The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 NOVEMBER 3, 2022 VOL. 387 NO. 18

Single-Dose Psilocybin for a Treatment-Resistant Episode of Major Depression

G.M. Goodwin, S.T. Aaronson, O. Alvarez, P.C. Arden, A. Baker, J.C. Bennett, C. Bird, R.E. Blom, C. Brennan, D. Brusch, L. Burke, K. Campbell-Coker, R. Carhart-Harris, J. Cattell, A. Daniel, C. DeBattista, B.W. Dunlop, K. Eisen, D. Feifel, M.K. Forbes, H.M. Haumann, D.J. Hellerstein, A.I. Hoppe, M.I. Husain, L.A. Jelen, J. Kamphuis, J. Kawasaki, J.R. Kelly, R.E. Key, R. Kishon, S. Knatz Peck, G. Knight, M.H.B. Koolen, M. Lean, R.W. Licht, J.L. Maples-Keller, J. Mars, L. Marwood, M.C. McElhiney, T.L. Miller, A. Mirow, S. Mistry, T. Mletzko-Crowe, L.N. Modlin, R.E. Nielsen, E.M. Nielson, S.R. Offerhaus, V. O'Keane, T. Páleníček, D. Printz, M.C. Rademaker, A. van Reemst, F. Reinholdt, D. Repantis, J. Rucker, S. Rudow, S. Ruffell, A.J. Rush, R.A. Schoevers, M. Seynaeve, S. Shao, J.C. Soares, M. Somers, S.C. Stansfield, D. Sterling, A. Strockis, J. Tsai, L. Visser, M. Wahba, S. Williams, A.H. Young, P. Ywema, S. Zisook, and E. Malievskaia

ABSTRACT

BACKGROUND

Psilocybin is being studied for use in treatment-resistant depression.

METHODS

In this phase 2 double-blind trial, we randomly assigned adults with treatment-resistant depression to receive a single dose of a proprietary, synthetic formulation of psilocybin at a dose of 25 mg, 10 mg, or 1 mg (control), along with psychological support. The primary end point was the change from baseline to week 3 in the total score on the Montgomery–Åsberg Depression Rating Scale (MADRS; range, 0 to 60, with higher scores indicating more severe depression). Secondary end points included response at week 3 (≥50% decrease from baseline in the MADRS total score), remission at week 3 (MADRS total score ≤10), and sustained response at 12 weeks (meeting response criteria at week 3 and all subsequent visits).

RESULTS

A total of 79 participants were in the 25-mg group, 75 in the 10-mg group, and 79 in the 1-mg group. The mean MADRS total score at baseline was 32 or 33 in each group. Least-squares mean changes from baseline to week 3 in the score were –12.0 for 25 mg, –7.9 for 10 mg, and –5.4 for 1 mg; the difference between the 25-mg group and 1-mg group was –6.6 (95% confidence interval [CI], –10.2 to –2.9; P<0.001) and between the 10-mg group and 1-mg group was –2.5 (95% CI, –6.2 to 1.2; P=0.18). In the 25-mg group, the incidences of response and remission at 3 weeks, but not sustained response at 12 weeks, were generally supportive of the primary results. Adverse events occurred in 179 of 233 participants (77%) and included headache, nausea, and dizziness. Suicidal ideation or behavior or self-injury occurred in all dose groups.

CONCLUSIONS

In this phase 2 trial involving participants with treatment-resistant depression, psilocybin at a single dose of 25 mg, but not 10 mg, reduced depression scores significantly more than a 1-mg dose over a period of 3 weeks but was associated with adverse effects. Larger and longer trials, including comparison with existing treatments, are required to determine the efficacy and safety of psilocybin for this disorder. (Funded by COMPASS Pathfinder; EudraCT number, 2017-003288-36; ClinicalTrials.gov number, NCT03775200.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Dr. Goodwin can be contacted at guy.goodwin@compasspathways.com.

This article was updated on November 3, 2022, at NEJM.org.

