential role of blinding, the within-arm effect of blinded vs open-label trials in both PAT and TADs was also compared.

RESULTS Of the initially retrieved PubMed 619 records, 24 met inclusion criteria. Contrary to the first of 3 hypotheses, PAT (8 trials; 249 patients) was no more effective than open-label TAD treatment (16 open-label TAD trials; 7921 patients), with an estimated difference of 0.3 favoring open-label TADs (95% CI, −1.39 to 1.98; $P = .73$). Open-label TADs were associated with better outcomes than blinded treatment (144 blinded TAD trials; 31 792 patients), with an estimated difference of 1.3 (95% CI, 0.07-2.51; $P = .04$;), but the same difference was not observed for PAT (0.67; 95% CI, −3.08 to 1.73; $P = .58$).

CONCLUSIONS AND RELEVANCE In trials of depression, PAT was not more effective than open-label TADs. Blinding made a difference for TADs, but not for PAT, confirming that PAT trials are effectively always open label. These results argue against highly optimistic narratives surrounding PAT and highlight the importance of blinding integrity.

Multimedia

Related article page 448

Supplemental content

Author Affiliations: Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles (Williams); Frist Center for Autism and Innovation, Vanderbilt University School of Engineering, Nashville, Tennessee (Williams); A.J. Drexel Autism Institute, Drexel University, Philadelphia, Pennsylvania (Williams); Centre for Psychedelic Research, Imperial College London, London, United Kingdom (Barnett, Szigeti); Department of Psychiatry and Behavioral Sciences, Weill Institute for Neurosciences, University of California, San Francisco (Szigeti).

Corresponding Author: Balázs Szigeti, PhD, Department of Psychiatry and Behavioral Sciences, Weill Institute for Neurosciences, University of California, San Francisco, 675 18th St, San Francisco, CA 94107 (balazs.szigeti@ucsf.edu).

JAMA Psychiatry. 2026;83(5):461-468. doi:10.1001/jamapsychiatry.2025.4809
Published online March 18, 2026.

Current treatments for depression include selective serotonin (SSRIs) and norepinephrine reuptake inhibitors (SNRIs) and a few other medications, such as mirtazapine. These are collectively referred to as traditional antidepressants (TADs). A comprehensive meta-analysis found a significant TAD vs placebo difference (also known as specific-treatment or between-arm effect), but the magnitude of the difference was approximately 2.4 units on the 17-item Hamilton Depression Rating Scale (HAM-D).[1,2] This small difference raised questions around whether TADs provide a clinically meaningful benefit over placebos.[3] Psychedelic-assisted therapy (PAT), which is the combined treatment of psychotherapy and psychedelics, has emerged as a novel depression treatment,[4] attracting much attention.[5] Unlike TAD studies, PAT studies have reported larger PAT vs placebo differences, with a mean effect of approximately 7.3 HAM-D units.[6-8]

In open-label trials, patients are aware of their treatment; in contrast, blinded trials conceal treatment allocation. Even when trials are formally blinded, patients can sometimes deduce their treatment due to adverse effects and/or other factors. This phenomenon is called functional unblinding and may inflate the drug-placebo difference.[9-11] Functional unblinding is exceedingly common in psychedelic trials due to the drugs' intense subjective effects. In a blinded trial, approximately 50% of patients can correctly guess their treatment allocation (assuming 2 arms and 1:1 allocation); in PAT trials, the correct guess rate is 90% to 95%,[12,13] even when an active placebo is used.[14] Functional unblinding played a prominent role in the US Food and Drug Administration's rejection of 3,4-methylenedioxymethamphetamine–assisted psychotherapy as a treatment for posttraumatic stress disorder.[15] Researchers continue to debate whether blinding is maintained in TAD trials. Two recent meta-analyses arrived at opposite conclusions.[16,17] However, the correct guess rate in blinded TAD trials is 63%, which is substantially lower than in PAT trials. Therefore, even if unblinding is present in TAD trials, its magnitude is much smaller than in trials of PAT.

To date, only 1 PAT vs TAD (psilocybin vs escitalopram) head-to-head trial has been conducted.[4] At the primary end point, there was no between-treatment difference in the primary outcome; however, PAT showed significantly better improvement in all secondary depression measures. This discrepancy between the trial's outcomes has fueled debate about the trial's methodologic limitations and the interpretation of its results.[18] The present meta-analysis compares PAT with open-label TADs, so potential benefit from effects associated with patients knowing their treatment is equal between interventions.

Key Points

Question How does psychedelic-assisted therapy (PAT) compare with open-label traditional antidepressants (TADs) in the treatment of major depression?

Findings This meta-analysis of 24 trials found no statistically significant difference in patient improvement following PAT or open-label TADs. Open-label trials were associated with slightly greater patient improvement for antidepressants, but the same difference was not observed for psychedelics.

Meaning These findings suggest that PAT is no more effective than TADs under equal-unblinding conditions for the treatment of depression and highlight the potential role of blinding integrity.

Methods

Preregistration and Hypothesis

We preregistered both the search process and the statistical analysis plan of the meta-analysis on a registration platform (AsPredicted).[19] Minor deviations from this plan were made, which are described in the eMethods in Supplement 1. This meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guideline. Previous work found that a difference of 3 to 5 HAM-D units corresponds to the minimal clinically important difference (MCID)[20]; we used the lower bound of this estimate (ie, MCID, 3 HAM-D units). The following hypotheses were registered: at the primary end point, the estimated mean difference between PAT and open-label TAD trials will exceed the MCID, favoring PAT (H1); at the primary end point, the estimated mean difference between blinded and open-label TAD trials will exceed the MCID, favoring open-label administration (H2) (to estimate the effectiveness of blinded TAD treatment, we used data from Cipriani et al[1]); and at the primary end point, the estimated mean difference between blinded and open-label PAT trials will not exceed the MCID (H3) (ie, open-label treatment is not superior to blinded treatment).

