

Review

Psilocybin in Treatment-Resistant Depression: Weighing Rapid Antidepressant Effects Against Uncertain Durability

Lukasz Szarpak ^{1,*}, Michal Pruc ², Dorota Jablonska-Moczulo ³, Iwona Niewiadomska ⁴, Bernard Rybczynski ³ and Wiesław J. Cudała ⁵

¹ Institute of Medical Science, Collegium Medicum, The John Paul II Catholic University of Lublin, 20-400 Lublin, Poland

² Institute of Pharmaceutical Science, Collegium Medicum, The John Paul II Catholic University of Lublin, 20-400 Lublin, Poland; michal.pruc@kul.pl

³ 6th Psychiatric Department, Prof. Jan Mazurkiewicz Masovian Specialist Hospital, 05-802 Pruszkow, Poland; jakotymo@gmail.com (D.J.-M.); bernard.rybczynski@mscz.pl (B.R.)

⁴ Institute of Psychology, The John Paul II Catholic University of Lublin, 20-037 Lublin, Poland; iwona.niewiadomska@kul.pl

⁵ Department of Psychiatry, Medical University of Gdansk, 80-211 Gdansk, Poland; wieslaw.j.cudala@gumed.edu.pl

* Correspondence: lukasz.szarpak@kul.pl

Highlights

What are the main findings?

- COMP001 met its primary efficacy endpoint, whereas EPISODE did not confirm its primary response hypothesis. Secondary findings require interpretation within each trial's design and testing plan.
- Long-term improvement in mean symptom scores does not establish sustained, treatment-free remission. Follow-up involving additional treatment or repeat dosing does not isolate durability after the initial course. No biomarker or retreatment schedule has been validated.

What are the implications of the main findings?

- COMP360 remained investigational when development status was rechecked on 25 September 2026. If authorized, it should initially be delivered in specialist programs with rigorous patient selection, prolonged monitoring, emergency procedures, and structured follow-up.
- Confirmatory trials should assess treatment guessing, standardize psychological support, and compare maintenance strategies using sustained remission, functioning, rescue treatment, cumulative exposure, and clinically relevant adverse events.

Abstract

Background/Objectives: Treatment-resistant depression (TRD) is heterogeneous and recurrent. Psilocybin administered with structured psychological support can reduce symptoms rapidly. Support protocols differ in content and intensity, and trial design complicates estimates of the specific drug effect and durability. We reviewed efficacy, durability, mechanisms, safety, and clinical delivery in TRD. **Methods:** This structured narrative review covered PubMed/MEDLINE literature through 8 August 2026 and preliminary phase 3 reports through 9 August. September checks concerned only COMP360 development status (25 September) and regulatory access (30 September), without extending the clinical evidence cut-off. Priority was given to primary unipolar TRD studies, follow-up reports, and evidence on masking, medication management, safety, and delivery. Sponsor-reported



Academic Editor: Nejat Düzgüneş

Received: 14 August 2026

Revised: 9 October 2026

Accepted: 12 October 2026

Published: 19 October 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

phase 3 findings were considered preliminary because incomplete reporting and the absence of peer review limited independent appraisal. No meta-analysis or formal risk-of-bias assessment was performed. **Results:** COMP001 met its prespecified week-3 primary endpoint for 25 mg versus 1 mg. EPISODE did not confirm its prespecified week-6 response hypothesis, although prespecified secondary continuous outcomes favored 25 mg and remain exploratory. Psilocybin in Depression Resistant to Standard Treatments (PsiDeR) provided exploratory efficacy estimates, including differences from the inert placebo, at weeks 3 and 6 within a feasibility trial. Differences in comparators, treatment guessing, endpoints, and psychological support limit generalization of effect estimates. Long-term follow-up showed improved group means, but sustained response was less common. Additional treatment during follow-up limited attribution to psilocybin alone. Mechanistic studies support 5-hydroxytryptamine receptor 2A (5-HT_{2A}) receptor engagement, plasticity, and network changes. No validated biomarker guides selection or retreatment. Safety requires screening, repeated assessment of suicidality, continuous observation during dosing, defined medical and psychiatric emergency pathways, and structured follow-up after discharge. **Conclusions:** Psilocybin administered with structured psychological support shows short-term antidepressant activity in carefully selected adults with TRD. The specific drug effect, durability, retreatment strategy, and comparative place in care remain uncertain. Confirmatory trials should assess masking, standardize psychological support, and measure sustained remission, functioning, rescue treatment, and cumulative safety.

Keywords: clinical trials; durability; functional unblinding; psilocybin; psychological support; retreatment; safety; treatment-resistant depression

1. Introduction

Treatment-resistant depression (TRD) usually denotes a major depressive episode that has not responded adequately to at least two antidepressant treatments of sufficient dose and duration. The definition is practical rather than biological. It does not capture pseudo-resistance, adherence, comorbidity, episode duration, or the difference between two failed medication trials and failure of multiple pharmacological, psychological, and somatic interventions [1]. Outcomes worsen as treatment steps accumulate, and remission becomes less likely to persist [2]. Any new intervention for TRD must therefore be judged against a disorder characterized by recurrence, functional impairment, and a wide range of prior treatment exposure.

Psilocybin is a prodrug of psilocin and a classic serotonergic psychedelic. Clinical protocols pair a single or limited number of oral doses with preparation, continuous support during a session lasting most of a working day, and one or more post-session visits. This package differs from ordinary drug prescribing. It also complicates causal interpretation because drug exposure, expectancy, interpersonal support, and the therapeutic setting occur together [3]. Throughout this review, the default term is “psilocybin administered with structured psychological support”. Here, structured refers to the support specified by each study protocol. It does not imply equivalent session content, contact time, staff training, or fidelity assessment across trials. We reserve “psychedelic-assisted therapy” and “psilocybin-assisted psychotherapy” for protocols that specify and deliver a defined psychotherapeutic intervention. Preparation, monitoring, and integration sessions are described according to their reported content and are not assumed to constitute psychotherapy.

The evidence base has changed materially since the first open-label reports. A large phase 2 trial identified a short-term dose-related signal [4]. Two randomized trials published

in 2026 produced different primary findings. The Efficacy and Safety of Psilocybin in Treatment-Resistant Major Depression (EPISODE) trial did not meet its primary week-6 response endpoint [5]. The Psilocybin in Depression Resistant to Standard Treatments (PsiDeR) trial, conducted at one National Health Service (NHS) site, reported a large exploratory difference from inert placebo [6]. At the same time, 12-month naturalistic data and sponsor-reported phase 3 results have sharpened questions about maintenance and retreatment. These data warrant a fresh appraisal rather than an extension of conclusions drawn from early, small studies.

This review examines antidepressant efficacy and the persistence of benefit in TRD. It also considers whether mechanistic findings can inform clinical decisions and what delivery conditions would be required if psilocybin entered routine care. The analysis also places the trial results within current safety and regulatory expectations [7,8].

2. Methods

We conducted a structured narrative review of the literature. PubMed/MEDLINE searches were initiated at the outset of manuscript preparation and were updated through 8 August 2026, which constituted the final literature evidence cut-off. Regulatory and sponsor websites were checked through 9 August 2026 for preliminary phase 3 communications included in the evidence synthesis. COMP360 development status was rechecked on 25 September 2026, and official regulatory sources concerning jurisdictional access to psilocybin were checked on 30 September 2026 during revision; neither post-cut-off check extended the clinical evidence cut-off. The literature search was iterative and question-driven, consistent with a narrative-review design, rather than based on a single prespecified systematic-review Boolean strategy. Search terms were used in clinically relevant combinations and included psilocybin, psilocin, treatment-resistant depression, major depressive disorder, randomized trial, placebo, active control, blinding, expectancy, response, remission, relapse, durability, retreatment, antidepressant discontinuation, adverse event, suicidality, psychotherapy, psychological support, set and setting, 5-HT_{2A}, neuroplasticity, brain-derived neurotrophic factor, functional connectivity, biomarker, mechanism, and implementation. No language filter was applied at the database-search stage. Reference lists of relevant primary studies and recent critical reviews were also screened for additional publications. For sponsor-reported analyses, we distinguished the announcement date, the reported follow-up time point, and the underlying trial data cut-off; calendar dates for the latter were not stated in the cited press releases and were therefore recorded as not reported.

Priority was given to peer-reviewed primary clinical studies enrolling adults with unipolar TRD and to their prespecified or naturalistic follow-up reports. TRD studies addressing safety, masking, medication management, psychological support, and treatment delivery were considered directly relevant. Evidence from broader major depressive disorder populations, healthy volunteers, other psychedelic studies, cellular systems, or animal models was used only when TRD-specific evidence was insufficient to address mechanisms, trial design, safety, or implementation. Such evidence was treated as extrapolated and was not considered direct evidence of efficacy, durability, patient selection, or retreatment benefit in unipolar TRD. Regulatory sources were limited to official documents. Sponsor communications were used only to document development status and explicitly attributed preliminary phase 3 findings; they were considered separately from peer-reviewed clinical evidence and were not used to establish comparative efficacy or a treatment recommendation.

The review focused on critical clinical interpretation rather than quantitative synthesis. It was not registered, and neither a formal risk-of-bias assessment nor a meta-analysis was

performed. Consistent with a narrative-review methodology, the search was intended to identify the most clinically and methodologically relevant evidence rather than to generate an exhaustive systematic-review study set. A PRISMA flow diagram, formal duplicate record-level screening, and quantitative screening counts were therefore not used. The author team selected and critically appraised the literature according to relevance to the predefined clinical questions, with attention to population and definition of treatment resistance, comparator, masking and expectancy, endpoint and estimand, attrition and missing data, exposure to rescue treatment or retreatment, adverse-event ascertainment, and specification of psychological support.

Patient and public involvement: Patients and people with lived experience were not involved as contributors to the preparation of this review, interpretation of the evidence, prioritization of outcomes, or formulation of implementation recommendations. These recommendations reflect the authors' assessment of the published evidence.

3. Pharmacology and Mechanistic Evidence

3.1. Receptor Pharmacology and Plasticity

After oral administration, psilocybin is dephosphorylated to psilocin. The acute subjective and physiological effects are attributed mainly to partial agonism at cortical serotonin 2A (5-HT_{2A}) receptors, with contributions from other serotonergic targets [9]. The resulting state includes altered perception, salience, self-processing, and emotional responsiveness. Receptor occupancy establishes target engagement, but it does not by itself explain why symptoms improve days or weeks after psilocin has been cleared.

Preclinical studies provide a plausible bridge from acute receptor activation to longer-lived cellular change. Psychedelic compounds increase neuritogenesis and spinogenesis in cultured neurons and animal models [10]. In mice, psilocybin produced rapid growth of dendritic spines in the frontal cortex, with a proportion of new spines persisting for at least one month [11]. More recent work implicates intracellular 5-HT_{2A} signaling and has also reported direct binding of psychedelics to the tropomyosin receptor kinase B (TrkB) receptor in preclinical models [12,13]. These experiments establish biological effects in cellular or animal systems [10–13]. They do not reproduce the prolonged illness, repeated treatment exposure, psychiatric comorbidity, or psychological context of human TRD. Differences in species, drug exposure, brain region, and measurement timing further limit direct extrapolation. Increased spine density or altered plasticity signaling therefore cannot be interpreted as evidence of remission, functional recovery, or durable benefit in patients. Whether these processes contribute to improvement in human TRD, and under which clinical conditions, remains to be established.

Plasticity should also not be treated as intrinsically therapeutic. A period of increased learning or cognitive flexibility may allow adaptive change, but its direction will depend on symptoms, expectations, relationships, and the environment in which new learning occurs. This is one reason that pharmacology and treatment context cannot be separated by rhetoric when they have not been separated experimentally.

3.2. Network Effects and Candidate Biomarkers

Human neuroimaging studies show marked acute changes in large-scale network organization, followed in some studies by post-acute changes associated with symptom improvement. Increased global integration and reduced modular segregation have been reported after psilocybin treatment for depression [14]. That analysis included an open-label TRD cohort and a separate randomized trial in a broader major depressive disorder (MDD) population, so its findings should not be interpreted as validation of an individual prediction rule for TRD [14]. Earlier TRD work identified changes in default mode network

connectivity after treatment [15]. The direction and location of findings vary across tasks, acquisition windows, and analytic methods, and a systematic appraisal found insufficient evidence for a reproducible imaging marker of clinical response [16].

Peripheral biomarkers are even further from clinical use. Small human blood-biomarker studies have reported post-psilocybin increases in plasma brain-derived neurotrophic factor (BDNF), whereas findings for several other peripheral biomarkers have been more variable [17]. More broadly, the clinical interpretation of peripheral BDNF in major depression is influenced by illness state, treatment, sampling, and assay conditions [18]. Evidence that a biological process occurs after dosing does not establish that measuring it will predict clinical benefit. A post-treatment biomarker change may reflect drug exposure, altered physiological state, or improvement already underway without mediating recovery. Likewise, an association between a baseline marker and outcome within the psilocybin group does not establish that the marker identifies greater benefit from psilocybin than from a comparator [19]. Within the evidence reviewed, no imaging or peripheral biomarker has been validated for patient selection, prediction of sustained remission, or retreatment decisions in TRD [16–18]. These measures should remain exploratory and should not substitute for clinical outcomes.

3.3. Subjective Experience and Psychological Process

Exploratory studies have associated later improvement with psychological flexibility [20], the quality of the acute psychedelic experience [21], and therapeutic alliance or rapport [22]. Post hoc analyses of COMP001 in TRD found associations between certain dimensions of the acute psychedelic experience and week-3 symptom improvement [23]. A subsequent path analysis in the 25 mg group found weaker associations for pre-dosing therapeutic alliance than for measures of the psychedelic experience. That analysis found no nominally significant direct effect of alliance on week-3 Montgomery-Åsberg Depression Rating Scale (MADRS) scores [24]. These findings are clinically interesting but remain exploratory and do not establish causal mediation or a necessary experiential pathway. Subjective intensity is measured after randomization, tracks dose and treatment guess, and may reflect both pharmacological exposure and expectancy. Standard regression cannot determine whether the experience causes improvement, marks adequate exposure, or is a correlate of participants learning that they received active treatment. No validated experiential threshold should guide redosing or be presented to patients as a requirement for benefit. Table 1 summarizes the translational status of the principal proposed mechanisms, distinguishing biological plausibility from evidence sufficient to guide clinical decisions.

Table 1. Translational status of proposed antidepressant mechanisms of psilocybin.