N Engl J Med 2022;387:1637-48.
DOI: 10.1056/NEJMoa2206443

A Quick Take is available at NEJM.org

Treatment-resistant depression is a challenging disorder to treat, as shown in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial.[1] Incidences of remission became progressively lower from the first course of antidepressant treatment (36.8%) to the second course (30.6%), third course (13.7%), and fourth course (13.0%).[1,2] Failure of two courses of treatment has generally been considered to define a group of patients who have treatment-resistant depression. Patients with treatment-resistant depression have greater severity and duration of illness, disability, physical illness, incidences of hospitalization, risk of suicide, and economic costs than patients with treatment-responsive depression.[1-3]

Psilocybin is a tryptamine alkaloid found in several species of psilocybe mushrooms.[4] Its potential antidepressant efficacy was suggested by preliminary studies involving patients with life-threatening cancer.[5-7] Amelioration of symptomatic depression in pilot studies of major depressive disorder, including those that compared psilocybin with escitalopram[8,9] and that investigated its use in treatment-resistant depression,[10] has suggested therapeutic potential for this agent. The objective of the current trial was to identify an acceptable efficacious dose and assess the safety of a synthetic, proprietary formulation of psilocybin, administered together with psychological support,[11] in patients with a treatment-resistant major depressive episode.

METHODS

TRIAL OVERSIGHT

This was a phase 2 double-blind, dose-finding, parallel-group, randomized clinical trial. The sponsor, COMPASS Pathfinder, designed and funded the trial and provided a proprietary pharmaceutical-grade synthetic psilocybin formulation, COMP360, which was analyzed for stability and purity. A contract research organization (Worldwide Clinical Trials), paid by the sponsor, supervised the conduct of the trial. An independent contract research organization (MedAvante-ProPhase) was responsible for assessment of participants using the Montgomery–Åsberg Depression Rating Scale (MADRS),[12] performed by trained remote raters who were unaware of the details of the trial and the trial-group assignments. The statistical analysis of the data was performed by the contract research organization and reviewed by the sponsor, and the interpretation and post hoc statistical analyses of the data were performed by the sponsor. The sponsor paid for professional writing assistance for the first draft of the manuscript. All the authors reviewed and approved the manuscript before submission and vouch for the adherence of the trial to the protocol (available with the full text of this article at NEJM.org), the completeness and accuracy of the data, and the reporting of adverse events. Confidentiality agreements were in place between the investigators and COMPASS Pathfinder. The roles of the authors are listed in the Supplementary Appendix, available at NEJM.org.

The trial was conducted in accordance with the International Council for Harmonisation (ICH) Good Clinical Practice guidelines and the ethical principles of the Declaration of Helsinki. The trial protocol was approved by independent ethics committees or institutional review boards at each participating site. All the participants provided written informed consent.

PARTICIPANTS

Men and women 18 years of age or older were eligible if they met *Diagnostic and Statistical Manual of Mental Disorders* (fifth edition) criteria for a single or recurrent episode of major depressive disorder, without psychotic features, on the basis of clinical assessment and medical records and as documented by the Mini-International Neuropsychiatric Interview (version 7.0.2).[13] Recruitment was conducted through referrals from primary care and specialized psychiatry services, online advertisements, and word of mouth. Participants were outpatients who met criteria for the diagnosis of treatment-resistant depression and had a current episode of depression that had not responded to two to four adequate trials in terms of both dose and duration (≥8 weeks) of treatment according to the Massachusetts General Hospital Antidepressant Treatment Response Questionnaire (MGH ATRQ).[14] Augmentation agents, or other antidepressants not included in the MGH ATRQ, qualified as a treatment failure if they failed to ameliorate depression, provided they had local regulatory approval as a treatment for major depressive disorder. Additional selec-

tion criteria and screening procedures are summarized in the trial protocol.

TRIAL DESIGN AND PROCEDURES

The trial was conducted at 22 sites in 10 countries in Europe (the Czech Republic, Denmark, Germany, Ireland, the Netherlands, Portugal, Spain, and the United Kingdom) and North America (Canada and the United States) from March 1, 2019, through September 27, 2021. All but one of the principal investigators was a psychiatrist. Both assisting and lead therapists, whose roles are described below, were recruited as psychologists with at least master's-level qualifications, psychiatrists, master's-level practitioners, nurses, diploma-level cognitive behavioral therapists, or doctorate-level mental health specialists. These therapists had experience in adult mental health, addiction, dementia, physical health, child or developmental health, family therapy, or eating disorders and experience with patients having severe psychological distress. The therapist-training program that was expressly prepared for the trial had four components: an online learning platform, in-person training, clinical training, and ongoing individual mentoring and webinars. Therapists were required to complete the first three components of the training program before they could lead sessions independently and to engage in the fourth component to continue their professional development.[11] Therapists in training could act as assisting therapists so that there were always two therapists present on the day of drug administration. All the therapists were unaware of the trial-group assignments, did not collect efficacy assessments, and were discouraged from speculating about doses.