Search Strategy and Data Extraction

We searched the PubMed database to identify trials of major depressive disorder in the adult population (mean age, >18 and <65 years) where the treatment was either open label with commonly prescribed antidepressants as listed by Cipriani et al[1] (ie, TADs) or blinded or open-label PAT with 1 of the following psychedelics: lysergic acid diethylamide; psilocybin; mescaline, San Pedro, or peyote; or 5-methoxy-N,N-dimethyltruptamine or ayahuasca.

Trials were excluded if they were conducted in the inpatient setting or included patients with psychotic depression or substantial comorbidity. An exception was made for comorbid anxiety, due to its frequent co-occurrence with depression.[21] Augmentation and combination trials were excluded. Trials with run-in periods were excluded; however, data from the run-in period were included as open-label TADs if otherwise the phase met all inclusion criteria (eg, the open-label phase of a discontinuation trial).

Retrieved articles were independently marked for inclusion or exclusion by B.S. and H.B., and then depression scores were extracted. References of the selected articles were scanned for additional trials. When data were missing, we attempted recovery by contacting study authors.

Data Synthesis

All depression scores were transformed to HAM-D equivalents[22-24]; all HAM-D scores refer to the 17-item version of the scale. We converted between the SDs of end point

and change scores using a conservative correlation coefficient of 0.5.[25]

Statistical Analysis

We analyzed the data using preregistered bayesian and frequentist models. The primary effect of interest was the within-arm effect (ie, the change from baseline to primary end point in HAM-D units). These within-arm effects were synthesized using meta-analytic models with a multilevel random-effects structure[26] that accounted for the multiple outcome measures (eg, Montgomery-Åsberg Depression Rating Scale and Quick Inventory of Depressive Symptomology). All models included the same random-effects structure, a fixed effect (and random slope) of baseline depression severity, and the focal variable.

Models assessing H1 had the binary focal variable of treatment type (PAT vs TADs). Models assessing H2 compared open-label vs blinded TAD studies, and models assessing H3 compared open-label vs blinded PAT studies. These models had the binary focal variable of blinding (open label vs blinded). Heterogeneity was quantified from the bayesian models using the intraclass correlation coefficient (ICC; mean and 95% credible interval [CrI]) with 10 000 bootstrapped iterations.

Bayesian model results are reported with posterior distributions of the focal variable,[27] summarized by the posterior median and 95% CrI. Posterior predictive distributions are shown in eFigure 4 in Supplement 1 to illustrate between-study heterogeneity. Frequentist results are reported with the mean and 95% CI of the focal variable. Secondary analyses were performed using frequentist models due to their lower computational demand. All *P* values are 2 sided, with significance set at .05. Detailed model outputs for all hypotheses are provided in eTable 1 in Supplement 1.

Results

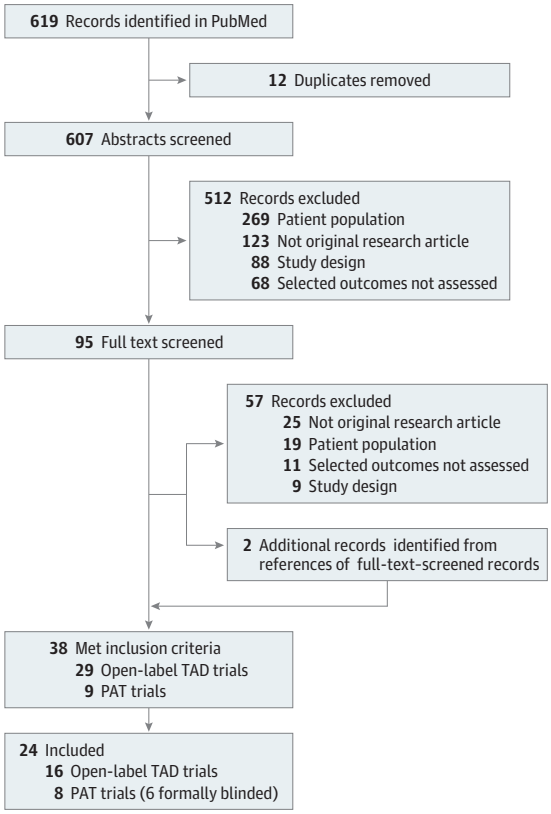
Characteristics of Included Trials

The search was conducted in March 2024. Of the 619 retrieved records, 38 met the inclusion criteria. Of these 38, the required variables could be extracted from 24 trials, including 16 open-label TAD[28-43] (7921 patients) and 8 PAT trials[4,6,8,44-48] (249 patients) (**Figure 1**). Among the 8 PAT trials, 6 were formally blinded (213 patients) and 2 were open label (36 patients). The mean (SD) time from baseline to the primary end point was 8.1 (1.5) weeks for TADs and 3.4 (2.2) weeks for PAT. At baseline, the mean (SD) HAM-D score was 22.7 (1.9) for TADs and 21.3 (3.9) for PAT.

We identified 4 additional PAT studies[7,49-51] that nearly met inclusion criteria: 3 in which cancer comorbidity and 1 in which both major depression and bipolar II disorder were present. Of these 4 additional trials, 2 were formally blinded (40 patients) and 2 were open label (28 patients). We included these trials in an extended dataset, which was used to test robustness of results.