Mechanistic Domain	Best Available Evidence	Principal Limitation	Current Clinical Relevance
5-HT2A receptor engagement	Psilocin is principally a partial agonist at cortical 5-HT2A receptors, accounting for much of the acute psychedelic state and establishing target engagement [9].	This explains acute psychoactive effects more directly than delayed antidepressant benefit. Within the literature reviewed, no study has established that 5-HT2A engagement predicts response or durability in TRD.	Confirms pharmacological activity but does not identify responders, define a therapeutic exposure threshold, or guide retreatment.

Table 1. Cont.

Mechanistic Domain	Best Available Evidence	Principal Limitation	Current Clinical Relevance
Synaptic plasticity and TrkB-related signaling	Psychedelics promote neurogenesis and spinogenesis; psilocybin induces persistent dendritic spine growth in mice. Intracellular 5-HT2A signaling and direct TrkB binding have been implicated [10–13].	Evidence is mainly cellular or animal; causal mediation of antidepressant response in humans has not been demonstrated, and the direction of plastic change may depend on treatment context.	Supports biological plausibility only. No validated measure identifies a therapeutic plasticity window or confirms dose adequacy.
Large-scale network reorganization	Human imaging studies report increased global integration and post-treatment default mode network changes after psilocybin [14,15].	Samples are small, methods vary, and independent replication is limited. No validated imaging measure currently provides reliable individual-level prediction [16].	Remains a research outcome. Neuroimaging cannot currently support patient selection, response monitoring, or relapse prediction.
Peripheral biomarkers including BDNF	Small human blood-biomarker studies have reported post-psilocybin increases in plasma BDNF; broader depression data remain sensitive to illness state, treatment, sampling, and assay conditions [17,18].	Studies are small, sampling varies, and peripheral concentrations may not reflect central plasticity. Predictive performance has not been established.	No current role in selecting patients, confirming target engagement, monitoring response, or determining when retreatment is indicated.
Subjective experience and psychological flexibility	Acute experience quality and psychological flexibility have correlated with improvement in exploratory analyses; findings for therapeutic alliance are mixed, including post hoc analyses of COMP001 [20–24].	These post-randomization measures track dose, expectancy, treatment guessing, and therapist interaction; associations cannot distinguish causal mediation from an exposure marker or functional unblinding.	No experiential threshold should define adequate treatment, be presented as necessary for benefit, or determine redosing.

Notes. Biological plausibility, pharmacodynamic activity, causal mediation, and predictive utility represent different levels of evidence. The candidate biomarkers summarized here remain exploratory and are not validated surrogate endpoints or tools for patient selection, dose selection, or retreatment. Abbreviations: 5-HT2A, 5-hydroxytryptamine receptor 2A; BDNF, brain-derived neurotrophic factor; TRD, treatment-resistant depression; TrkB, tropomyosin receptor kinase B.

4. Clinical Evidence in Treatment-Resistant Depression

4.1. Early and Phase 2 Studies

The first studies in treatment-resistant depression (TRD) were designed to establish feasibility and identify a clinical signal. They were useful for deciding whether larger trials were justified, but they were not designed to provide a stable estimate of efficacy. In the open-label study by Carhart-Harris and colleagues, 20 participants received 10 mg of psilocybin followed 1 week later by 25 mg, with psychological support during both sessions. Nine participants met response criteria and four met remission criteria at week 5. Mean symptom scores remained below baseline at 6 months [25]. Without a concurrent control, however, the contributions of expectancy, repeated clinical contact, regression to the mean, and subsequent treatment cannot be separated from the effect of psilocybin.

Rosenblat and colleagues randomized 31 participants to immediate treatment or a waitlist. Thirty received psilocybin, and 29 contributed to the week-2 analysis. The treated sample included 26 participants with major depressive disorder and four with bipolar II disorder. Feasibility was the primary objective. The secondary efficacy analysis showed a mean reduction in MADRS score of 9.6 points with immediate treatment versus 3.0 points during the waiting period [26]. The later flexible course, which allowed one to three administrations, offered preliminary information about repeat dosing but no longer preserved a controlled comparison. A waitlist also controls poorly for therapeutic contact and treatment expectation, two factors that are especially relevant when the acute effects make allocation obvious.

Small studies in more refractory populations showed that substantial improvement is possible, while also making the heterogeneity of individual outcomes visible. In a 12-participant open-label study of severe TRD after at least five failed treatments, a single 25 mg dose was followed by least-squares mean MADRS changes of -15.8 points at week 3 and -17.2 points at week 12 [27]. In a separate pilot with seven completers who received two 25 mg sessions, two participants had a sustained response, three relapsed, and two did not respond [28]. These observations are clinically informative, but neither the large mean changes nor associations with features of the acute experience should be used to select patients until replicated in prospective, adequately powered studies.

The multicenter phase 2b study by Goodwin and colleagues provided the first large controlled estimate in TRD. After discontinuing antidepressants, 233 participants were randomized to a single 25 mg, 10 mg, or 1 mg dose with structured psychological support. At week 3, the adjusted MADRS difference between 25 mg and 1 mg was -6.6 points (95% confidence interval [CI], -10.2 to -2.9 ; $p < 0.001$), whereas 10 mg did not separate from 1 mg [4]. The week-3 comparison of 25 mg with 1 mg met the prespecified primary endpoint. Sustained response through week 12 was a prespecified secondary endpoint, but formal testing of secondary endpoints was not undertaken after the 10 mg primary comparison failed within the testing hierarchy. These secondary findings therefore remain descriptive and do not establish maintenance efficacy after a single administration [4]. The 1 mg condition may have preserved some uncertainty before dosing, yet it was not inert and cannot be assumed to have preserved masking once the subjective effects emerged.

4.2. What the 2026 Randomized Trials Add

EPISODE tested psilocybin under more restrictive conditions than the early trials. This two-center, triple-masked phase 2b trial randomized 144 adults. Of these, 143 received at least one dose, and 142 entered the main efficacy analysis under the prespecified while-receiving-treatment estimand. The primary assessment took place at week 6, before the second administration [5]. At week 6, response, defined as a reduction of at least 50% from baseline in the 17-item Hamilton Depression Rating Scale (HAMD-17) score, occurred in 8/47 participants (17.0%) after 25 mg, 6/48 (12.5%) after 5 mg, and 5/47 (10.6%) after nicotinamide. The prespecified comparison between 25 mg and nicotinamide was not significant (adjusted odds ratio, 1.73; 95% CI, 0.53–6.23; $p = 0.19$) [5]. The prespecified superiority hypothesis was therefore not confirmed. This does not establish absence of an antidepressant effect: the confidence interval includes both no advantage and a potentially substantial increase in the odds of response. The primary result remains inconclusive, and exploratory continuous outcomes cannot replace the failed confirmatory test [5].

Continuous HAMD-17 scores nevertheless favored 25 mg at week 6, with estimated differences of -4.60 points versus nicotinamide (95% CI, -7.01 to -2.18) and -3.09 points versus 5 mg (95% CI, -5.50 to -0.69) [5]. These continuous outcomes were prespecified secondary endpoints, but their interpretation remains exploratory because the primary

comparison did not meet the criterion required to proceed through the fixed testing sequence [5]. The time course also matters. Response at week 1 was 34.0% after 25 mg, compared with 10.4% after 5 mg and 6.4% after nicotinamide, followed by attenuation at week 6. A second administration was given after the primary assessment, and all groups had received at least one 25 mg dose by week 12 [5]. No between-group difference was then detected, and the study did not establish an incremental benefit from a second 25 mg administration within the observed period.

PsiDeR produced a larger numerical estimate in a different setting. This single-site National Health Service feasibility trial randomized 60 adults to 25 mg or inert placebo with preparation, support during dosing, and integration in both groups. All 60 randomized participants received their allocated treatment. Follow-up MADRS assessments were available for 59 participants, comprising 30 psilocybin recipients and 29 placebo recipients [6]. The adjusted between-group difference was -10.41 points at week 3 (95% CI, -14.86 to -5.95) and -12.92 points at week 6 (95% CI, -17.37 to -8.47). The EPISODE and PsiDeR results are not opposing verdicts on the same experiment. At week 3, response occurred in 13/30 participants receiving psilocybin (43.3%) and 1/29 receiving placebo (3.4%). Remission occurred in 12/30 (40.0%) and 1/29 (3.4%), respectively [6]. The clinical efficacy outcomes were prespecified secondary outcomes and remain exploratory because the trial was designed primarily to assess feasibility and estimate outcome variance [6]. Assessment of participant blinding was introduced part-way through the trial and covered only 23 of 60 randomized participants. In this post hoc analysis, 13/13 psilocybin recipients and 7/10 placebo recipients correctly identified their allocation. Independent outcome raters and instructions not to disclose allocation beliefs provided some protection against observer bias, but the reported assessment does not establish that rater blinding was maintained or quantify the contribution of expectancy to the treatment difference [6].

These estimates also differ in what they measure. A response odds ratio describes a relative contrast in the odds of crossing a prespecified threshold; it does not quantify absolute symptom improvement or the absolute probability of benefit. MADRS and HAMD-17 point differences retain their respective scale units, whereas standardized mean differences depend on the variability used for standardization. These measures should therefore be interpreted alongside arm-level scores, absolute response and remission proportions, assessment timing, and the estimand, rather than ranked as interchangeable measures of efficacy. EPISODE used two comparator conditions and a categorical week-6 endpoint. PsiDeR used inert placebo, treated continuous MADRS change as an exploratory efficacy outcome, and was conducted at one site. Their samples, placebo trajectories, and support models also differed. The resulting contrast is clinically important because it shows that the estimated magnitude of benefit is sensitive to comparator choice, endpoint definition, and timing. It does not provide a principled reason to regard the larger estimate as the effect most likely to be observed in routine care.

Treatment resistance also differed across these populations. Entry after two adequate medication failures does not establish comparability with patients who remain depressed after multiple augmentation strategies, psychotherapy, or somatic treatments. Some protocols counted augmentation, psychotherapy, or treatment intolerance towards eligibility, whereas others restricted the number of previous failures or excluded particular treatment histories [4–6,27,29,30]. These differences limit indirect comparisons of efficacy. They should not, however, be assumed to explain differences between trial results without appropriate analyses. Small studies enrolling more refractory patients do not establish efficacy after failure of ECT or other neuromodulation, and the available evidence does not validate prior treatment burden as a predictor of individual response.

EPISODE also showed substantial functional unblinding despite its triple-masked design. Among participants who answered the relevant question after the first administration, 38/44 (86.4%) in the 25 mg group correctly identified receipt of the high dose. Therapists correctly identified it in 71/87 ratings (81.6%), with two therapists potentially contributing ratings for each participant. Allocation guesses were not collected from outcome raters, and patient expectancy was not measured [5]. The descriptive comparison of week-6 outcomes by allocation belief cannot establish that expectancy was unimportant, particularly because only six high-dose recipients did not identify the high dose. Blinding percentages should not be compared directly with PsiDeR because the questions, sampling, and available treatment conditions differed [5,6].

Recognizing allocation may change expectations, engagement with psychological support, symptom reporting, and staff behavior. Randomization supports comparison of the assigned interventions under the trial conditions, but it does not isolate a pharmacological effect independent of these pathways. Expectancy-related improvement may form part of the clinical effect of the treatment package, whereas differential reporting or rating can introduce measurement bias. Limited improvement in a comparator group is compatible with disappointment effects but does not demonstrate them. Neither trial provides sufficient evidence to quantify these contributions or attribute the observed between-group difference to pharmacological action alone [5,6,8,31].

Masking remains the most persistent threat to interpretation. A 2026 systematic review found that only 33 of 112 psychedelic randomized trials assessed blinding integrity. Blinding failure rates exceeding 90% were frequently reported in studies of psilocybin, lysergic acid diethylamide, and ayahuasca. No comparator strategy consistently preserved masking [32]. An indirect meta-analysis sought to address differences in functional unblinding by comparing symptom change in trials of psychedelic interventions with that in open-label trials of conventional antidepressants. The estimated difference was 0.3 HAMD-17 points favoring antidepressants (95% CI, -1.39 to 1.98 ; $p = 0.73$) [33]. The analysis did not identify a statistically significant difference. It included depression populations beyond TRD and cannot substitute for a head-to-head trial. Its relevance is narrower: conspicuous psychoactive effects limit how confidently an observed group difference can be assigned to drug action alone. These design problems are considered further in Section 5. A 2026 living systematic review and meta-analysis included 15 randomized trials with 801 participants. Twelve trials (585 participants) contributed to the primary model, which favored psilocybin for depressive symptoms (Hedges' $g = -0.90$) [34]. However, many included studies were small or at risk of bias, and pooling across heterogeneous populations and comparators does not resolve the TRD-specific effect estimate or the problem of functional unblinding.

Taken together, the published randomized evidence supports short-term antidepressant activity after 25 mg in a supervised treatment package. It does not yet provide a single transportable effect size, a dependable individual response probability, or an evidence-based rule for a second administration. For a clinician, that is a more useful conclusion than classifying trials as simply positive or negative.

4.3. Phase 3 Development Status

At the review cut-off of 9 August 2026, the phase 3 findings summarized here were available as sponsor communications rather than peer-reviewed trial reports. In its announcement dated 23 June 2025, Compass Pathways reported that COMP005 had dosed 258 participants at 32 US sites. The reported between-group difference in MADRS change at week 6 was -3.6 points for a single 25 mg administration versus placebo (95% CI, -5.7 to -1.5 ; $p < 0.001$) [35]. In its announcement dated 17 February 2026, the sponsor reported that COMP006 had dosed 581 participants and compared two administrations, 3 weeks

apart, of 25 mg, 10 mg, or 1 mg. The reported week-6 difference in MADRS change for 25 mg versus 1 mg was -3.8 points (95% CI, -5.8 to -1.8 ; $p < 0.001$) [36].

In the sponsor's report, 39% of COMP006 participants in the 25 mg arm met a threshold of at least 25% MADRS reduction at week 6 after two scheduled administrations. Subsequent benefit was described as maintained on average through week 26 [37]. This is below the conventional 50% response threshold and does not establish continuous individual response. Retreatment was permitted during Part B. The reported follow-up therefore concerns a treatment strategy; it does not isolate persistence after the initial course. Participant-level trajectories, retreatment exposure, and missing-data methods are needed to interpret this claim.