Eligible participants completed a run-in period of 3 to 6 weeks, during which antidepressants and other prohibited medications affecting the central nervous system were tapered and discontinued at least 2 weeks before the baseline visit (the day before psilocybin administration). During this period, the participant met with a therapist at least three times to build trust, receive psychoeducation, and prepare for the psychedelic experience. Participants who continued to meet eligibility criteria were randomly assigned in a 1:1:1 ratio to receive a single dose of psilocybin of 25 mg, 10 mg, or 1 mg (control). Randomization was performed at a central location and stratified according to country and the participant's previous experience with psilocybin. The administration session (day 1) lasted 6 to 8 hours, with the lead therapist who had prepared the participant for the intervention and an assisting therapist in attendance. A trial psychiatrist was available on site for consultation. Administration rooms were designed to provide a nonclinical, calming atmosphere. During the administration session, participants listened to a specially designed music playlist while wearing eyeshades to help direct attention internally. After at least 6 hours and when the psychedelic effects of the drug had fully dissipated, participants returned home.

The trial followed participants for 12 weeks after treatment. Participants received two integration sessions, with the same lead and assisting therapists at the day 2 visit and with the lead therapist at the week 1 visit. The goal of the integration sessions was to support participants in deriving their own insights and solutions from the experience with psilocybin. Therapists were advised to remain open and supportive, without active guiding.[11] Participants were requested to remain off antidepressant treatment during the first 3 weeks after the trial-drug administration; however, these medications could be started at any time during the trial if deemed clinically necessary by a physician investigator. (A schedule of the assessments is provided in Table S1 in the Supplementary Appendix.)

EFFICACY END POINTS

The primary end point was the change from baseline (day –1, the day before trial-drug administration) to 3 weeks in the MADRS total score (range, 0 to 60, with higher scores indicating greater severity of depression).[12] The primary analysis was of the 25-mg dose and 10-mg dose each compared with the 1-mg dose. The MADRS was administered by experienced mental health clinician raters by telephone at baseline, on day 2, and at weeks 1, 3 (primary end-point assessment), 6, 9, and 12. The Structured Interview Guide for the MADRS provided structured probes to ensure standardization of administration and comprehensive coverage of the 10 questions.[15] Three key secondary efficacy end points were response (≥50% decrease from baseline to week 3

in the MADRS total score), remission (MADRS total score ≤10 at week 3), and sustained response (week 3 response maintained through week 12).

SAFETY END POINTS

Adverse events were evaluated at every visit and were recorded and coded with the use of the *Medical Dictionary for Regulatory Activities* (MedDRA), version 23.0. All visits were in conducted in person except for the week 6 and 9 visits, which were conducted remotely. Adverse events that emerged or worsened after trial-drug administration were categorized as serious adverse events on the basis of the ICH Good Clinical Practice criteria and with the use of additional information from the Columbia Suicide Severity Rating Scale.[16] Suicidal ideation with intent or endorsement of any items in the suicidal-behavior section, including nonsuicidal self-injurious behavior, was reported as a serious adverse event. Safety assessments also included evaluation of vital signs (at screening, baseline, day 1, and day 2), clinical laboratory tests (including urine drug screening) (at screening, day 2, and week 3), and 12-lead electrocardiography (ECG) at screening and day 2.

STATISTICAL ANALYSIS

Using a two-sample t-test, we calculated that a sample of 216 participants (72 per group) would provide 90% power at a two-sided alpha level of 0.05 to detect a 6-point difference in the mean change from baseline to week 3 in the MADRS total score between the 25-mg group or the 10-mg group and the 1-mg group, assuming a common standard deviation of 11.0 (see the Supplementary Appendix). Efficacy analyses were performed in the modified intention-to-treat analysis set, which included all randomly assigned participants who received treatment and had at least one postbaseline efficacy assessment.