H2, blind vs open-label TADs, used data from Cipriani et al[1] to quantify the within-arm effect of blind TAD treatment. We could extract the data required for our analysis from 144 trials, involving 31 792 patients; eMethods in Supplement 1 provide additional details.

Figure 1. Flowchart of Study Identification



During the screening process, multiple reasons could be tagged for exclusion, and these were not consolidated.

Bayesian Model of H1 (Open-Label TADs vs PAT)

The bayesian estimate of the within-arm effect of open-label TAD treatment was −12.5 HAM-D units (95% CRI, −12.9 to −12.2; standardized mean difference [SMD], −2.7), while for PAT it was −11.8 (95% CRI, −13.3 to −10.3; SMD, −2.6). The posterior distribution of the PAT-TAD difference had a mean of β = 0.25 (95% CRI, −1.90 to 2.45; ICC, 0.361; 95% CRI, 0.345-0.376), favoring TADs. The posterior probability that PAT decreased depression by 3 or more HAM-D units over open-label TADs (ie, probability that H1 is true) was 0.2%. The posterior probability of the difference falling within plus or minus 3 HAM-D units (ie, the region of practical equivalence[52]) was 99.1% (**Figure 2**). When using the extended dataset, the difference was similar (β = 0.43; 95% CRI, −0.17 to 2.55) (eFigure 1 in Supplement 1). These findings provided strong evidence against H1.

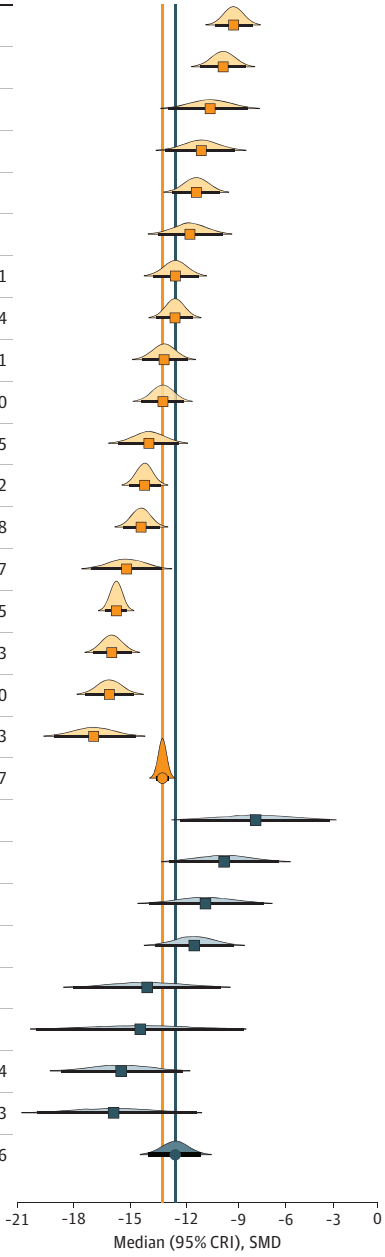
Frequentist Model of H1 (Open-Label TADs vs PAT)

In line with the bayesian analysis, the frequentist model did not find a significant difference between PAT and open-label TADs (estimated mean difference, 0.3; 95% CI, −1.39 to 1.98; *P* = .73), failing to confirm H1. The results did not change qualitatively when the frequentist model for H1 was applied to the

Figure 2. Forest Plot of Difference Between Psychedelic-Assisted Therapy (PAT) and Open-Label Traditional Antidepressant (TAD) Treatment in Bayesian Model

A PAT vs TAD Within-arm effect sizes (primary end point minus baseline)

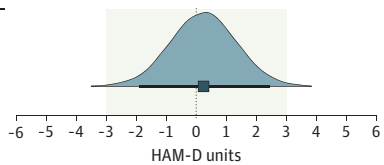
Study	Median (95% CRI), SMD
Souery et al,[38] 2015 (venlafaxine)	-8.4 (-9.5 to -7.3), -1.6
Kennedy et al,[34] 2019 (escitalopram)	-9.0 (-10.3 to -7.7), -1.8
Gargoloff et al,[31] 2016 (agomelatine)	-9.8 (-12.2 to -7.6), -1.4
Thase et al,[43] 1997 (fluoxetine)	-10.3 (-12.4 to -8.3), -2.2
Lesser et al,[35] 2010 (citalopram)	-10.6 (-12.0 to -9.2), -1.5
Husain et al,[33] 2023 (escitalopram)	-10.9 (-12.8 to -9.0), -2.4
Arns et al,[28] 2016 (venlafaxine)	-11.8 (-13.1 to -10.5), -3.1
Trivedi et al,[42] 2013 (escitalopram)	-11.8 (-12.9 to -10.7), -3.4
Arns et al,[28] 2016 (escitalopram)	-12.4 (-13.7 to -11.1), -3.1
Arns et al,[28] 2016 (sertraline)	-12.5 (-13.8 to -11.3), -3.0
Trivedi et al,[41] 2006 (citalopram)	-13.3 (-15.1 to -11.6), -2.5
Chokka and Legault,[30] (escitalopram)	-13.6 (-14.5 to -12.7), -3.2
Hudson et al,[32] 2007 (duloxetine)	-13.8 (-14.8 to -12.7), -3.8
Pecenak and Novotny,[36] 2013 (agomelatine)	-14.6 (-16.7 to -12.6), -3.7
Cervera-Enguix et al,[29] (venlafaxine)	-15.2 (-15.9 to -14.6), -2.5
Thase et al,[40] 2001 (mirtazapine)	-15.5 (-16.6 to -14.3), -4.3
Taragano et al,[39] 1990 (sertraline)	-15.6 (-17.0 to -14.2), -2.0
Pinto et al,[37] 2007 (escitalopram)	-16.6 (-18.9 to -14.1), -3.3
Antidepressant summary effect	-12.5 (-12.9 to -12.2), -2.7
Sloshower et al,[48] 2023 (psilocybin)	-7.1 (-11.5 to -2.8), -0.9
von Rotz et al,[8] 2023 (psilocybin)	-9.0 (-12.1 to -5.8), -2.0
Goodwin et al,[46] 2022 (psilocybin)	-10.0 (-13.3 to -6.7), -2.3
Carhart-Harris et al,[4] 2021 (psilocybin)	-10.7 (-13.0 to -8.4), -2.7
Davis et al,[45] 2021 (psilocybin)	-13.4 (-17.8 to -9.2), -3.4
Palhano-Fontes et al,[47] 2019 (ayahuasca)	-13.8 (-19.9 to -7.8), -2.6
Carhart-Harris et al,[44] 2016 (psilocybin)	-15.0 (-18.4 to -11.3), -3.4
Raison et al,[6] 2023 (psilocybin)	-15.4 (-19.9 to -10.6), -3.3
Psychedelic summary effect	-11.8 (-13.3 to -10.3), -2.6