The sponsor-reported differences of 3.6 and 3.8 MADRS points indicate statistical separation between treatment groups at week 6. Their clinical importance requires separate assessment. A between-group mean difference does not describe the improvement experienced by an individual patient or identify how many patients obtained a worthwhile benefit. Thresholds for meaningful within-patient change cannot be applied directly to between-group mean differences [38]. These results therefore neither establish clinically meaningful benefit for each treated patient nor exclude substantial improvement in a subset. Interpretation requires the distribution of individual changes, conventional response and remission proportions with denominators and confidence intervals, and corresponding changes in functioning and quality of life.

The estimates were obtained within supervised treatment programs providing psychological support. Comparable support across randomized groups does not by itself invalidate the treatment contrast, but it limits extrapolation to administration outside that setting. Possible functional unblinding may also influence expectancy, participant behavior, and interactions with staff; the available results do not quantify its contribution to the observed differences. The cited sponsor releases do not provide arm-level baseline and week-6 scores, conventional response and remission counts, or complete estimand specifications. These omissions limit independent assessment of adherence to prespecified analyses, handling of missing data and subsequent treatment, multiplicity control, and completeness of safety reporting. The absence of peer review further limits external scrutiny. Reported p -values and confidence intervals describe the announced analyses but do not establish robustness to alternative analytical assumptions or exclude selective reporting. Full trial reports, protocols, statistical analysis plans, and sensitivity analyses are needed to assess the clinical benefit-harm balance. Until these materials permit independent appraisal, the findings should remain provisional and should not, by themselves, support treatment recommendations [35–37].

After the clinical evidence cut-off, Compass Pathways released 52-week open-label COMP005 Part C topline data on 9 September 2026 [39]. Because these sponsor-reported data were released after the prespecified clinical evidence cut-off and remain non-peer-reviewed, they were not incorporated into the efficacy or durability synthesis. Development status was rechecked on 25 September 2026; subsequent public sponsor communications continued to describe COMP360 as investigational [40]. This limited post-cut-off check did not extend the clinical evidence cut-off.

4.4. Durability Is an Estimand, Not a Single Observation

Long-term mean symptom change and individual durability outcomes answer different clinical questions. A lower group mean at 6 or 12 months can reflect persistent partial improvement, sustained remission in a minority, or improvement and relapse occurring in different participants. It does not establish that the same patients remained well throughout follow-up. Sustained response requires maintenance of a specified reduction from baseline,

whereas sustained remission requires maintenance below a defined symptom threshold over a prespecified interval. Remission at one late visit establishes neither. A durability estimand must additionally specify the population, treatment comparison, handling of rescue treatment and retreatment, and population-level summary measure [41]. Table 2 distinguishes these outcomes and the limitations of the available follow-up evidence.

Table 2. Study populations and characterization of treatment resistance in psilocybin trials.

Study and Population	Treatment Resistance: Eligibility and Observed History	Chronicity	Previous Psychotherapy and Somatic Treatment
Carhart-Harris et al., 2018 [25]. Twenty adults with unipolar TRD.	At least two unsuccessful courses of distinct antidepressants during the current episode, each lasting at least 6 weeks. Mean lifetime medication failures: 4.6 (SD 2.6).	Mean reported illness duration: 17.7 years. The authors acknowledged difficulty distinguishing episode duration from chronic illness.	Previous psychotherapy: 17/20. Its adequacy was not standardized. Previous ECT or TMS: NR.
Goodwin et al., 2022, COMP001 [4,30]. 233 adults with unipolar TRD.	Two to four unsuccessful pharmacological treatments during the current episode, with adequate dose and duration of at least 8 weeks assessed using the MGH-ATRQ. Approved augmentation could count as an additional treatment.	A current episode lasting more than 1 year was reported by 86%.	Previous psychotherapy adequacy: NR. ECT or ketamine during the current episode was exclusionary.
Rosenblat et al., 2024 [26]. 31 randomized and 30 treated, including 26 with MDD and four with bipolar II depression.	At least two adequate pharmacological treatment failures during the current episode, without an upper limit. Mean reported failed medication trials: 11.27 (SD 5.59).	Current episode at least 3 months. Mean reported time with depression: 18.3 years (SD 11.5), which is not a standardized current-episode duration.	Previous psychotherapy dose and adequacy: NR. Stable pre-existing psychotherapy was permitted.
Aaronson et al., 2025 [27]. Twelve adults with severe unipolar TRD.	At least five adequate treatment failures during the current episode. Mean medication failures during that episode: 5.6 (SD 1.4).	Mean current-episode duration: 9.1 years (SD 8.8).	Completed ECT or TMS courses and up to one psychotherapy course could contribute to eligibility. Seven of 12 participants were receiving psychotherapy at enrollment.
Mertens et al., 2026, EPISODE [5]. 144 adults randomized.	Two adequate pharmacological treatments during the current episode, each lasting at least 6 weeks at a minimum effective dose. Augmentation could count as the second treatment. Mean lifetime antidepressant agents tried: 5.43 (SD 2.83). Mean augmentation agents tried: 1.16 (SD 1.77).	Mean time since first MDD diagnosis: 13.86 years (SD 9.65). Current-episode duration: NR.	Previous psychotherapy adequacy: NR. ECT within 12 months, ketamine or esketamine within 6 months, and any previous VNS or DBS were exclusionary.
Rucker et al., 2026, PsiDeR [6,29]. Sixty adults randomized and treated.	Two antidepressant trials or one antidepressant trial plus psychotherapy, without an upper limit. The published protocol specified at least 6 weeks at a minimum effective antidepressant dose and at least six psychotherapy sessions, but also allowed intolerance to count as failure. Median MSM score: 8.0 (IQR 7.0–9.5).	Mean time with significant depression: 20.0 years (SD 10.2). Current-episode duration: NR.	Psychotherapy could contribute to eligibility. Actual previous psychotherapy exposure was not separately summarized. No upper limit on the number or type of previous treatments was specified.

Table 2. Cont.

Study and Population	Treatment Resistance: Eligibility and Observed History	Chronicity	Previous Psychotherapy and Somatic Treatment
COMP005 [35,36] and COMP006 [36,37]. Respectively, 258 and 581 participants dosed. Preliminary sponsor reports.	Registered eligibility required two to four adequate pharmacological treatment failures during the current episode, assessed using the MGH-ATRQ. Actual failure distributions: NR in the cited sponsor reports.	For a first lifetime episode, registered eligibility required a duration of 3 months to 2 years. Actual episode-duration distributions: NR in the cited sponsor reports.	Previous psychotherapy exposure: NR in the cited sponsor reports. ECT, DBS, or VNS during the current episode, and TMS within the preceding 6 months, were exclusionary.

Notes. Eligibility criteria are distinguished from observed treatment histories. Lifetime treatment exposure, failures during the current episode, and different measures of chronicity are not interchangeable. Previous psychotherapy does not establish an adequate unsuccessful course. COMP005 and COMP006 entries combine registered eligibility criteria with preliminary sponsor reports. NR indicates information not reported in the reviewed sources, rather than absence of an assessment or treatment. Abbreviations: DBS, deep brain stimulation; ECT, electroconvulsive therapy; IQR, interquartile range; MDD, major depressive disorder; MGH-ATRQ, Massachusetts General Hospital Antidepressant Treatment Response Questionnaire; MSM, Maudsley Staging Method; NR, not reported; SD, standard deviation; TMS, transcranial magnetic stimulation; TRD, treatment-resistant depression; VNS, vagus nerve stimulation.

The handling of additional treatment changes the clinical question. Outcomes recorded after rescue medication, psychotherapy, or repeat psilocybin administration describe the course under that treatment strategy. They cannot establish persistence attributable to the initial psilocybin course alone. Treatment-free remission requires remission over a specified interval after completion of the initial protocol, without ongoing or additional depression treatment, including further psychedelic dosing. Excluding participants who receive additional treatment, or censoring their observations when treatment begins, may preferentially remove those whose symptoms worsen. Such analyses require explicit assumptions and sensitivity analyses. Missing assessments also require prespecified handling and should not be treated as evidence that improvement continued.

COMP004 enrolled 66 participants, including 58 from COMP001 and eight from COMP003, for observational follow-up extending to 52 weeks after the original administration. The prespecified time-to-event analysis included all 233 participants from COMP001. Median time to a depressive event was 92, 83, and 62 days in the original 25 mg, 10 mg, and 1 mg groups, respectively [42]. These estimates were descriptive. A separate post hoc analysis restricted to the 58 COMP001 participants entering the extension yielded medians of 189, 43, and 21 days. Selective extension participation and uncontrolled subsequent treatment limit interpretation of this subgroup comparison.

The naturalistic EPISODE follow-up included 126 of the original 144 randomized participants who attended at least one assessment at 6 or 12 months after the first administration. All participants in this follow-up cohort had received at least one 25 mg dose during the main trial [43]. Mean HAMD-17 change from the original baseline was −7.93 points at 6 months (95% CI, −9.17 to −6.70) and −7.74 points at 12 months (95% CI, −9.04 to −6.43), with no difference according to original allocation [43]. Study-defined sustained HAMD-17 response from week 12 was reported in 28/126 participants (22.2%) at 6 months and 20/126 (15.9%) at 12 months. Sustained remission, defined using HAMD-17 <8, was less frequent. Summing the published arm-level counts gives 20/126 participants (15.9%) at 6 months and 10/126 (7.9%) at 12 months [43]. These findings concern maintenance of the respective criteria at scheduled assessments and do not establish uninterrupted remission between visits or remission without subsequent treatment. By 12 months, 32% had initiated antidepressant pharmacotherapy, 25% reported further psychedelic use, and 68% had attended at least one optional integration group [43]. Statistical adjustment for

recorded treatment did not recreate a randomized long-term comparison. The study describes outcomes after exposure to psilocybin and an extended therapeutic program. It does not establish a 12-month drug effect.

A small trial in recurrent major depressive disorder, rather than TRD, provides useful contextual evidence because it retained a randomized niacin comparison for 12 months. Among 35 participants, the MADRS difference favored 25 mg through day 42 but not at day 365 (−3.68 points; 95% CI, −9.30 to 1.94; $p = 0.20$). Self-reported separation persisted through day 102 and shortened to day 72 when observations were censored after antidepressant initiation [44]. The population and sample size preclude direct extrapolation to TRD. The result nonetheless shows why persistent within-group improvement and persistent randomized separation are not the same claim.

Current evidence therefore does not define a maintenance schedule, a symptom threshold for retreatment, or the cumulative risk of repeated exposure in TRD. Psilocybin is better understood at present as a potentially rapid intervention delivered within a resource-intensive clinical program, with maintenance needs that remain unresolved. Describing it as a one-time treatment with durable efficacy goes beyond the published data. Tables 2–4 summarize the clinical evidence and distinguish longer-term outcomes according to retention, subsequent treatment, and the definition of persistence or relapse. Table S1 provides arm-level symptom scores, response and remission denominators, treatment contrasts with confidence intervals, and the estimand specifications available for the principal randomized trials.

Table 3. Short-term efficacy outcomes in trials of psilocybin administered with psychological support.

Study and Treatment	Outcome, Timing, and Analysis Population	Main Finding	Interpretation
Carhart-Harris et al., 2018 [25]. Open-label administration of 10 mg followed by 25 mg 1 week later.	QIDS-SR-16 change 5 weeks after 25 mg. Longitudinal analysis: 19 participants.	Mean change: −9.2 points (95% CI, −11.8 to −6.6).	Uncontrolled within-group improvement. The specific treatment effect cannot be estimated.
COMP001 [4]. Double-blind comparison of single doses of 25 mg, 10 mg, and 1 mg.	Prespecified primary endpoint: MADRS change at week 3. Analysis population: 233.	Adjusted difference versus 1 mg: −6.6 points for 25 mg (95% CI, −10.2 to −2.9) and −2.5 points for 10 mg (95% CI, −6.2 to 1.2).	The 25 mg comparison met the primary endpoint. The 10 mg comparison did not, stopping the testing hierarchy. Secondary findings remained descriptive.
Rosenblat et al., 2024 [26]. Open-label randomized comparison of immediate 25 mg treatment with a 2-week waitlist.	Feasibility was primary. MADRS change at week 2 was a secondary efficacy outcome, assessed in 29 participants.	Mean MADRS reduction: 9.6 points with immediate treatment versus 3.0 points during the waitlist period.	The waitlist did not control for expectancy or therapeutic contact. Findings include bipolar II depression.
Aaronson et al., 2025 [27]. Open-label single 25 mg administration.	Primary endpoint: MADRS change at week 3. Twelve participants.	Least-squares mean change: −15.8 points (95% CI, −25.4 to −6.3).	A small uncontrolled study in a highly refractory population. Comparative efficacy remains uncertain.
EPISODE [5]. Two-center, triple-masked trial. First administration: 25 mg, 5 mg, or nicotinamide 100 mg.	Prespecified primary endpoint: ≥50% HAM-D-17 reduction at week 6, before the second administration. Of 144 randomized participants, 143 received treatment and 142 entered the primary analysis.	Response: 8/47 (17.0%), 6/48 (12.5%), and 5/47 (10.6%), respectively. Adjusted OR for 25 mg versus nicotinamide: 1.73 (95% CI, 0.53 to 6.23; $p = 0.19$).	The prespecified hypothesis was not confirmed. The confidence interval indicates substantial uncertainty. Continuous secondary findings remain exploratory.

Table 3. Cont.

Study and Treatment	Outcome, Timing, and Analysis Population	Main Finding	Interpretation
PsiDeR [6]. Single-site, double-blind comparison of 25 mg with inert placebo.	Primary outcomes concerned feasibility and MADRS variance. Efficacy outcomes were prespecified secondary outcomes. Follow-up MADRS data: 59/60 participants.	Adjusted MADRS difference: −10.41 points at week 3 (95% CI, −14.86 to −5.95) and −12.92 at week 6 (95% CI, −17.37 to −8.47).	Exploratory efficacy estimates from a feasibility trial. Small sample size, limited placebo improvement, and functional unblinding constrain interpretation.
COMP005 [35,36]. Double-blind phase 3 comparison of single-dose COMP360 25 mg with placebo. Sponsor announcements: 23 June 2025 and 17 February 2026.	Sponsor-reported primary endpoint: MADRS change at week 6. Dosed population: 258.	Reported difference: −3.6 points (95% CI, −5.7 to −1.5; $p < 0.001$).	Preliminary sponsor-reported statistical separation. Individual clinical benefit requires fuller reporting of response, remission, functioning, and masking.
COMP006 [36]. Double-blind phase 3 comparison of two fixed administrations of 25 mg, 10 mg, or 1 mg, 3 weeks apart. Sponsor announcement: 17 February 2026.	Sponsor-reported primary endpoint: MADRS change at week 6 after the first administration, 3 weeks after the second. Dosed population: 581.	Reported difference for 25 mg versus 1 mg: −3.8 points (95% CI, −5.8 to −1.8; $p < 0.001$).	Preliminary findings concern a two-dose course. Statistical separation does not establish clinically important improvement for each participant.