A "hypothetical strategy" estimand was applied in which MADRS total scores for participants who initiated a new antidepressant treatment were imputed at visits after initiation with the use of a missing-not-at-random mechanism that progressively worsened the MADRS total score. The aim was to hypothesize what would have happened to the MADRS total score had a new treatment for depression not been available to use. This same method was also applied to missing MADRS total scores after trial withdrawal for reasons of lack of efficacy or adverse events. All other missing data on MADRS total scores, both intermittent and after trial withdrawal for other reasons, were imputed with the use of a missing-at-random mechanism.

The primary efficacy end point (change from baseline to week 3 in the MADRS total score) was evaluated with the use of a mixed model for repeated measures (MMRM) analysis comparing the 25-mg dose with the 1-mg dose and comparing the 10-mg dose with the 1-mg dose. The MMRM analysis included treatment, visit, pooled trial site, treatment-by-visit interaction, baseline MADRS total score, and an unstructured correlation matrix. The estimates of the least-squares means and mean differences and 95% confidence intervals were then pooled with the use of Rubin's combination rules. This analysis method combined the between-imputation variability with the within-imputation variability to obtain one single point and confidence interval estimate to address imputation uncertainty.

Response and remission were analyzed with the use of a generalized linear mixed model, and sustained response was analyzed with the use of a logistic-regression model. A "composite strategy" estimand was applied, whereby participants who initiated a new antidepressant treatment or withdrew from the trial for reasons of lack of efficacy or adverse events were classified as not having a response, remission, or a sustained response at all visits after these events.

To control the overall type I error rate, a hierarchical test procedure was applied across the primary and three key secondary efficacy end points. The 25-mg group and then the 10-mg group were sequentially examined for each end point before proceeding to the next end point. All testing was done at the two-sided 0.05 alpha level. Descriptive statistics were used to analyze safety data from all randomly assigned participants who received single-dose treatment (safety analysis set), including adverse events, concomitant medications, evaluation of vital signs, clinical laboratory tests, findings from 12-lead ECG, and suicidality assessments.

RESULTS

PARTICIPANTS

A total of 428 participants were screened, and 233 were enrolled, underwent randomization, and received psilocybin treatment (safety analy-