B Posterior density of PAT minus TAD difference

Mean difference (95% CI)

0.25 (-1.90 to 2.45)



A, Primary end point minus baseline within-arm effect sizes are shown as model-based posterior densities, with each study summarized using its posterior median (orange and dark blue squares), 95% highest-density credible interval (CrI; horizontal bar), and standardized mean difference (SMD; scaled by baseline SD). Pooled within-arm summary effects for TAD and PAT conditions are also depicted, with point estimates represented by vertical colored lines. B, Posterior distribution curve of between-treatment difference, with negative values favoring PAT (ie, larger decrease in symptoms). The probability that hypothesis 1 is true (ie, PAT − TADs < −3) is 0.002. Shaded area spans region of practical equivalence (ie, minimum clinically important difference of 3 Hamilton Depression Rating Scale [HAM-D] units in both directions). TAD trials had larger sample sizes (total of 8111 patients in TAD and 249 in PAT trials), resulting in narrower CrI estimates.

following variations: extended dataset of PAT vs open-label TADs (β = 0.24; 95% CI, −1.39 to 1.87; *P* = .77); PAT vs open-label TADs for non-treatment-resistant depression (TRD; β = 0.64; 95% CI, −1.15 to 2.43; *P* = .48); PAT vs open-label TADs, but only using HAM-D data (ie, the preregistered analysis, but without the conversion between depression scales; β = 0.64; 95% CI, −2.05 to 3.33; *P* = .64); PAT vs open-label SSRI (citalopram, escitalopram, fluoxetine, and sertraline) trials (β = 1.02; 95% CI, −0.81 to 2.86; *P* = .28; we also tested PAT vs each SSRI separately and found no significant difference between PAT and any SSRI); and PAT vs open-label SNRI (duloxetine and venlafaxine) trials (β = −0.36; 95% CI, −3.32 to 2.6; *P* = .81; we also tested PAT vs each SNRI separately and found no significant difference between PAT and any SNRI).

Bayesian Model of H2 (Blinded vs Open-Label TADs)

The bayesian model estimated that the within-arm effect of TADs with blinded administration was −12.3 HAM-D units (95% CRI, −12.5 to −12.1; SMD, −2.6), while with open-label administration it was −12.7 (95% CRI, −13.1 to −12.3; SMD, −2.8). The posterior distribution of the blinded vs open-label TAD difference had a mean of β = 0.85 (95% CRI, −1.09 to 2.57; ICC, 0.199; 95% CrI, 0.187-0.211), favoring open label. The posterior probability that open-label TADs decreased depression by 3 or more HAM-D units over blinded treatment (ie, probability that H2 is true) was 1.4%. The posterior probability of the difference falling within plus or minus 3 HAM-D units was 98.4% (eFigure 2 in Supplement 1). These findings provided evidence against H2.

Frequentist Model of H2 (Blinded vs Open-Label TADs)

In the frequentist model, there was a statistically significant difference between blinded and open-label TAD treatment. The estimated mean difference was β = 1.29 HAM-D units (95% CI, 0.07-2.51; *P* = .04), favoring open-label administration and supporting H2. However, this effect corresponds to approximately half of the MCID, and the entire 95% CI falls within the region of plus or minus 3 HAM-D units, meaning that the magnitude of this effect is practically negligible.[53]

Bayesian Model of H3 (Blinded vs Open-Label PAT)

The bayesian model estimated that PAT's within-arm effect with blinded administration was −10.8 HAM-D units (95% CRI, −12.9 to −8.8; SMD, −2.2), while with open-label administration it was −13.3 (95% CRI, −15.4 to −11.2; SMD, −3.2). The posterior distribution of the blinded-open-label PAT difference had a mean of β = 2.14 (95% CRI, −1.86 to 5.79; ICC, 0.422; 95% CRI, 0.411-0.432) HAM-D units, favoring open label. The posterior probability that the difference is within plus or minus 3 HAM-D units (ie, probability that H3 is true) was 68.0%. When using the extended dataset, the difference was smaller (β = 0.08; 95% CRI, −3.38 to 3.57). These results provided evidence for H3 (eFigure 3 in Supplement 1).