Notes. Negative symptom-score changes indicate improvement. Within-group changes, between-group differences, and odds ratios are not directly comparable. Statistical significance alone does not establish individual clinical benefit. Phase 3 findings were preliminary and not peer-reviewed at the sponsor-source cut-off of 9 August 2026. Announcement dates are not trial data cut-off dates, which were not reported. Table S1 provides detailed efficacy data and available estimand specifications. Abbreviations: CI, confidence interval; COMP360, investigational synthetic psilocybin formulation; HAMD-17, 17-item Hamilton Depression Rating Scale; MADRS, Montgomery-Åsberg Depression Rating Scale; OR, odds ratio; QIDS-SR-16, 16-item Quick Inventory of Depressive Symptomatology, self-report version.

Table 4. Durability of clinical improvement, subsequent treatment, and retreatment in psilocybin studies.

Study and Assessment	Durability Outcome and Finding	Follow-Up Completeness	Subsequent Treatment and Psychedelic Exposure	Interpretation
Carhart-Harris et al., 2018 [25]. Six months.	Late-visit status among nine week-5 responders. Three met the study-defined relapse criterion of QIDS-SR-16 ≥ 6 at 6 months.	Nineteen of 20 completed all assessments. Six-month QIDS-SR-16 data were available for all 20.	Six started antidepressants, five received later psychotherapy, and five used nontrial psilocybin. One continued venlafaxine throughout.	Late-visit status does not establish uninterrupted response or treatment-free remission.
COMP001 [4]. Twelve weeks.	Sustained response required $\geq 50\%$ MADRS reduction at week 3 and subsequent assessments. Rates for 25/10/1 mg: 16/79, 4/75, and 8/79. OR for 25 mg versus 1 mg: 2.2 (95% CI, 0.9 to 5.4).	Withdrawals by week 12: 5/79, 9/75, and 10/79, respectively.	New depression treatment before week 3: 4/79, 9/75, and 14/79. After week 3 through week 12: 26/79, 18/75, and 16/79. No protocol redosing.	Descriptive secondary outcome. A between-group difference in sustained response was not established. New treatment counted against sustained response.

Table 4. Cont.

Study and Assessment	Durability Outcome and Finding	Follow-Up Completeness	Subsequent Treatment and Psychedelic Exposure	Interpretation
Rosenblat et al., 2024 [26]. Six months.	Flexible retreatment after previous benefit, acceptable tolerability, and depressive relapse symptoms lasting at least 2 weeks. Additional doses were associated with further MADRS reductions.	Twenty-one of 30 treated participants completed 6 months. Eight withdrew after the week-2 assessment to begin other treatment.	Seventeen received a second dose and five a third, each with psychotherapy.	Clinician-selected retreatment and withdrawal for other treatment prevent estimation of single-dose durability or the independent benefit of redosing.
Aaronson et al., 2025 [27]. Twelve weeks.	Late-visit MADRS change: −17.2 points (95% CI, −25.2 to −9.1). Formal relapse endpoint: NR.	Eleven of 12 completed week 12.	Two restarted antidepressant regimens. One used nontrial psilocybin, and one restarted a cannabinoid. Additional psychotherapy exposure was not quantified.	Later symptom improvement was observed alongside subsequent treatment in some participants. Continuous response and treatment-free remission were not established.
Goodwin et al., 2025, COMP004 [42]. Up to 52 weeks after parent-trial dosing.	Prespecified time to a composite depressive event in all 233 COMP001 participants. Median times after 25/10/1 mg: 92, 83, and 62 days, respectively.	Sixty-six entered the extension: 58 from COMP001 and eight from COMP003. Forty-eight completed it, including 45/58 COMP001 entrants.	Subsequent treatment was unrestricted. New depression treatment in the COMP001 cohort: 38/79, 29/75, and 34/79 across follow-up. Further psychedelic exposure: NR.	Descriptive time-to-event findings. The composite included treatment initiation and deterioration, rather than relapse after remission alone. Selective extension entry and censoring limit inference.
Mertens et al., 2026, EPISODE follow-up [43]. Six and 12 months after the first administration.	HAMD-17 point response: 43/126 (34.1%) and 47/126 (37.3%). Response sustained from week 12: 28/126 (22.2%) and 20/126 (15.9%), respectively. Sustained HAMD-17 remission from week 12: 20/126 (15.9%) and 10/126 (7.9%), respectively, calculated from the published arm-level counts.	Follow-up cohort: 126/144 randomized participants with at least one follow-up assessment. HAMD-17 data: 118/126 at 6 months and 117/126 at 12 months, with 8 and 9 assessments missing, respectively.	By 12 months, 40/126 had initiated antidepressants, 32/126 had used classical psychedelics, and 86/126 had attended an optional integration group.	All follow-up participants had received at least one 25 mg dose. Subsequent treatment and incomplete follow-up limit attribution to the original course.
COMP005 [36]. Twenty-six weeks. Sponsor announcement: 17 February 2026.	The sponsor reported that 25% of the 25 mg arm achieved ≥25% MADRS reduction at week 6, with subsequent benefit through week 26 after one or two doses. Exact outcome n/N and CI: NR.	Week-26 retention and outcome-specific denominators: NR in the cited communication.	A second dose was permitted. Additional antidepressant treatment, psychotherapy, and nontrial psychedelic exposure: NR.	Preliminary evidence concerning a treatment strategy allowing retreatment. The threshold is below conventional response and does not establish remission.

Table 4. Cont.

Study and Assessment	Durability Outcome and Finding	Follow-Up Completeness	Subsequent Treatment and Psychedelic Exposure	Interpretation
COMP006 [37]. Twenty-six weeks. Sponsor announcement: 7 July 2026.	The sponsor reported that 39% of the 25 mg arm achieved $\geq 25\%$ MADRS reduction at week 6, with benefit maintained on average through at least week 26. Exact outcome n/N and CI: NR.	Week-26 retention and outcome-specific denominators: NR in the cited communication.	Two initial administrations were followed by permitted retreatment in Part B. Additional antidepressant treatment, psychotherapy, and nontrial psychedelic exposure: NR.	Preliminary findings do not establish uninterrupted conventional response in 39% or durability after a single administration.

Notes. Point response describes status at one visit. Sustained response requires response at successive specified assessments. Neither establishes treatment-free remission, which requires remission over a defined interval without ongoing or additional depression treatment. EPISODE response and remission proportions, including sustained outcomes, use the follow-up cohort denominator (N = 126), consistent with Table 3 of the original publication [43]. Participants with missing HAMD-17 assessments remain in the denominator but do not contribute to the corresponding endpoint numerator. Missing scores were not imputed. These proportions represent documented outcomes; missing assessments do not establish nonresponse or relapse. Phase 3 findings were preliminary and not peer-reviewed at the sponsor-source cut-off of 9 August 2026. NR does not establish absence of treatment or events. Abbreviations: CI, confidence interval; HAMD-17, 17-item Hamilton Depression Rating Scale; MADRS, Montgomery-Åsberg Depression Rating Scale; NR, not reported in the reviewed sources; OR, odds ratio; QIDS-SR-16, 16-item Quick Inventory of Depressive Symptomatology, self-report version.

5. Interpreting the Treatment Effect

5.1. Comparator Choice and Functional Unblinding

A credible placebo for a conspicuous psychoactive intervention remains unresolved. Inert placebo maximizes pharmacological contrast but offers little protection against treatment guessing. Very low-dose psilocybin may produce sensations that support masking, yet it can have biological or subjective effects. Niacin and other active placebos reproduce selected bodily sensations without reproducing the psychedelic state. Each comparator changes the estimand.

Treatment guessing can influence symptom reporting, clinician ratings, and the responses of session staff. Nominal masking of outcome assessors does not establish that expectancy effects have been controlled [31]. Future trials should measure expectancy and prior psychedelic exposure before randomization and obtain treatment guesses separately from participants, session monitors, and outcome assessors at prespecified, repeated time points. Participants and monitors should be assessed after acute effects have subsided and again during follow-up. Assessors should record their guesses after completing the relevant symptom rating. Each assessment should include confidence, an “uncertain” option, and the reasons for the guess. Results should be reported by randomized group, respondent role, and time point, with denominators and missing assessments.

Analyses should examine how treatment guesses relate to baseline expectancy, subsequent changes in expectancy, and later clinical outcomes, while documenting symptom changes that preceded each guess. These associations require cautious interpretation: improvement may influence a treatment guess, as well as follow from expectations associated with that guess. Treatment guess is a post-randomization variable. Adjusting for it, or restricting analyses to participants who guessed incorrectly, can remove part of the effect of treatment assignment or introduce selection bias, including collider bias [45]. The prespecified randomized comparison should therefore remain the primary analysis. Additional mediation or sensitivity analyses should state their causal estimand and assumptions ex-

plicitly; they cannot, by themselves, restore masking or establish an effect independent of expectancy.

Consent procedures should acknowledge that acute effects may reveal allocation without suggesting that a particular subjective experience is necessary for clinical benefit [46].

Future trials should use independent central outcome raters who are separate from dosing and support teams and have no access to session records, acute adverse-effect descriptions, or allocation guesses. Participants should receive neutral instructions to avoid disclosing these details during efficacy assessments, while safety information remains fully available to the treating team. Inadvertent disclosures should be documented. Planned psychological contact should be comparable across groups, with adherence to the support protocol assessed and additional contact recorded. Candidate active comparators should be evaluated for credibility and safety before confirmatory use, recognizing that their own activity changes the treatment contrast. These measures can reduce avoidable bias but cannot ensure equivalent expectancy or establish a pharmacological effect independent of the treatment context [8,31].

5.2. The Treatment Package

Preparation, music, room design, interpersonal conduct, crisis management, and post-session contact vary across studies. These variables can alter anxiety, disclosure, meaning-making, and retention. The Reporting of Setting in Psychedelic Clinical Trials (ReSPCT) guidelines identify 30 setting variables across the physical environment, dosing procedure, therapeutic framework, and subjective experience that should be reported [47]. A systematic review of psilocybin frameworks likewise found substantial variation in therapeutic models and their specification [48].

Psychological support should be described according to its purpose and content. Safety support includes monitoring, reassurance, management of acute distress, and access to medical care. Nonspecific therapeutic contact includes attention, continuity, empathic listening, and the therapeutic relationship. Formal psychotherapy involves specified techniques intended to change processes that maintain depression. These functions may overlap within a session, and the labels “preparation” and “integration” do not establish which were delivered. Inclusion of a defined psychotherapy within a study protocol does not establish its independent contribution to the observed treatment effect. A manual can standardize practice and support safeguarding [49], but its existence does not demonstrate adherence or establish the antidepressant efficacy of its psychological components.

Table S2 summarizes the preparation, dosing support, integration, staff qualifications, training, and fidelity procedures reported for COMP001, EPISODE, and PsiDeR. Protocol requirements are distinguished from evidence about delivery, and unavailable details are identified. Building on the Template for Intervention Description and Replication (TIDieR) checklist and ReSPCT guidelines, we propose a minimum reporting set for psychological support [47,50]. Preparation, dosing support, and integration should each be described by their purpose, content, delivery format, and planned session number, duration, and timing. Reports should specify staff roles, qualifications, training, supervision, and continuity across visits. Any manual and patient materials should be identified by version and made accessible. The description should distinguish safety procedures from psychotherapeutic techniques and specify how expectancy, acute distress, and interpersonal boundaries are addressed.

For each treatment arm, reports should provide actual attendance and contact time, additional support, concurrent psychotherapy, protocol deviations, and adaptations across sites. Fidelity reporting should include the assessment method, sessions sampled, assessor independence, and adherence results. The existence of a manual or completion of staff

training does not establish that the intervention was delivered as intended. This reporting set is intended to improve reproducibility and does not establish the minimum effective amount of psychological support. Randomized comparisons of psychological components, including factorial or dismantling designs where feasible, are needed to estimate their contribution while retaining necessary safety support. Current efficacy findings concern drug administration within the tested clinical setting and support arrangements; they do not establish the outcome of administering psilocybin without those arrangements.

5.3. Population, Medication Management, and Generalizability

Published samples are narrower than routine TRD populations. Trials commonly exclude psychotic disorders, bipolar I disorder or mania, active substance dependence, acute suicide risk, unstable cardiovascular disease, and clinical circumstances that would make medication withdrawal or a prolonged session unsafe [4–6,25–28]. Participants are also unusually willing to receive a psychedelic intervention and may enter with strong expectations. The resulting evidence applies most directly to carefully screened adults treated in specialized programs.

Assessment of apparent treatment resistance was not uniform. Diagnostic interviews and reviews of treatment history addressed some causes of pseudo-resistance, but documentation of prescribed dose and duration does not itself verify adherence. The reviewed reports do not establish a common procedure for excluding pseudo-resistance or independently confirming adherence to previous treatments. Nonresponse despite adequate exposure should also be distinguished from inability to complete treatment because of adverse effects. The published PsiDeR protocol, for example, allowed intolerance to contribute to its definition of treatment failure [29].

Assessment of psychiatric comorbidity also differed. COMP001 used a structured diagnostic interview and the McLean Screening Instrument for Borderline Personality Disorder, with specified personality disorders and recent alcohol or substance abuse among its exclusions [4,30]. EPISODE used the Structured Clinical Interview for DSM-5 Personality Disorders and reported personality disorders in 27/144 participants, despite excluding cluster A and borderline personality disorders. Its protocol specified the Childhood Trauma Questionnaire, excluded post-traumatic stress disorder and current substance use disorder, and included urine drug screening [5]. PsiDeR excluded diagnosed personality disorders and drug or alcohol dependence, although the published reports did not describe a separate dimensional assessment of personality pathology or trauma-related symptoms [6,29].

These procedures should not be interpreted as equivalent assessments of clinical complexity. Excluding a diagnosis, recording childhood adversity, and measuring current trauma-related symptoms address different questions. Likewise, a negative drug screen does not characterize longitudinal substance use. The available reporting does not permit consistent comparisons of these factors as modifiers of efficacy or safety. Their limited characterization should be considered when applying trial findings to routine TRD services.