Frequentist Model of H3 (Blinded vs Open-Label PAT)

In concordance with the bayesian analysis, the frequentist model did not find a significant difference between blinded and open-label PAT (estimated mean difference β = 0.67; 95% CI, −3.08 to 1.73; *P* = .58), confirming H3. The result did not change qualitatively when the same model was applied to the extended dataset (β = 0.09; 95% CI, −2.13 to 2.32; *P* = .93).

Discussion

The premise of this preregistered meta-analysis was that comparison of an open-label with a blind intervention would be biased, as the open-label intervention would unfairly benefit from effects associated with patients knowing their treatment. Therefore, we compared the effectiveness of PAT, which is effectively always open label, with that of open-label TADs for the treatment of major depression.

In contrast with H1, we did not find treatment with PAT to be better than treatment with open-label TADs by a clinically meaningful margin. Not only was the difference not clinically meaningful, but there was practically no difference at all; according to both bayesian and frequentist estimates, the difference was negligible at approximately 0.3 HAM-D units. This result was robust across variations in study selection, including when we removed PAT trials of TRD.

We also assessed the impact of blinding in both PAT and TAD trials. We found that for TADs (H2), but not for PAT (H3), open-label administration was associated with better outcomes than blinded treatment. This finding is in line with our premise that TADs can be effectively blinded, and thus blinding should lead to a difference, whereas PAT is effectively always open label, so formal blinding should make no difference. The estimate of the open-label vs blinded difference for TADs was 0.85 HAM-D units in the bayesian model and 1.29 HAM-D units in the frequentist model, approximately half of the MCID. Thus, both models for H2 indicated that the effect of blinding was not clinically meaningful.

PAT trials reported a 7.3-HAM-D unit difference from placebo,[6-8,49] while TAD trials reported a difference of 2.4.[1] Thus, PAT is 4.9 (ie, 7.3–2.4) or approximately 5 points better than TADs when measured against placebo. Two key factors explain the failure of H1 despite the 5-point difference of between-arm differences. (1) Open-label TADs are approximately 1.29 HAM-D units more effective than blinded treatment (H2). This effect can be interpreted as the influence of a patient knowing their treatment assignment, including positive expectancy. (2) Recent meta-analyses of depression trials found that the within-arm effect size of the control conditions in PAT is considerably lower than in TAD trials.[54,55] One analysis estimated that the placebo response is 4.0 HAM-D units lower in psychedelic trials.[55] This suppressed placebo response leads to an inflated between-arm difference, as the treatment arm is measured against a lower floor.

The sum of these 2 effects is 5.2 (ie, 1.29 + 4.0) HAM-D units, which equals the difference of the reported between-arm effects (7.3 − 2.4 = ~5), potentially explaining why H1 failed. We speculate that this suppressed placebo response in PAT trials is attributable to the so-called know-cebo effect (ie, the disappointment patients feel when they realize they are in the control group).[56] This effect could be magnified by PAT's treatment model, where patients undergo extensive therapy in

preparation for a transformative, spiritual experience. Then, on dosing day, patients experience none of these effects and are often bored for the 6 to 8 hours they have to remain on site. This placebo suppression accounts for ~55% (4.0 of 7.3) of the total between-arm effect. That is, ~55% of PAT's between-arm effect is explained by the lack of improvement in the placebo arm. In some psychedelic trials, depression worsened in the placebo group.[57,58] In contrast, the most comprehensive meta-analysis of TAD trials involved 304 placebo groups, and patients improved in all 304 of them.[1]

Limitations

This study has limitations. Some PAT trials exclusively recruited patients with TRD, while no TAD trial focused on TRD. Baseline scores in open-label TAD trials were slightly higher, at a mean (SD) of 22.7 (1.9) compared with 21.3 (3.9) HAM-D units in PAT trials. The lack of difference persisted in sensitivity analyses that excluded PAT trials exclusively recruiting patients with TRD. Furthermore, all models controlled for baseline scores, separating variance due to baseline depression from between-condition contrasts. Therefore, the difference in depression severity is unlikely to explain the lack of a between-treatment difference.

Other differences between PAT and TAD trials included that the primary end point was much later in TAD trials (3.4 weeks for PAT vs 8.1 weeks for TAD), which likely favored PAT. Demographic differences are also likely to exist between the populations. PAT trials are known to disproportionately include patients with high levels of education and to underrepresent minorities,[59] which again may bias the results favoring PAT.[60]

The correct guess rate for treatment assignment in PAT trials is approximately 90% to 95%, and it is likely that some patients had less than absolute confidence. Therefore, there is some blinding in PAT trials, unlike in open-label TAD trials, where the treatment is known with certainty. Relatedly, 2 equally unblinded treatments may not contain equal expectancy effects, in particular, if 1 treatment has better public perception than the other.[61] Arguably, psychedelic medicine has more positive media coverage than traditional antidepressants, suggesting that unblinded PAT may involve larger expectancy effects.

The conversion between depression scales may have introduced artifacts due to, for example, content differences. However, qualitative conclusions remained the same when we used data from the HAM-D scale only, which was the most widely used measure.

For H2 (blinded vs open-label TADs), we used data from Cipriani et al[1] to assess the effect of blind TADs. That meta-analysis used similar, but not identical, inclusion criteria to ours. Thus, confounders may exist in this comparison, such as baseline severity and enrollment decisions, which may bias results.

We only examined symptom reductions. Other key treatment considerations include side effects and functional improvements.[62] At 6-month follow-up, the escitalopram vs psilocybin trial found no difference in depression, but the psilocybin group experienced greater functional improvements,[63] highlighting potential differences beyond core depressive symptoms.

We modeled within-person effect sizes, which include all health changes that result after administration of the treatment, including nonspecific effects such as natural history and placebo. Thus, the within-arm effects reported here should be understood as composite effects, containing both treatment-specific and nonspecific effects; our methodology is not capable of separating the various components.

Lastly, we scanned only PubMed to find relevant studies. Although PubMed is a comprehensive library, it may not include all relevant trials.

Conclusions

In trials of depression, PAT was not more effective than open-label TADs. Blinding made a difference for TADs, but not for PAT, confirming that PAT trials are effectively always open label. PAT's lack of superiority compared with TADs under equal-unblinding conditions argues against overly optimistic narratives about PAT's potential and highlights the importance of blinding integrity when evaluating novel treatments.

ARTICLE INFORMATION

Accepted for Publication: October 13, 2025.

Published Online: March 18, 2026. doi:10.1001/jamapsychiatry.2025.4809

Author Contributions: Drs Williams and Szigeti had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Drs Williams and Szigeti are co–first authors.
Concept and design: Szigeti.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: All authors.
Critical review of the manuscript for important intellectual content: Williams, Szigeti.
Statistical analysis: Williams, Szigeti.
Administrative, technical, or material support: All authors.
Supervision: Szigeti.

Conflict of Interest Disclosures: Dr Williams reported receiving personal fees for consulting from Roche outside the submitted work. No other disclosures were reported.

Data Sharing Statement: See Supplement 2.

REFERENCES

1. 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. *Focus (Am Psychiatr Publ)*. 2018;16(4):420-429. doi:10.1176/appi.focus.16407

2. Stone MB, Yaseen ZS, Miller BJ, Richardville K, Kalaria SN, Kirsch I. Response to acute monotherapy for major depressive disorder in randomized, placebo controlled trials submitted to the US Food and Drug Administration: individual participant data analysis. *BMJ*. 2022;378:e067606. doi:10.1136/bmj-2021-067606

3. Hengartner MP, Plöderl M. Statistically significant antidepressant-placebo differences on subjective symptom-rating scales do not prove that the drugs work: effect size and method bias matter! *Front Psychiatry*. 2018;9:517. doi:10.3389/fpsyt.2018.00517

4. Carhart-Harris R, Giribaldi B, Watts R, et al. Trial of psilocybin versus escitalopram for depression. *N Engl J Med*. 2021;384(15):1402-1411. doi:10.1056/NEJMoa2032994

5. Burke MJ, Blumberger DM. Caution at psychiatry's psychedelic frontier. *Nat Med*. 2021;27(10):1687-1688. doi:10.1038/s41591-021-01524-1

6. Raison CL, Sanacora G, Woolley J, et al. Single-dose psilocybin treatment for major depressive disorder: a randomized clinical trial. *JAMA*. 2023;330(9):843-853. doi:10.1001/jama.2023.14530

7. 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(12):1165-1180. doi:10.1177/0269881116675512

8. 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

9. Muthukumaraswamy SD, Forsyth A, Lumley T. Blinding and expectancy confounds in psychedelic randomized controlled trials. *Expert Rev Clin Pharmacol*. 2021;14(9):1133-1152. doi:10.1080/17512433.2021.1933434

10. Szigeti B, Nutt D, Carhart-Harris R, Erritzoe D. The difference between 'placebo group' and 'placebo control': a case study in psychedelic microdosing. *Sci Rep*. 2023;13(1):12107. doi:10.1038/s41598-023-34938-7

11. Van Elk M, Fried EI. History repeating: a roadmap to address common problems in psychedelic science. *PsyArXiv*. Preprint posted on March 10, 2023. doi:10.31234/osf.io/ak6gx

12. Back AL, Freeman-Young TK, Morgan L, et al. Psilocybin therapy for clinicians with symptoms of depression from frontline care during the COVID-19 pandemic: a randomized clinical trial. *JAMA Netw Open*. 2024;7(12):e2449026. doi:10.1001/jamanetworkopen.2024.49026

13. 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(3):215-223. doi:10.1016/j.biopsych.2022.08.025

14. Bogenschutz MP, Ross S, Bhatt S, et al. Percentage of heavy drinking days following psilocybin-assisted psychotherapy vs placebo in the treatment of adult patients with alcohol use disorder: a randomized clinical trial. *JAMA Psychiatry*. 2022;79(10):953-962. doi:10.1001/jamapsychiatry.2022.2096

15. US Food and Drug Administration. Psychopharmacologic Drugs Advisory Committee (PDAC) meeting June 4, 2024. Accessed November 25, 2024. https://www.youtube.com/watch?v=JqQKP8gcY1E

16. Lin YH, Sahker E, Shinohara K, et al. Assessment of blinding in randomized controlled trials of antidepressants for depressive disorders 2000-2020: a systematic review and meta-analysis. *EClinicalMedicine*. 2022;50:101505. doi:10.1016/j.eclinm.2022.101505

17. Scott AJ, Sharpe L, Colagiuri B. A systematic review and meta-analysis of the success of blinding in antidepressant RCTs. *Psychiatry Res*. 2022;307:114297. doi:10.1016/j.psychres.2021.114297

18. Lieberman JA. Back to the future—the therapeutic potential of psychedelic drugs. *N Engl J Med*. 2021;384(15):1460-1461. doi:10.1056/NEJMe2102835

19. AsPredicted. Psychedelic therapy vs. open-label antidepressant treatment (AsPredicted #142,691). September 5, 2023. Accessed January 22, 2026. https://aspredicted.org/35V_DWY