Antidepressant discontinuation is both a clinical hazard and an interpretive problem. Withdrawal symptoms can worsen mood and anxiety near baseline, while prior serotonergic treatment may alter acute drug effects. An open-label study of 19 participants receiving psilocybin while continuing a selective serotonin reuptake inhibitor found no serious adverse events and an antidepressant signal [51]. Exploratory post hoc analyses of antidepressant discontinuation have yielded conflicting findings. In a major depressive disorder (MDD) trial, participants who had discontinued SSRIs or serotonin–norepinephrine reuptake inhibitors (SNRIs) before psilocybin showed smaller treatment effects than participants who entered the trial unmedicated, although causality could not be inferred [52]. By contrast, a post hoc analysis of the COMP001 TRD dataset found no evidence that

antidepressant discontinuation worsened baseline depression or compromised psilocybin efficacy [53]. A 2025 scoping review found no consistent serotonin-syndrome signal with concomitant classic psychedelics, especially psilocybin, but the evidence was small and heterogeneous [54]. Routine continuation, mandatory washout, and taper duration therefore remain trial questions rather than settled rules.

Within the literature reviewed, we identified no head-to-head study comparing psilocybin with electroconvulsive therapy, repetitive transcranial magnetic stimulation, ketamine or esketamine, evidence-based augmentation, or an optimized psychotherapy pathway. Separate placebo-controlled trials cannot determine sequence, substitution, or comparative value. Current data do not support placing psilocybin ahead of established TRD treatments or restricting it to a precise line of therapy.

6. Safety and Patient Selection

The common acute adverse effects are predictable but clinically relevant: headache, nausea, anxiety, dizziness, transient increases in blood pressure and heart rate, and short periods of fear, confusion, or perceptual disturbance. A systematic review identified 214 unique studies of classic psychedelics, of which 114 reported analyzable adverse-event data. Serious adverse events were uncommon in monitored research settings. However, adverse-event ascertainment was frequently incomplete, and events requiring medical or psychiatric attention did occur [55].

A review focused on psilocybin trials reached a similar conclusion: adverse-event definitions, collection windows, and management were inconsistent [56]. EPISODE illustrates the distinction between acute tolerability and serious psychiatric harm. Across two dosing sessions, 143 participants received 280 administrations. Severe adverse events on dosing days were reported after 44/160 administrations of 25 mg psilocybin (27.5%), 3/72 administrations of 5 mg (4.2%), and 4/48 administrations of nicotinamide (8.3%). These denominators represent administrations, not unique participants or parallel randomized groups. The trial also reported two serious adverse reactions following 25 mg, including one case of hallucinogen-persisting perception disorder (HPPD) with broader psychological destabilization [5].

Future randomized trials should follow the Consolidated Standards of Reporting Trials (CONSORT) Harms 2022 recommendations [57]. Reports should specify how events were elicited and coded, define observation windows, and distinguish severity, seriousness, and assessed relatedness. Results should include affected participants and denominators by treatment arm, total event counts, and events leading to withdrawal, rescue medication, or hospitalization. With repeated dosing, per-administration findings should accompany cumulative participant-level risk. Acute effects should be distinguished from delayed psychiatric and perceptual symptoms, with onset, duration, resolution, and follow-up completeness reported. We additionally recommend independent review of serious psychiatric events. Missing safety information should be distinguished from confirmed absence of events.

Suicidality requires separate reporting of treatment-emergent ideation, worsening on a standardized scale, suicidal behavior, and serious adverse events. Treatment-emergent ideation denotes new or worsened suicidal thoughts after dosing and does not establish causation. A higher score on the Columbia-Suicide Severity Rating Scale (C-SSRS) describes a change in the assessed suicidal state; it should not be equated with an attempt. Preparatory behavior, interrupted or aborted attempts, actual attempts, and nonsuicidal self-injury should be identified separately where the source permits. Seriousness depends on regulatory and protocol criteria, including hospitalization or immediate danger to life,

and is distinct from symptom intensity. A serious adverse reaction additionally implies a suspected causal relationship with treatment [4,5].

Table S3 reports absolute counts, denominators, assessment windows, baseline risk, and ascertainment methods for the randomized TRD trials. These findings warrant continued monitoring beyond the dosing session. Neither similar event frequencies between groups nor improvement in mean suicidality scores excludes clinically important deterioration in individual patients. The trials were not designed to estimate uncommon suicidal outcomes precisely, and selection against acute suicide risk limits their applicability to patients requiring urgent intervention.

An open-label study enrolled 20 adults with chronic suicidal ideation, MDD, and at least two prior antidepressant treatment failures, while excluding recent ideation with a specific plan and imminent intent [58]. Following one 25 mg administration, mean Modified Scale for Suicidal Ideation (MSSI) scores improved through 12 weeks. However, 2/20 participants had scores above baseline during follow-up: one transiently at week 3 and one at week 12, when the score increased from 17 to 23. No SAEs were reported among the 20 participants. Study clinicians assessed suicidality using the MSSI and C-SSRS; AE relatedness was investigator-assessed. These individual deteriorations should be distinguished from both the improvement in group means and the absence of SAEs. The uncontrolled design and initiation or resumption of psychotropic treatment in 12/20 participants at or after week 3 limit causal interpretation.

For future delivery programs, we propose structured suicide-risk assessment at screening, during any medication taper, immediately before dosing, before discharge once acute drug effects have subsided, and at scheduled or symptom-triggered follow-up. The C-SSRS should supplement clinical assessment of suicidal intent, planning, access to means, previous behavior, and available support [5,59,60]. New suicidal intent or inability to remain safe should prompt immediate psychiatric assessment and a documented escalation plan. Improvement in depression scores should not replace assessment of the individual patient's current safety.

Screening in a future clinical service would need to verify the diagnosis and adequacy of previous treatment. Assessment should cover mania, psychosis, substance use, cognition, trauma-related vulnerability, suicide risk, cardiovascular status, pregnancy risk, medication interactions, and social support. It should also establish whether the patient can give informed consent for an intense and partly unpredictable experience. Family history of psychosis or bipolar disorder, unstable personality organization, and prior adverse psychedelic experiences require individualized judgment because direct trial evidence is sparse. Established safety guidance remains conservative [59,60]. A 2026 systematic review of 23 human studies found a low but clinically meaningful risk of transient hypomanic or manic symptoms in susceptible individuals, with higher-risk signals in bipolar I disorder, familial vulnerability, polysubstance use, and unsupervised use. Long-term outcomes and repeated exposure remain insufficiently characterized [61].

Pregnant and breastfeeding patients should be excluded from routine psilocybin treatment because reproductive and infant safety have not been established. Available human evidence is insufficient to characterize pregnancy outcomes, exposure through breast milk, or effects on the nursing infant [62]. Screening should address pregnancy intentions and breastfeeding, with pregnancy testing before each administration when applicable and contraceptive counselling consistent with the study protocol [5,6]. Consent should explain these uncertainties and discuss established alternatives for treating depression during pregnancy and lactation. No evidence-based interval for interrupting and resuming breastfeeding can currently be specified. Inadvertent exposure warrants individualized obstetric or teratology advice and, during breastfeeding, assessment of potential infant exposure.

The July 2026 US Food and Drug Administration guidance for psychedelic clinical investigations recommends observation by two monitors throughout dosing. The lead monitor should be an independently licensed healthcare professional with graduate-level training and psychotherapy experience. If the lead monitor is not a physician, an on-call physician should be able to reach the site within 15 min [8]. These recommendations concern clinical trials. Drawing on this guidance and established research safety recommendations [60], we propose minimum infrastructure for future authorized clinical services. Consent should address potentially prolonged psychological effects, increased vulnerability, and suggestibility.

Services should provide a private room with environmental hazards minimized, continuous observation, equipment for measuring blood pressure, pulse, and oxygen saturation, basic resuscitation equipment, and staff trained in basic life support. Written protocols should specify who can prescribe and administer rescue medication, indications for urgent medical or psychiatric assessment, and arrangements for emergency transport and hospital transfer. Persistent severe agitation, psychosis, suicidal intent, or cardiovascular instability should trigger prompt assessment and escalation. Discharge should require recovery of orientation, clinically stable observations, an individualized safety assessment, an accompanying adult, and written follow-up and emergency instructions. Patients who do not meet discharge criteria require continued observation or transfer to an appropriate care setting.

Postdischarge care should include planned contact rather than relying solely on patients to report difficulties. A reasonable service model would include contact within 24 h and reassessment during the first week, with additional reviews according to symptoms and baseline risk [60]. This schedule is a clinical proposal and has not been established by comparative trials. Reviews should assess sleep, anxiety, derealization, perceptual symptoms, mood elevation, suicidal thoughts, medication changes, and substance use. Mild symptoms with preserved insight may be managed with reassurance, grounding, sleep support, and closer follow-up. Persistent or worsening symptoms require clinical reassessment. Reduced need for sleep with mood elevation, psychosis, suicidal intent, or impaired ability to remain safe warrants urgent psychiatric assessment [59,60]. Persistent perceptual symptoms require evaluation for HPPD and alternative psychiatric, neurological, or ophthalmic causes. Patients should receive written contact instructions and a clear route to urgent care. Unresolved symptoms should be followed beyond the first week, and further dosing deferred pending assessment.

Therapist conduct is part of safety. Altered judgment, emotional intensity, and dependence on session staff create an unusual power imbalance. Training must cover boundaries, touch, recording and chaperone policies, adverse-event recognition, and routes for reporting misconduct [63]. Safety cannot be inferred from the molecule alone when containment depends on the people and system around it.

7. Clinical Delivery and Place in Care

If approved, psilocybin would function more like a scheduled procedure than a take-home prescription. A treatment episode includes diagnostic confirmation, medication planning, preparation, a 6 to 8 h dosing visit, continuous observation, discharge assessment, transport home, integration or follow-up, and access to urgent psychiatric care. The room, staff, documentation, and emergency pathway are parts of the intervention. Clinical translation cannot assume that efficacy from high-contact trials will persist after these elements are reduced to improve throughput.

This model creates practical trade-offs. Under the FDA clinical-development framework, dosing sessions involve two monitors [8]. Comparable requirements in routine care

would commit substantial staff time and room capacity to each administration. Patient burden includes preparation, prolonged observation, transport, recovery time, follow-up, and medication tapering when required by the protocol. Tapering also requires assessment of withdrawal symptoms and clinical deterioration [51–54]. Evaluation should therefore include functioning, quality of life, participation in work and usual activities, and the patient's own treatment goals. Treatment burden should be measured prospectively through time spent attending and recovering from treatment, travel, caregiver involvement, out-of-pocket expenditure, and treatment discontinuation attributable to these demands. These outcomes should be assessed over a common follow-up period that includes retreatment and additional care. Symptom improvement alone cannot establish clinical value or cost-effectiveness. Programs also require controlled-drug governance, staff supervision, and review of psychological and medical adverse events. Access and continuity of care remain central implementation concerns [7,63,64].

Regulatory status and medical access to psilocybin for TRD vary internationally. Australia permits psilocybin-assisted psychotherapy for TRD under its Authorised Prescriber scheme, although the products remain unapproved therapeutic goods [65]. In Canada, access to psilocybin may be granted through the Special Access Program on a case-by-case basis [66], whereas Switzerland permits restricted medical use through individual exceptional authorisations [67]. In Czechia, a legal framework effective from 1 January 2026 permits psilocybin-assisted psychotherapy using individually prepared medicinal products in specialised healthcare facilities, including for clinically significant depressive disorder after documented failure or intolerance of prior treatment [68]. These forms of legal access should be distinguished from marketing authorisation. Psilocybin remains subject to Schedule I controls under the 1971 Convention on Psychotropic Substances, which restrict medical use and require specific authorisations [69]. The need to manage acute psychological distress and provide continuous observation during prolonged dosing sessions adds to the requirements for safe implementation [8]. Availability also depends on trained clinical teams, suitable facilities, emergency support and funding for preparation, dosing and follow-up. In its 23 March 2026 statement, the Czech National Institute of Mental Health identified finalisation of reimbursement arrangements as a prerequisite for clinical rollout [70].

Liechti et al. describe the Swiss limited medical use program, established in 2014, with psilocybin added in 2021 [71]. In 2024, authorisations covered 348 patients treated with psilocybin. The report provides practical information on physician training, treatment organisation and financing. The clinical population included several diagnoses, and psilocybin outcome data were limited to 21 feedback reports within the registry and quality-assessment program. These observations inform service development but do not establish the efficacy or durability of psilocybin specifically in TRD.

The appropriate position of psilocybin in a TRD pathway remains unknown. Established interventions have more developed frameworks for monitoring and continuation care, although maintenance evidence varies by treatment and population [72–77]. Table 5 summarizes these differences without ranking efficacy. For psilocybin, retreatment criteria, concomitant medication management, and cumulative safety remain unresolved. Current evidence does not establish superiority or a preferred treatment sequence relative to established options. If authorized, initial delivery in specialist services could support careful selection and follow-up, but this would be a service-design judgment rather than a validated treatment sequence. Treatment decisions should consider clinical urgency, previous treatment, contraindications, patient preference, and the feasibility of completing the proposed care pathway.

Table 5. Delivery requirements, maintenance evidence, and unresolved questions for interventions considered in treatment-resistant depression.

Intervention	Route and Delivery	Monitoring and Patient Burden	Maintenance Framework and Evidence	Selected Contraindications or Precautions	Unresolved Questions
Psilocybin administered with structured psychological support [4–8,37–43,59,60]	Oral administration within a program of preparation, supervised dosing, and follow-up.	Dosing visits commonly last 6–8 h; continuous observation, transport home, and subsequent contact are required. Medication tapering may add burden.	Follow-up studies and retreatment programs are available; no validated maintenance schedule or criteria for repeat dosing.	Typical trial exclusions include psychotic or bipolar disorders, unstable cardiovascular illness, pregnancy, and breastfeeding.	Retreatment timing, cumulative safety, concomitant medication management, and sustained functional benefit.
Electroconvulsive therapy [72,73]	Electrical stimulation under general anaesthesia and muscle relaxation.	Anaesthetic and cardiovascular monitoring, repeated cognitive assessment, recovery time, and transport arrangements.	Continuation pharmacotherapy and/or continuation or maintenance ECT after acute benefit; treatment frequency is individualized.	High anaesthetic or cardiovascular risk; vulnerability to cognitive adverse effects.	Selection of maintenance strategy, optimal frequency and duration, and balance between relapse prevention and cognitive burden.
Repetitive transcranial magnetic stimulation [74,75]	Outpatient magnetic stimulation without general anaesthesia.	Frequent attendance during the acute course; monitoring of tolerability and seizure risk; hearing protection.	Randomized maintenance data include a 24-week trial; protocols and schedules remain heterogeneous.	Incompatible metal or electronic implants near the stimulation field; seizure risk requires individual assessment.	Optimal maintenance schedule, duration, and criteria for reintroducing an acute course.
Intravenous racemic ketamine [72,76]	Intravenous infusions, usually delivered as a repeated treatment course.	Blood pressure, pulse, respiratory status, dissociation, and recovery monitoring; transport after treatment.	Repeated-infusion strategies are used, but controlled maintenance evidence and standardized schedules remain limited.	Poorly controlled hypertension, unstable cardiovascular or respiratory disease, psychotic symptoms, and misuse risk require assessment.	Maintenance frequency and duration; cumulative urinary, cognitive, and misuse risks.
Intranasal esketamine [77]	Supervised intranasal administration with induction and maintenance phases.	Under the US label, at least 2 h of observation, including respiratory monitoring; blood pressure and mental-state assessment; transport arrangements.	Randomized withdrawal evidence in stabilized patients receiving an oral antidepressant; labeled maintenance dosing is weekly, then weekly or every 2 weeks according to response.	Contraindications include aneurysmal vascular disease, arteriovenous malformation, previous intracerebral haemorrhage, and hypersensitivity.	Optimal treatment duration, discontinuation strategy, and cumulative patient burden.