20. Hengartner MP, Plöderl 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(2):69-73. doi:10.1136/bmjebm-2020-111600

21. Lamers F, van Oppen P, Comijs HC, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):341-348. doi:10.4088/JCP.10m06176blu

22. Furukawa TA, Reijnders M, Kishimoto S, et al. Translating the BDI and BDI-II into the HAMD and vice versa with equipercentile linking. *Epidemiol Psychiatr Sci*. 2019;29:e24. doi:10.1017/S2045796019000088

23. Leucht S, Fennema H, Engel RR, Kaspers-Janssen M, Szegedi A. Translating the HAM-D into the MADRS and vice versa with equipercentile linking. *J Affect Disord*. 2018;226:326-331. doi:10.1016/j.jad.2017.09.042

24. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. 2003;54(5):573-583. doi:10.1016/S0006-3223(02)01866-8

25. Higgins J, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd ed. John Wiley & Sons, Inc; 2019. doi:10.1002/9781119536604

26. Pastor DA, Lazowski RA. On the multilevel nature of meta-analysis: a tutorial, comparison of software programs, and discussion of analytic choices. *Multivariate Behav Res*. 2018;53(1):74-89. doi:10.1080/00273171.2017.1365684

27. Kruschke JK, Liddell TM. The bayesian new statistics: hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychon Bull Rev*. 2018;25(1):178-206. doi:10.3758/s13423-016-1221-4

28. Arns M, Bruder G, Hegerl U, et al. EEG alpha asymmetry as a gender-specific predictor of outcome to acute treatment with different antidepressant medications in the randomized iSPOT-D study. *Clin Neurophysiol*. 2016;127(1):509-519. doi:10.1016/j.clinph.2015.05.032

29. Cervera-Enguix S, Soutullo CA, Landecho I, Murillo-Jelsbak R. Quality of Life in 833 outpatients with major depression treated with open-label venlafaxine extended release: an observational 24-week study. *Int J Psychiatry Clin Pract*. 2003;7(3):193-197. doi:10.1080/13651500310001446

30. Chokka P, Legault M. Escitalopram in the treatment of major depressive disorder in primary-care settings: an open-label trial. *Depress Anxiety*. 2008;25(12):E173-E181. doi:10.1002/da.20458

31. Gargoloff PD, Corral R, Herbst L, Marquez M, Martinotti G, Gargoloff PR. Effectiveness of agomelatine on anhedonia in depressed patients: an outpatient, open-label, real-world study. *Hum Psychopharmacol*. 2016;31(6):412-418. doi:10.1002/hup.2557

32. Hudson JI, Perahia DG, Gilaberte I, Wang F, Watkin JG, Detke MJ. Duloxetine in the treatment of major depressive disorder: an open-label study. *BMC Psychiatry*. 2007;7:43. doi:10.1186/1471-244X-7-43

33. Husain MI, Ledwos N, Fellows E, et al. Serotonergic psychedelics for depression: what do we know about neurobiological mechanisms of action? *Front Psychiatry*. 2023;13:1076459. doi:10.3389/fpsyt.2022.1076459

34. Kennedy SH, Lam RW, Rotzinger S, et al; CAN-BIND Investigator Team. Symptomatic and functional outcomes and early prediction of response to escitalopram monotherapy and sequential adjunctive aripiprazole therapy in patients with major depressive disorder: a CAN-BIND-1 report. *J Clin Psychiatry*. Published online February 5, 2019. doi:10.4088/JCP.18m12202

35. Lesser IM, Myers HF, Lin KM, et al. Ethnic differences in antidepressant response: a prospective multi-site clinical trial. *Depress Anxiety*. 2010;27(1):56-62. doi:10.1002/da.20619

36. Pecenak J, Novotny V. Agomelatine as monotherapy for major depression: an outpatient, open-label study. *Neuropsychiatr Dis Treat*. 2013;9:1595-1604. doi:10.2147/NDT.S49062

37. Pinto C, Trivedi JK, Vankar GK, Sharma PSVN, Narasimha V. An open-label multicentric study of the tolerability and response to escitalopram treatment in Indian patients with major depressive disorder. *J Indian Med Assoc*. 2007;105(7):364, 366, 368 passim.

38. Souery D, Calati R, Papageorgiou K, et al. What to expect from a third step in treatment resistant depression: a prospective open study on escitalopram. *World J Biol Psychiatry*. 2015;16(7):472-482. doi:10.3109/15622975.2014.987814

39. Taragano F, Lyketsos CG, Paz J, Schapira M, Comesaña Díaz E, Klimovsky S. An open-label trial of sertraline for the treatment of major depression in primary care. *Ann Clin Psychiatry*. 1999;11(2):67-71. doi:10.3109/10401239909147051

40. Thase ME, Nierenberg AA, Keller MB, Panagides J; Relapse Prevention Study Group. Efficacy of mirtazapine for prevention of depressive relapse: a placebo-controlled double-blind trial of recently remitted high-risk patients. *J Clin Psychiatry*. 2001;62(10):782-788. doi:10.4088/JCP.v62n1006

41. Trivedi MH, Rush AJ, Wisniewski SR, et al; STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry*. 2006;163(1):28-40. doi:10.1176/appi.ajp.163.1.28

42. Trivedi MH, Cutler AJ, Richards C, et al. A randomized controlled trial of the efficacy and safety of lisdexamfetamine dimesylate as augmentation therapy in adults with residual symptoms of major depressive disorder after treatment with escitalopram. *J Clin Psychiatry*. 2013;74(8):802-809. doi:10.4088/JCP.13m08360