Table 5. *Cont.*

Intervention	Route and Delivery	Monitoring and Patient Burden	Maintenance Framework and Evidence	Selected Contraindications or Precautions	Unresolved Questions
Lithium augmentation [72,73]	Daily oral medication added to antidepressant treatment.	Serum lithium concentrations; renal, thyroid, and calcium monitoring; assessment of interactions and toxicity.	Continuation evidence includes relapse prevention after ECT; applicability depends on the population and prior treatment.	Renal impairment, dehydration, and interacting medication that increases lithium concentrations.	Optimal duration and benefit–risk balance in different TRD populations.
Atypical-antipsychotic augmentation [72,73]	Daily oral medication; agent and dose selected individually.	Weight, metabolic measures, movement symptoms, and agent-specific cardiac monitoring; ongoing medication burden.	Acute augmentation evidence is stronger than evidence defining the optimal duration of maintenance treatment in TRD.	Agent-specific metabolic risks, QT prolongation, akathisia, and other movement disorders.	Duration of continuation, withdrawal strategy, and cumulative metabolic and neurological harm.
Evidence-based psychotherapy [72,73]	Structured individual or group sessions, including CBT; relapse-prevention approaches include CBT and MBCT.	Repeated attendance, between-session practice, and monitoring of symptoms, functioning, and risk.	Relapse-prevention evidence is established in recurrent depression; evidence specific to strictly defined TRD is less extensive.	No general medical contraindication; outpatient treatment requires an appropriate plan for acute risk and capacity to participate.	Optimal intensity, duration, and sequencing or combination with biological treatment in TRD.

The table describes care requirements and the maturity of maintenance evidence; it does not provide an indirect efficacy comparison or a recommended treatment sequence. Study populations and definitions of treatment resistance differ. Listed precautions are not uniformly absolute contraindications, and psilocybin trial exclusions should not be interpreted as validated prescribing criteria. Esketamine monitoring and dosing details refer to the US prescribing information revised in March 2026; local requirements may differ. CBT, cognitive behavioural therapy; MBCT, mindfulness-based cognitive therapy; TRD, treatment-resistant depression.

8. Priorities for Confirmatory Research

The next trials should be designed around decisions faced in practice. Confirmatory efficacy studies need a justified comparator, prespecified assessment of masking, centralized ratings, and both continuous and categorical outcomes. Functioning, quality of life, work participation, health service use, and treatment burden should accompany symptom scales. Analyses should distinguish participants who never respond, respond and relapse, respond only after retreatment, and maintain treatment-free remission.

Head-to-head trials should compare complete treatment pathways, including preparation, medication changes, acute treatment, maintenance, and rescue care over the same follow-up period. Sequencing trials should establish when psilocybin should be considered after an inadequate response to established interventions. Alongside symptoms and relapse, these studies should assess functioning, quality of life, patient and caregiver time, treatment discontinuation, and health-service costs.

For future phase 3 trials, we propose a common set of durability outcomes over a prespecified follow-up period, preferably 12 months. Cumulative time in remission should estimate the total days or weeks spent in remission, including remission regained after relapse, across the full randomized population. Sustained remission should require

the remission criterion to remain satisfied across a predefined sequence of assessments. Relapse-free survival should measure time from a specified starting point after confirmed remission to a prospectively defined relapse event. These outcomes require explicit symptom thresholds, confirmation rules, assessment intervals, and methods for handling missing observations. Isolated late visits cannot reliably establish cumulative remission time or uninterrupted remission.

Comparative maintenance trials should randomize participants after confirmed remission, with maintenance randomization defining the start of follow-up. Restricting an initial randomized comparison to participants who subsequently remit does not preserve the original randomization. Protocols should distinguish outcomes under a treatment strategy permitting rescue or retreatment from remission maintained without additional treatment. The former requires continued assessment after these interventions. For the latter, additional treatment should constitute a prespecified failure event rather than trigger routine censoring. Reports should provide absolute relapse-free probabilities, between-group differences in cumulative remission time, confidence intervals, and sensitivity analyses for missing data [8,41].

Studies of concomitant antidepressants should compare continuation with carefully managed tapering rather than excluding the question through protocol. Enrolment should broaden only alongside adequate safeguards and prospective assessment of mania, psychosis, persistent perceptual symptoms, suicidality, and interpersonal harms.

In confirmatory TRD efficacy trials, candidate mechanistic biomarkers should currently be included only as prespecified exploratory outcomes or within nested mechanistic substudies. Protocols should define candidate measures, baseline and post-dose sampling times, standardized acquisition and assay procedures, and methods for addressing multiple testing and missing data. Analyses should distinguish baseline predictors from treatment-induced changes and document medication exposure and relevant clinical characteristics. Proposed predictors should demonstrate reproducible performance beyond standard clinical variables in an independent TRD sample. Claims that a baseline biomarker identifies preferential benefit from psilocybin should be evaluated through a prespecified biomarker-by-treatment interaction in randomized data [19]. Primary efficacy conclusions should remain based on clinical outcomes. Unvalidated biomarker results should not determine eligibility, dosing, or retreatment. Psychological support should be manualized and fidelity-rated, with enough detail to distinguish safety support from psychotherapy. Table 6 links the main inferential problems to the evidence required for clinical decisions.

Table 6. Design questions that determine clinical interpretation.

Domain	What Current Evidence Permits	Main Inferential Problem	Evidence Needed for a Clinical Decision
Comparator and masking	Short-term separation has been observed against 1 mg psilocybin and inert placebo; EPISODE missed its binary endpoint against nicotineamide.	The acute state reveals allocation, and each comparator estimates a different contrast.	Repeated treatment guesses with confidence and reasons from participants, session staff, and raters; baseline and follow-up expectancy; associations with subsequent outcomes; a justified comparator; and sensitivity analyses that explicitly address the risks of conditioning on post-randomization variables.
Endpoint and estimand	Continuous symptom change may favor 25 mg when a categorical response endpoint does not.	A threshold discards information; late mean scores do not show continuous individual benefit.	Continuous and categorical outcomes, cumulative time in remission, sustained remission, functioning, and rescue treatment reported together.

Table 6. *Cont.*

Domain	What Current Evidence Permits	Main Inferential Problem	Evidence Needed for a Clinical Decision
Psychological support	Preparation and in-session support are components of contemporary clinical protocols, while post-session support or integration is common but varies across studies.	Safety support and psychotherapy are usually bundled, variably manualized, and rarely compared.	Planned and delivered preparation, dosing, and integration contact; staff qualifications and training; manual content; fidelity methods and results; additional support and protocol deviations; and randomized evaluation of psychological components with necessary safety support retained.
Durability and retreatment	Some participants remain improved for months; sustained individual response is less common than late-visit response.	Naturalistic follow-up allows additional treatment; optimal redosing and cumulative risk are unknown.	Prespecified cumulative time in remission, sustained remission, and relapse-free survival. Explicit definitions, assessment schedules, time origins, and handling of rescue treatment, retreatment, and missing data. Randomized maintenance comparisons after confirmed remission, with cumulative exposure and safety reported alongside efficacy.
Medication management	Small studies suggest concomitant SSRIs may be feasible; attenuation of acute effects is inconsistent.	Tapering can cause withdrawal or symptom worsening and may alter baseline; continuation may change exposure.	Randomized continuation versus standardized tapering, with withdrawal measures, pharmacokinetics, and interaction surveillance.
Patient selection	Evidence applies mainly to screened adults without psychosis, mania, acute high suicide risk, or unstable medical illness.	Common comorbidity and greater clinical acuity are underrepresented; no responder profile is validated.	Pragmatic cohorts with staged inclusion, prespecified subgroup analyses, and external validation of any prediction model.
Safety ascertainment	Common acute events are expected; serious events are uncommon in specialist settings.	Rare psychiatric, cardiovascular, perceptual, and interpersonal harms require much larger denominators and longer follow-up.	Harmonized definitions, active surveillance, independent adjudication, registries, and reporting of context and management.
Place in care and value	The intervention can be delivered in a public health setting with high retention and specialist staffing.	No head-to-head evidence, uncertain capacity, and costs driven by staff time and retreatment.	Comparative-effectiveness trials, resource-use measurement, patient-reported burden, and cost per sustained functional recovery.

9. Limitations of This Review

This is a structured narrative review, not a systematic review. The literature search was iterative and question-driven, and study selection and weighting involved clinical judgment. No formal duplicate screening, PRISMA flow diagram, risk-of-bias assessment, or certainty framework was used. The review therefore cannot be interpreted as an exhaustive systematic identification of every publication relevant to psilocybin and depression. Rapid publication in 2026 also means that bibliographic details and development status can change after the prespecified evidence cut-off. Our appraisal of phase 3 efficacy and durability was restricted to sponsor communications available by 9 August 2026. Post-cut-off sponsor communications identified during revision were used only to document that new information had appeared and to recheck investigational status; they were not incorporated into the clinical evidence synthesis. The limited methodological detail in sponsor communications prevented independent assessment of complete analyses and whether all relevant efficacy and safety findings had been reported. Selective reporting

could affect the apparent benefit–harm balance. These limitations reduced the weight assigned to phase 3 findings in the clinical conclusions of this review. Heterogeneity in populations, comparators, support models, endpoints, and follow-up made quantitative pooling unlikely to answer the clinical questions emphasized here.

10. Conclusions

Psilocybin administered with structured psychological support shows short-term antidepressant activity in some adults with TRD. Sponsor-reported phase 3 findings provide additional preliminary evidence, but complete reporting and independent appraisal are required before they can establish clinical benefit, durability, or safety. The size of the specific drug effect is uncertain because treatment guessing, comparator choice, and the surrounding care model materially influence the estimate. Published long-term studies show that group means can remain improved, but they do not establish uninterrupted, treatment-free benefit for most patients. No biomarker selects responders, no retreatment schedule has been validated, and no trial defines where psilocybin belongs relative to established TRD treatments.

The evidence is mature enough to justify confirmatory and implementation research but not to simplify psilocybin into a rapid-acting prescription. Its clinical value will depend on whether reproducible benefit can be delivered as a safe procedure, maintained over the course of a recurrent illness, and achieved at a resource burden that health systems and patients can sustain.

Further research on the medical use of psychedelics would benefit from greater international coordination and at least partial harmonisation of research standards and regulatory evidence requirements. For psilocybin in TRD, this should include consistent definitions of treatment resistance, clear specification of psychological support, and common methods for assessing sustained clinical benefit and psychiatric adverse events. Agreement on these minimum standards would improve comparability across studies and provide a more consistent basis for decisions about clinical use while allowing for differences between national healthcare systems.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/therapeutics3040021/s1>, Table S1. Symptom scores and efficacy outcomes in the principal randomized trials of psilocybin for treatment-resistant depression; Table S2. Psychological support and psychotherapy procedures reported for COMP001, EPISODE, and PsiDeR; Table S3. Suicidality and selected serious adverse-event findings in randomized trials of psilocybin for TRD. Reference [78] is cited in the Supplementary Materials.

Author Contributions: Conceptualization and design, L.S. and M.P.; literature analysis and interpretation, L.S., M.P., D.J.-M., I.N., B.R. and W.J.C.; drafting of the manuscript, L.S. and M.P.; critical revision for important intellectual content, L.S., M.P., D.J.-M., I.N., B.R. and W.J.C.; supervision, L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This article is a narrative review of previously published work.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: During preparation of this manuscript, the authors used Quillbot v. 3.1.1 to assist with language editing. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the submitted manuscript.

Conflicts of Interest: Wiesław Jerzy Cubała declares grants/research support from Acadia, Alkermes, Allergan, Auspex Pharmaceuticals, BMS, Celon, Cephalon, Cortexyme, Ferrier, Forest Laboratories, Gedeon Richter, GW Pharmaceuticals, KCR, Lilly, Lundbeck, Minerva, NIH, Orion, and Servier; honoraria from Adamed, AstraZeneca, BMS, Celon, GSK, KRKA, Lekam, Lundbeck, Minerva, NeuroCog, Orion, Pfizer, Polfa Tarchomin, Sanofi, Servier, Zentiva, and Recognify Life Sciences; advisory-board participation for Celon (terminated) and Sanofi; and consultancy for Douglas Pharmaceuticals, GH Research, and Janssen. All other authors declare no conflicts of interest. The disclosed relationships had no role in the study design, literature selection, analysis, interpretation, manuscript preparation, or decision to submit the article for publication.

Abbreviations

The following abbreviations are used in this manuscript:

5-HT2A	5-hydroxytryptamine receptor 2A
BDI	Beck Depression Inventory
BDNF	brain-derived neurotrophic factor
CI	confidence interval
EPISODE	Efficacy and Safety of Psilocybin in Treatment-Resistant Major Depression
FDA	US Food and Drug Administration
HAMD-17	17-item Hamilton Depression Rating Scale
HPPD	hallucinogen-persisting perception disorder
MADRS	Montgomery-Åsberg Depression Rating Scale
MDD	major depressive disorder
OR	odds ratio
PsiDeR	Psilocybin in Depression Resistant to Standard Treatments
QIDS-SR-16	16-item Quick Inventory of Depressive Symptomatology-Self-Report
RCT	randomized controlled trial
ReSPCT	Reporting of Setting in Psychedelic Clinical Trials
SNRI	serotonin-norepinephrine reuptake inhibitor
SSRI	selective serotonin reuptake inhibitor
TRD	treatment-resistant depression
TrkB	tropomyosin receptor kinase B

References

1. McAllister-Williams, R.H.; Arango, C.; Blier, P.; Demyttenaere, K.; Falkai, P.; Gorwood, P.; Hopwood, M.; Javed, A.; Kasper, S.; Malhi, G.S.; et al. The identification, assessment and management of difficult-to-treat depression: An international consensus statement. *J. Affect. Disord.* **2020**, *267*, 264–282. [[CrossRef](#)] [[PubMed](#)]
2. Rush, A.J.; Trivedi, M.H.; Wisniewski, S.R.; Nierenberg, A.A.; Stewart, J.W.; Warden, D.; Niederhe, G.; Thase, M.E.; Lavori, P.W.; Lebowitz, B.D.; et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *Am. J. Psychiatry* **2006**, *163*, 1905–1917. [[CrossRef](#)] [[PubMed](#)]
3. Fonzo, G.A.; Wolfgang, A.S.; Barksdale, B.R.; Krystal, J.H.; Carpenter, L.L.; Kraguljac, N.V.; Grzenda, A.; McDonald, W.M.; Widge, A.S.; Rodriguez, C.I.; et al. Psilocybin: From psychiatric pariah to perceived panacea. *Am. J. Psychiatry* **2025**, *182*, 54–78. [[CrossRef](#)] [[PubMed](#)]
4. Goodwin, G.M.; Aaronson, S.T.; Alvarez, O.; Arden, P.C.; Baker, A.; Bennett, J.C.; Bird, C.; Blom, R.E.; Brennan, C.; Brusch, D.; et al. Single-dose psilocybin for a treatment-resistant episode of major depression. *N. Engl. J. Med.* **2022**, *387*, 1637–1648. [[CrossRef](#)] [[PubMed](#)]
5. Mertens, L.J.; Koslowski, M.; Betzler, F.; Brand, M.; Evens, R.; Kärtner, L.; Jungaberle, A.; Jungaberle, H.; Majić, T.; Schmitz, C.N.; et al. Efficacy and safety of psilocybin in treatment-resistant major depression: The EPISODE randomized clinical trial. *JAMA Psychiatry* **2026**, *83*, 448–460. [[CrossRef](#)] [[PubMed](#)]
6. Rucker, J.J.; Mantingh, T.; Kerr-Gaffney, J.; Bird, C.; Chu, P.; Modlin, N.L.; Campbell-Coker, K.; Hambleton, S.; Seath, P.; Hignett, R.; et al. Psilocybin-assisted therapy for treatment-resistant major depressive disorder in a public healthcare setting: A randomized controlled trial. *Nat. Med.* **2026**, *32*, 3430–3437. [[CrossRef](#)] [[PubMed](#)]
7. Hosein, M.M.; Reid, M.J.; Walser, S.; Charney, S.; Fonzo, G.A.; Lewis, B.R.; Yaden, D.B.; Suppes, T.; Cordner, Z.A.; Barrett, F.S. Considerations and cautions for the integration of psilocybin into routine clinical care: A consensus statement from the US

- National Network of Depression Centers' Task Group on Psychedelics and Related Compounds. *EClinicalMedicine* **2025**, *89*, 103517. [CrossRef] [PubMed]
8. U.S. Food and Drug Administration. *Psychedelic Drugs: Considerations for Clinical Investigations. Final Guidance for Industry*; U.S. Food and Drug Administration (FDA): Silver Spring, MD, USA, 2026. Available online: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/psychedelic-drugs-considerations-clinical-investigations> (accessed on 9 August 2026).
 9. Nichols, D.E. Psychedelics. *Pharmacol. Rev.* **2016**, *68*, 264–355. [CrossRef] [PubMed]
 10. Ly, C.; Greb, A.C.; Cameron, L.P.; Wong, J.M.; Barragan, E.V.; Wilson, P.C.; Burbach, K.F.; Soltanzadeh Zarandi, S.; Sood, A.; Paddy, M.R.; et al. Psychedelics promote structural and functional neural plasticity. *Cell Rep.* **2018**, *23*, 3170–3182. [CrossRef] [PubMed]
 11. Shao, L.X.; Liao, C.; Gregg, I.; Davoudian, P.A.; Savalia, N.K.; Delagarza, K.; Kwan, A.C. Psilocybin induces rapid and persistent growth of dendritic spines in frontal cortex in vivo. *Neuron* **2021**, *109*, 2535–2544.e4. [CrossRef] [PubMed]
 12. Vargas, M.V.; Dunlap, L.E.; Dong, C.; Carter, S.J.; Tombari, R.J.; Jami, S.A.; Cameron, L.P.; Patel, S.D.; Hennessey, J.J.; Saeger, H.N.; et al. Psychedelics promote neuroplasticity through the activation of intracellular 5-HT_{2A} receptors. *Science* **2023**, *379*, 700–706. [CrossRef] [PubMed]
 13. Moliner, R.; Giry, M.; Brunello, C.A.; Kovaleva, V.; Biojone, C.; Enkavi, G.; Antenucci, L.; Kot, E.F.; Goncharuk, S.A.; Kaurinkoski, K.; et al. Psychedelics promote plasticity by directly binding to BDNF receptor TrkB. *Nat. Neurosci.* **2023**, *26*, 1032–1041. [CrossRef] [PubMed]
 14. Daws, R.E.; Timmermann, C.; Giribaldi, B.; Sexton, J.D.; Wall, M.B.; Erritzoe, D.; Roseman, L.; Nutt, D.; Carhart-Harris, R. Increased global integration in the brain after psilocybin therapy for depression. *Nat. Med.* **2022**, *28*, 844–851. [CrossRef] [PubMed]
 15. Carhart-Harris, R.L.; Roseman, L.; Bolstridge, M.; Demetriou, L.; Pannekoek, J.N.; Wall, M.B.; Tanner, M.; Kaelen, M.; McGonigle, J.; Murphy, K.; et al. Psilocybin for treatment-resistant depression: fMRI-measured brain mechanisms. *Sci. Rep.* **2017**, *7*, 13187. [CrossRef] [PubMed]
 16. Berkovitch, L.; Fauvel, B.; Preller, K.H.; Gaillard, R. Neurocognitive effects of psilocybin: A systematic and comprehensive review of neuroimaging studies in humans. *Neurosci. Biobehav. Rev.* **2025**, *175*, 106239. [CrossRef] [PubMed]
 17. Constantino, J.L.; van Dalen, J.H.; Massetti, S.; Kamphuis, J.; Schoevers, R.A. Neurobiological mechanisms of antidepressant properties of psilocybin: A systematic review of blood biomarkers. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2025**, *136*, 111251. [CrossRef] [PubMed]
 18. Cavaleri, D.; Moretti, F.; Bartocetti, A.; Mauro, S.; Crocamo, C.; Carrà, G.; Bartoli, F. The role of BDNF in major depressive disorder, related clinical features, and antidepressant treatment: Insight from meta-analyses. *Neurosci. Biobehav. Rev.* **2023**, *149*, 105159. [CrossRef] [PubMed]
 19. Ballman, K.V. Biomarker: Predictive or prognostic? *J. Clin. Oncol.* **2015**, *33*, 3968–3971. [CrossRef] [PubMed]
 20. Sloshower, J.; Zeifman, R.J.; Guss, J.; Krause, R.; Safi-Aghdam, H.; Pathania, S.; Pittman, B.; D'Souza, D.C. Psychological flexibility as a mechanism of change in psilocybin-assisted therapy for major depression: Results from an exploratory placebo-controlled trial. *Sci. Rep.* **2024**, *14*, 8833. [CrossRef] [PubMed]
 21. Roseman, L.; Nutt, D.J.; Carhart-Harris, R.L. Quality of acute psychedelic experience predicts therapeutic efficacy of psilocybin for treatment-resistant depression. *Front. Pharmacol.* **2018**, *8*, 974. [CrossRef] [PubMed]
 22. Murphy, R.; Kettner, H.; Zeifman, R.; Giribaldi, B.; Kärtner, L.; Martell, J.; Read, T.; Murphy-Beiner, A.; Baker-Jones, M.; Nutt, D.; et al. Therapeutic alliance and rapport modulate responses to psilocybin assisted therapy for depression. *Front. Pharmacol.* **2022**, *12*, 788155. [CrossRef] [PubMed]
 23. Goodwin, G.M.; Aaronson, S.T.; Alvarez, O.; Carhart-Harris, R.; Chai-Rees, J.; Croal, M.; DeBattista, C.; Dunlop, B.W.; Feifel, D.; Hellerstein, D.J.; et al. The role of the psychedelic experience in psilocybin treatment for treatment-resistant depression. *J. Affect. Disord.* **2025**, *372*, 523–532. Erratum in *J. Affect. Disord.* **2026**, *393*, 120352. <https://doi.org/10.1016/j.jad.2025.120352>. [CrossRef] [PubMed]
 24. Goodwin, G.M.; Aaronson, S.T.; Alvarez, O.; Carhart-Harris, R.; Croal, M.; Feifel, D.; Hellerstein, D.J.; Husain, M.I.; Kelly, J.R.; Kirlic, N.; et al. The role of therapeutic alliance in psilocybin treatment for treatment-resistant depression: A post hoc path analysis. *J. Affect. Disord.* **2026**, *406*, 121662. [CrossRef] [PubMed]
 25. Carhart-Harris, R.L.; Bolstridge, M.; Day, C.M.J.; Rucker, J.; Watts, R.; Erritzoe, D.E.; Kaelen, M.; Giribaldi, B.; Bloomfield, M.; Pilling, S.; et al. Psilocybin with psychological support for treatment-resistant depression: Six-month follow-up. *Psychopharmacology* **2018**, *235*, 399–408. [CrossRef] [PubMed]
 26. Rosenblat, J.D.; Meshkat, S.; Doyle, Z.; Kaczmarek, E.; Brudner, R.M.; Kratiuk, K.; Mansur, R.B.; Schulz-Quach, C.; Sethi, R.; Abate, A.; et al. Psilocybin-assisted psychotherapy for treatment resistant depression: A randomized clinical trial evaluating repeated doses of psilocybin. *Med* **2024**, *5*, 190–200.e5. [CrossRef] [PubMed]
 27. Aaronson, S.T.; van der Vaart, A.; Miller, T.; LaPratt, J.; Swartz, K.; Shoultz, A.; Lauterbach, M.; Suppes, T.; Sackeim, H.A. Single-dose psilocybin for depression with severe treatment resistance: An open-label trial. *Am. J. Psychiatry* **2025**, *182*, 104–113. [CrossRef] [PubMed]

28. Meikle, S.; Carter, O.; Liknaitzky, P.; Johansen, L.; Iyer, R.; Strauss, N.; Williams, M.; Castle, D.; Rossell, S.L. Psilocybin with psychotherapeutic support for treatment-resistant depression: A pilot clinical trial. *Ther. Adv. Psychopharmacol.* **2025**, *15*, 20451253251377187. [CrossRef] [PubMed]
29. Rucker, J.; Jafari, H.; Mantingh, T.; Bird, C.; Modlin, N.L.; Knight, G.; Reinholdt, F.; Day, C.; Carter, B.; Young, A. Psilocybin-assisted therapy for the treatment of resistant major depressive disorder (PsiDeR): Protocol for a randomised, placebo-controlled feasibility trial. *BMJ Open* **2021**, *11*, e056091. [CrossRef] [PubMed]
30. COMPASS Pathways, Ltd. *The Safety and Efficacy of Psilocybin in Participants with Treatment-Resistant Depression (P-TRD)*; COMPASS Pathways, Ltd.: London, UK, 2019. Available online: https://cdn.clinicaltrials.gov/large-docs/00/NCT03775200/Prot_000.pdf (accessed on 25 September 2026).
31. Aday, J.S.; Heifets, B.D.; Pratscher, S.D.; Bradley, E.; Rosen, R.; Woolley, J.D. Great expectations: Recommendations for improving the methodological rigor of psychedelic clinical trials. *Psychopharmacology* **2022**, *239*, 1989–2010. [CrossRef] [PubMed]
32. Orsini, D.K.; Wong, S.; Di Luch, S.; Chan, B.; Vasudeva, S.; Lovell, G.F.M.; Le, G.H.; Jones, B.D.M.; Chisamore, N.; Mollica, A.; et al. Blinding integrity in psychedelic randomized clinical trials: A systematic review. *JAMA Psychiatry* **2026**, *83*, 755–769. [CrossRef] [PubMed]
33. Williams, Z.J.; Barnett, H.; Szigeti, B. Psychedelic therapy vs antidepressants for the treatment of depression under equal unblinding conditions: A systematic review and meta-analysis. *JAMA Psychiatry* **2026**, *83*, 461–468. [CrossRef] [PubMed]
34. Singleton, S.P.; Sevchik, B.L.; Lahey, A.; Cuijpers, P.; Harrer, M.; Jones, M.T.; Nayak, S.M.; Strain, E.C.; Vandeckar, S.N.; Dworkin, R.H.; et al. A living systematic review, meta-analysis and open-data resource of randomized controlled trials of psilocybin treatment for symptoms of depression. *Nat. Ment. Health* **2026**, *4*, 870–880. [CrossRef]
35. Compass Pathways. Compass Pathways Successfully Achieves Primary Endpoint in First Phase 3 Trial Evaluating COMP360 Psilocybin for Treatment-Resistant Depression. 23 June 2025. Available online: <https://ir.compasspathways.com/News--Events-/news/news-details/2025/Compass-Pathways-Successfully-Achieves-Primary-Endpoint-in-First-Phase-3-Trial-Evaluating-COMP360-Psilocybin-for-Treatment-Resistant-Depression/default.aspx> (accessed on 9 August 2026).
36. Compass Pathways. Compass Pathways Successfully Achieves Primary Endpoint in Second Phase 3 Trial Evaluating COMP360 Psilocybin for Treatment-Resistant Depression. 17 February 2026. Available online: <https://ir.compasspathways.com/News--Events-/news/news-details/2026/Compass-Pathways-Successfully-Achieves-Primary-Endpoint-in-Second-Phase-3-Trial-Evaluating-COMP360-Psilocybin-for-Treatment-Resistant-Depression/default.aspx> (accessed on 25 September 2026).
37. Compass Pathways. Compass Pathways Announces Six-Month Data from Second Phase 3 Trial Confirming Rapid and Durable Profile. 7 July 2026. Available online: <https://ir.compasspathways.com/News--Events-/news/news-details/2026/Compass-Pathways-Announces-Six-Month-Data-from-Second-Phase-3-Trial-Confirming-Rapid-and-Durable-Profile/default.aspx> (accessed on 9 August 2026).
38. Duru, G.; Fantino, B. The clinical relevance of changes in the Montgomery-Asberg Depression Rating Scale using the minimum clinically important difference approach. *Curr. Med. Res. Opin.* **2008**, *24*, 1329–1335. [CrossRef] [PubMed]
39. Compass Pathways. Compass Pathways' New 52-Week Topline Open-Label Data from Phase 3 COMP005 Trial Demonstrates Additional Benefit from Another COMP360 Dose in Part C Extending Durability Out to 1 Year. 9 September 2026. Available online: <https://ir.compasspathways.com/News--Events-/news/news-details/2026/Compass-Pathways-New-52-Week-Topline-Open-Label-Data-from-Phase-3-COMP005-Trial-Demonstrates-Additional-Benefit-from-Another-COMP360-Dose-in-Part-C-Extending-Durability-Out-to-1-Year/default.aspx> (accessed on 25 September 2026).
40. Compass Pathways. Compass Pathways to Participate in TD Cowen 6th Annual Novel Mechanisms in Neuropsychiatry Summit on September 23, 2026. 16 September 2026. Available online: <https://ir.compasspathways.com/News--Events-/news/news-details/2026/Compass-Pathways-to-Participate-in-TD-Cowen-6th-Annual-Novel-Mechanisms-in-Neuropsychiatry-Summit-on-September-23-2026/default.aspx> (accessed on 25 September 2026).
41. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Harmonised Guideline E9(R1): Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials. 2019. Available online: https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf (accessed on 25 September 2026).
42. Goodwin, G.M.; Nowakowska, A.; Atli, M.; Dunlop, B.W.; Feifel, D.; Hellerstein, D.J.; Marwood, L.; Shabir, Z.; Mistry, S.; Stansfield, S.C.; et al. Results from a long-term observational follow-up study of a single dose of psilocybin for a treatment-resistant episode of major depressive disorder. *J. Clin. Psychiatry* **2025**, *86*, 24m15449. [CrossRef] [PubMed]
43. Mertens, L.J.; Betzler, F.; Brand, M.; Evens, R.; Jungaberle, A.; Jungaberle, H.; Kärtner, L.; Majić, T.; Schmitz, C.N.; Ströhle, A.; et al. Long-term efficacy of psilocybin with adjunct psychotherapy in treatment-resistant major depression (EPIsoDE): 6- and 12-month naturalistic follow-up of a phase 2b trial. *Psychother. Psychosom.* **2026**, 1–18. [CrossRef] [PubMed]
44. Yngwe, H.; Plavén-Sigra, P.; Ekman, C.J.; Henje, E.; Berglund, A.; Tiger, M.; Beckman, M.; Lundberg, J. Short-term and late-term effects of psilocybin on symptoms in major depression: A randomized clinical trial. *JAMA Netw. Open* **2026**, *9*, e2612589. [CrossRef] [PubMed]