43. Thase ME, Blomgren SL, Birkett MA, Apter JT, Tepner RG. Fluoxetine treatment of patients with major depressive disorder who failed initial treatment with sertraline. *J Clin Psychiatry*. 1997;58(1):16-21. doi:10.4088/JCP.v58n0103

44. Carhart-Harris RL, Bolstridge M, Rucker J, et al. Psilocybin with psychological support for treatment-resistant depression: an open-label feasibility study. *Lancet Psychiatry*. 2016;3(7):619-627. doi:10.1016/S2215-0366(16)30065-7

45. 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(5):481-489. doi:10.1001/jamapsychiatry.2020.3285

46. 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(18):1637-1648. doi:10.1056/NEJMoa2206443

47. 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(4):655-663. doi:10.1017/S0033291718001356

48. Sloshower J, Skosnik PD, Safi-Aghdam H, et al. Psilocybin-assisted therapy for major depressive disorder: an exploratory placebo-controlled, fixed-order trial. *J Psychopharmacol*. 2023;37(7):698-706. doi:10.1177/02698811231154852

49. 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(12):1181-1197. doi:10.1177/0269881116675513

50. Lewis BR, Garland EL, Byrne K, et al. HOPE: a pilot study of psilocybin enhanced group psychotherapy in patients with cancer. *J Pain Symptom Manage*. 2023;66(3):258-269. doi:10.1016/j.jpainsymman.2023.06.006

51. Rosenblat JD, Meshkat S, Doyle Z, et al. Psilocybin-assisted psychotherapy for treatment resistant depression: a randomized clinical trial evaluating repeated doses of psilocybin. *Med*. 2024;5(3):190-200.e5. doi:10.1016/j.medj.2024.01.005

52. Kruschke JK. Rejecting or accepting parameter values in bayesian estimation. *Adv Methods Pract Psychol Sci*. 2018;1(2):270-280. doi:10.1177/2515245918771304

53. Lakens D, McLatchie N, Isager PM, Scheel AM, Dienes Z. Improving inferences about null effects with Bayes factors and equivalence tests. *J Gerontol B Psychol Sci Soc Sci*. 2020;75(1):45-57. doi:10.1093/geronb/gby065

54. Hieronymus F, López E, Werin Sjögren H, Lundberg J. Control group outcomes in trials of psilocybin, SSRIs, or esketamine for depression: a meta-analysis. *JAMA Netw Open*. 2025;8(7):e2524119. doi:10.1001/jamanetworkopen.2025.24119

55. Hsu TW, Tsai CK, Kao YC, et al. Comparative oral monotherapy of psilocybin, lysergic acid diethylamide, 3,4-methylenedioxymethamphetamine, ayahuasca, and escitalopram for depressive symptoms: systematic review and bayesian network meta-analysis. *BMJ*. 2024;386:e078607. doi:10.1136/bmj-2023-078607

56. Butler M, Jelen L, Rucker J. Expectancy in placebo-controlled trials of psychedelics: if so, so what? *Psychopharmacology (Berl)*. 2022;239(10):3047-3055. doi:10.1007/s00213-022-06221-6

57. Cybin. Cybin announces unprecedented positive phase 2 interim data for CYB003 in major depressive disorder meeting primary efficacy endpoint with rapid and significant improvements in depression symptoms after single dose. October 31, 2023. Accessed January 24, 2025. https://www.businesswire.com/news/home/20231031220551/en/Cybin-Announces-Unprecedented-Positive-Phase-2-Interim-Data-for-CYB003-in-Major-Depressive-Disorder-Meeting-Primary-Efficacy-Endpoint-with-Rapid-and-Significant-Improvements-in-Depression-Symptoms-After-Single-Dose

58. GH Research. GH Research announces primary endpoint met in phase 2b trial with GH001 in TRD demonstrating −15.5 point placebo-adjusted MADRS reduction. February 3, 2025. Accessed March 28, 2025. https://www.globenewswire.com/news-release/2025/02/03/3019385/0/en/GH-Research-Announces-Primary-Endpoint-Met-in-Phase-2b-Trial-with-GH001-in-TRD-Demonstrating-15-5-Point-Placebo-adjusted-MADRS-Reduction.html

59. Haft SL, Downey AE, Raymond-Flesch M, et al. A systematic review of participant diversity in psychedelic-assisted psychotherapy trials. *Psychiatry Res*. 2025;345:116359. doi:10.1016/j.psychres.2025.116359

60. Mills JA, Suresh V, Chang L, et al. Socioeconomic predictors of treatment outcomes among adults with major depressive disorder. *Psychiatr Serv*. 2022;73(9):965-969. doi:10.1176/appi.ps.202100559

61. Szigeti B, Heifets BD. Expectancy effects in psychedelic trials. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2024;9(5):512-521. doi:10.1016/j.bpsc.2024.02.004

62. Veal C, Tomlinson A, Cipriani A, et al. Heterogeneity of outcome measures in depression trials and the relevance of the content of outcome measures to patients: a systematic review. *Lancet Psychiatry*. 2024;11(4):285-294. doi:10.1016/S2215-0366(23)00438-8

63. Erritzoe D, Barba T, Greenway KT, et al. Effect of psilocybin versus escitalopram on depression symptom severity in patients with moderate-to-severe major depressive disorder: observational 6-month follow-up of a phase 2, double-blind, randomised, controlled trial. *EClinicalMedicine*. 2024;76:102799. doi:10.1016/j.eclinm.2024.102799