45. Pearl, J. Conditioning on post-treatment variables. *J. Causal Inference* **2015**, *3*, 131–137. [CrossRef]
46. Matvey, M.; Kelley, D.P.; Bradley, E.R.; Chiong, W.; O'Donovan, A.; Woolley, J. Modifying informed consent to help address functional unmasking in psychedelic clinical trials. *JAMA Psychiatry* **2025**, *82*, 311–318. [CrossRef] [PubMed]
47. Pronovost-Morgan, C.; Greenway, K.T.; Roseman, L.; ReSPCT Experts. An international Delphi consensus for reporting of setting in psychedelic clinical trials. *Nat. Med.* **2025**, *31*, 2186–2195. [CrossRef] [PubMed]
48. Kittur, M.E.; Burgos, M.L.A.; Jones, B.D.M.; Blumberger, D.M.; Mulsant, B.H.; Rosenblat, J.D.; Husain, M.I. Mapping psilocybin therapy: A systematic review of therapeutic frameworks, adaptations, and standardization across contemporary clinical trials. *J. Affect. Disord.* **2025**, *391*, 119952. [CrossRef] [PubMed]
49. Kirlić, N.; Lennard-Jones, M.; Atli, M.; Malievskaja, E.; Modlin, N.L.; Knatz Peck, S.; Gaillard, A.; Goodwin, G.M.; Koelpin, D. Compass Psychological Support Model for COMP360 psilocybin treatment of serious mental health conditions. *Am. J. Psychiatry* **2025**, *182*, 126–132. [CrossRef] [PubMed]
50. Hoffmann, T.C.; Glasziou, P.P.; Boutron, I.; Milne, R.; Perera, R.; Moher, D.; Altman, D.G.; Barbour, V.; Macdonald, H.; Johnston, M.; et al. Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *BMJ* **2014**, *348*, g1687. [CrossRef] [PubMed]
51. Goodwin, G.M.; Croal, M.; Feifel, D.; Kelly, J.R.; Marwood, L.; Mistry, S.; O'Keane, V.; Knatz Peck, S.; Simmons, H.; Sisa, C.; et al. Psilocybin for treatment resistant depression in patients taking a concomitant SSRI medication. *Neuropsychopharmacology* **2023**, *48*, 1492–1499. [CrossRef] [PubMed]
52. Erritzoe, D.; Barba, T.; Spriggs, M.J.; Rosas, F.E.; Nutt, D.J.; Carhart-Harris, R. Effects of discontinuation of serotonergic antidepressants prior to psilocybin therapy versus escitalopram for major depression. *J. Psychopharmacol.* **2024**, *38*, 458–470. [CrossRef] [PubMed]
53. Marwood, L.; Croal, M.; Mistry, S.; Simmons, H.; Tsai, J.; Young, M.B.; Goodwin, G.M. The impact of antidepressant discontinuation prior to treatment with psilocybin for treatment-resistant depression. *J. Psychiatr. Res.* **2024**, *180*, 198–203. [CrossRef] [PubMed]
54. Tap, S.C.; Thomas, K.; Páleníček, T.; Stenbæk, D.S.; Oliveira-Maia, A.J.; van Dalen, J.H.; Schoevers, R.A. Concomitant use of antidepressants and classic psychedelics: A scoping review. *J. Psychopharmacol.* **2025**, *39*, 1072–1088. [CrossRef] [PubMed]
55. Hinkle, J.T.; Graziosi, M.; Nayak, S.M.; Yaden, D.B. Adverse events in studies of classic psychedelics: A systematic review and meta-analysis. *JAMA Psychiatry* **2024**, *81*, 1225–1235. [CrossRef] [PubMed]
56. Bukovsky, D.; Amaev, A.; Song, J.; Kyte, S.; Carmona-Torres, E.; Ueno, F.; Deluca, V.; Straffella, A.P.; Husain, M.I.; Graff-Guerrero, A.; et al. Adverse event reporting and management in psilocybin therapy clinical trials: A systematic review to guide clinical and research protocol development. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2025**, *143*, 111541. [CrossRef] [PubMed]
57. Junqueira, D.R.; Zorzela, L.; Golder, S.; Loke, Y.; Gagnier, J.J.; Julious, S.A.; Li, T.; Mayo-Wilson, E.; Pham, B.; Phillips, R.; et al. CONSORT Harms 2022 statement, explanation, and elaboration: Updated guideline for the reporting of harms in randomised trials. *BMJ* **2023**, *381*, e073725. [CrossRef] [PubMed]
58. van der Vaart, A.; LaPratt, J.; Swartz, K.; Shoultz, A.; Lauterbach, M.; Suppes, T.; Sackeim, H.A.; Aaronson, S.T. Efficacy and safety of a single dose of psilocybin for chronic suicidal ideation: An open-label trial. *J. Clin. Psychiatry* **2026**, *87*, 26m16338. [CrossRef] [PubMed]
59. Reiff, C.M.; Richman, E.E.; Nemeroff, C.B.; Carpenter, L.L.; Widge, A.S.; Rodriguez, C.I.; Kalin, N.H.; McDonald, W.M.; Work Group on Biomarkers and Novel Treatments, a Division of the American Psychiatric Association Council of Research. Psychedelics and psychedelic-assisted psychotherapy. *Am. J. Psychiatry* **2020**, *177*, 391–410. [CrossRef] [PubMed]
60. Johnson, M.W.; Richards, W.A.; Griffiths, R.R. Human hallucinogen research: Guidelines for safety. *J. Psychopharmacol.* **2008**, *22*, 603–620. [CrossRef] [PubMed]
61. Eskinazi, M.; Nasserdine, R.; Cusin, R.M.; Baniotoupoulos, P.; Saccaro, L.F.; De Pieri, M.; Corino, T.; Seragnoli, F.; Briefer, J.-F.; Aboulafia Brakha, T.; et al. Psychedelic-induced hypomania and mania: A systematic review and meta-analysis. *Mol. Psychiatry* **2026**, *31*, 5749–5763. [CrossRef] [PubMed]
62. Albert, O.M.; Arthur, A. Psychedelic exposure in pregnancy: A scoping review to inform perinatal drug safety and clinical counseling. *Ther. Adv. Drug Saf.* **2026**, *17*, 20420986261436104. [CrossRef] [PubMed]
63. Barber, G.S.; Dike, C.C. Ethical and practical considerations for the use of psychedelics in psychiatry. *Psychiatr. Serv.* **2023**, *74*, 838–846. [CrossRef] [PubMed]
64. Jacobs, E.; Zahid, Z.; Hinkle, J.; Nayak, S.; Yaden, D.B. Psychedelic medicine: Mechanisms, evidence, and translation to practice. *BMJ* **2026**, *392*, e081723. [CrossRef] [PubMed]
65. Therapeutic Goods Administration. Accessing MDMA (3,4-Methylenedioxy-Methamphetamine) and Psilocybine as a Psychiatrist. Updated 22 June 2026. Available online: <https://www.tga.gov.au/resources/guidance/accessing-mdma-34-methylenedioxy-methamphetamine-and-psilocybine-psychiatrist> (accessed on 30 September 2026).
66. Health Canada. Psilocybin and Psilocin (Magic Mushrooms). Available online: <https://www.canada.ca/en/health-canada/services/substance-use/controlled-illegal-drugs/magic-mushrooms.html> (accessed on 30 September 2026).

67. Federal Office of Public Health. Limited Medical Use of Banned Narcotics. Available online: <https://www.bag.admin.ch/en/limited-medical-use-of-banned-narcotics> (accessed on 30 September 2026).
68. Ministry of Health of the Czech Republic. *Věstník MZD ČR 21/2025: Doporučený Postup PS-ČNPS Pro Asistovanou Psychoterapii Psilocybinem Pro Léčebné Použití*; Ministry of Health of the Czech Republic: Prague, Czech Republic, 2025; pp. 69–98. Available online: <https://mzd.gov.cz/vestnik/vestnik-21-2025/> (accessed on 30 September 2026).
69. United Nations. Convention on Psychotropic Substances, 1971. Article 7 and Schedule I. Available online: https://www.unodc.org/pdf/convention_1971_en.pdf (accessed on 30 September 2026).
70. National Institute of Mental Health, Czechia. Official Statement on the Availability of Psilocybin Treatment. 23 March 2026. Available online: <https://www.nudz.cz/en/news-and-events/official-statement-on-the-availability-of-psilocybin-treatment> (accessed on 30 September 2026).
71. Liechti, M.E.; Gasser, P.; Aicher, H.D.; Mueller, F.; Hawrot, T.; Schmid, Y. Implementing psychedelic-assisted therapy: History and characteristics of the Swiss limited medical use program. *Neurosci. Appl.* **2025**, *4*, 105525. [CrossRef] [PubMed]
72. Lam, R.W.; Kennedy, S.H.; Adams, C.; Bahji, A.; Beaulieu, S.; Bhat, V.; Blier, P.; Blumberger, D.M.; Brietzke, E.; Chakrabarty, T.; et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2023 Update on Clinical Guidelines for Management of Major Depressive Disorder in Adults. *Can. J. Psychiatry* **2024**, *69*, 641–687. [CrossRef] [PubMed]
73. National Institute for Health and Care Excellence. *Depression in Adults: Treatment and Management*; National Institute for Health and Care Excellence (NICE): London, UK, 2022. Available online: <https://www.nice.org.uk/guidance/ng222> (accessed on 25 September 2026).
74. Noda, Y.; Wada, M.; Mimura, Y.; Taniguchi, K.; Tarumi, R.; Moriyama, S.; Arai, N.; Tsugawa, S.; Thorpe, K.E.; Daskalakis, Z.J.; et al. Repetitive transcranial magnetic stimulation as maintenance treatment of depression: The MAINT-R randomized clinical trial. *JAMA Netw. Open* **2025**, *8*, e2515881. [CrossRef] [PubMed]
75. Perera, T.; George, M.S.; Grammer, G.; Janicak, P.G.; Pascual-Leone, A.; Wirecki, T.S. The Clinical TMS Society consensus review and treatment recommendations for TMS therapy for major depressive disorder. *Brain Stimul.* **2016**, *9*, 336–346. [CrossRef] [PubMed]
76. Sanacora, G.; Frye, M.A.; McDonald, W.; Mathew, S.J.; Turner, M.S.; Schatzberg, A.F.; Summergrad, P.; Nemeroff, C.B.; American Psychiatric Association (APA) Council of Research Task Force on Novel Biomarkers and Treatments. A consensus statement on the use of ketamine in the treatment of mood disorders. *JAMA Psychiatry* **2017**, *74*, 399–405. [CrossRef] [PubMed]
77. Janssen Pharmaceuticals, Inc. *SPRAVATO (Esketamine) Nasal Spray, CIII: Full Prescribing Information*; Janssen Pharmaceuticals, Inc.: Titusville, NJ, USA, 2026. Available online: <https://www.jnjlabels.com/package-insert/product-monograph/prescribing-information/SPRAVATO-pi.pdf> (accessed on 25 September 2026).
78. Tai, S.J.; Nielson, E.M.; Lennard-Jones, M.; Johanna Ajantaival, R.-L.; Winzer, R.; Richards, W.A.; Reinholdt, F.; Richards, B.D.; Gasser, P.; Malievskaia, E. Development and evaluation of a therapist training program for psilocybin therapy for treatment-resistant depression in clinical research. *Front. Psychiatry* **2021**, *12*, 586682. [CrossRef] [PubMed]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.