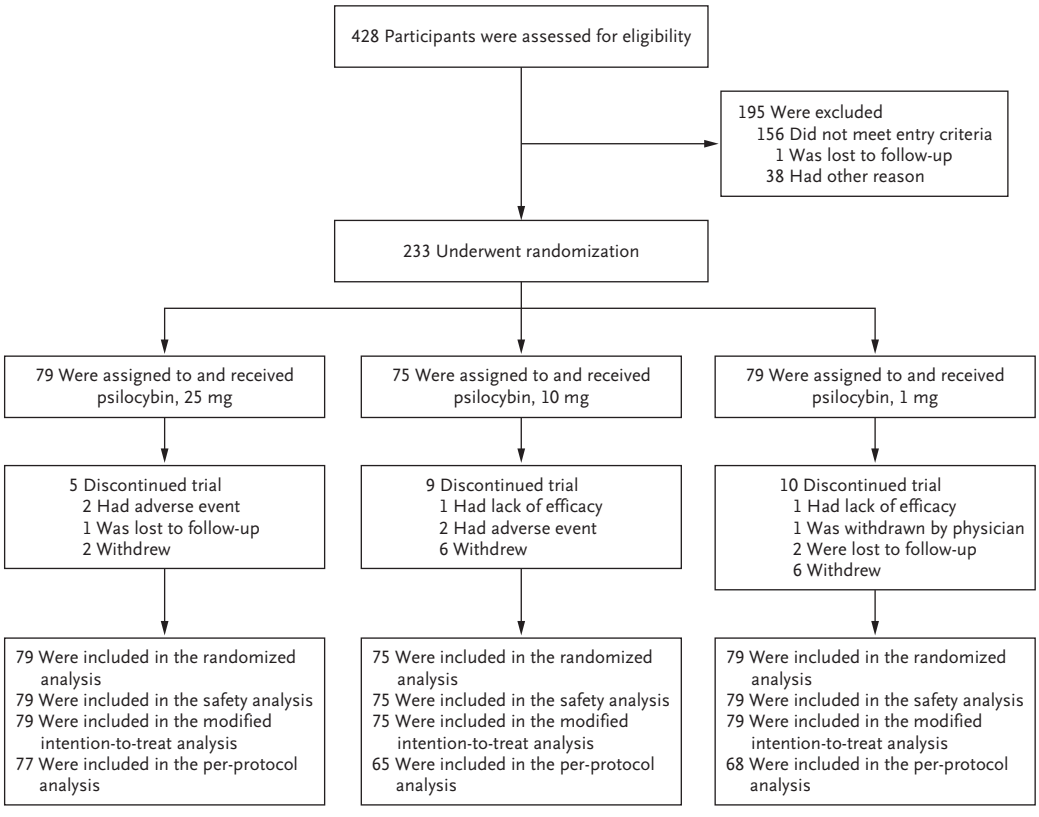

Figure 1. Enrollment, Randomization, and Follow-up of the Participants.

Randomly assigned participants received a single dose of a proprietary, synthetic formulation of psilocybin, which was administered together with psychological support.

sis set) and had at least one postbaseline efficacy evaluation (modified intention-to-treat analysis set). A total of 79 participants were assigned to the 25-mg group, 75 to the 10-mg group, and 79 to the 1-mg group (Fig. 1). By week 12, a total of 5 participants (6%) in the 25-mg group, 9 (12%) in the 10-mg group, and 10 (13%) in the 1-mg group had withdrawn from the trial.

The demographic and clinical characteristics of the participants at baseline were similar across the three groups (Table 1); the mean age was 39.8 years, 52% were female, and 92% were White. A total of 95% of the participants reported previous depressive episodes, with a mean of 6.9 lifetime depressive episodes, and 86% of the participants reported a duration of the current depressive episode of longer than 1 year. These characteristics were similar to what has been observed in population studies involving persons with treatment-resistant depression, and the representativeness of the trial population is shown in Table S11. Two thirds of the participants were receiving antidepressant treatment at screening. At baseline, depression was moderate (MADRS total score, 20 to 30) in 30% of the participants and severe (MADRS total score, ≥31) in 68% of the participants. Mean MADRS total scores at baseline were 31.9 in the 25-mg group, 33.0 in the 10-mg group, and 32.7 in the 1-mg group. A total of 6% of the participants had previous exposure to psilocybin.

Before the week 3 primary end-point assessment, initiation of treatment for depression was reported by 4 participants (5%) in the 25-mg group, 9 (12%) in the 10-mg group, and 14 (18%) in the 1-mg group. After week 3 and up to week 12, the number of participants initiating a treatment for depression was 26 (33%) in the 25-mg group, 18 (24%) in the 10-mg group, and 16 (20%) in the 1-mg group.

Table 1. Demographic and Clinical Characteristics of the Participants at Baseline (Safety Population).*

Characteristic	Psilocybin, 25 mg (N=79)	Psilocybin, 10 mg (N=75)	Psilocybin, 1 mg (N=79)	Overall (N=233)
Demographic characteristics				
Female sex — no. (%)	44 (56)	41 (55)	36 (46)	121 (52)
Age — yr	40.2±12.2	40.6±12.8	38.7±11.7	39.8±12.2
White race — no. (%)†	70 (89)	72 (96)	73 (92)	215 (92)
Body-mass index‡	26.52±6.13	28.26±8.20	27.26±6.02	27.34±6.86
Previous psilocybin use — no. (%)	5 (6)	5 (7)	4 (5)	14 (6)
Psychiatric history				
Recurrent MDD episode — no. (%)	75 (95)	74 (99)	73 (92)	222 (95)
Lifetime depressive episodes — no.				
Mean	7.3±8.6	7.8±9.1	5.7±4.4	6.9±7.6
Median	5.0	4.0	4.0	5.0
Duration of current depressive episode — no. (%)				
<1 yr	12 (15)	10 (13)	10 (13)	32 (14)
1 yr to <2 yr	33 (42)	28 (37)	33 (42)	94 (40)
≥2 yr	34 (43)	37 (49)	36 (46)	107 (46)
Failed treatments for current depressive episode — no. (%)				
2	66 (84)	62 (83)	63 (80)	191 (82)
3 or 4	12 (15)	11 (15)	14 (18)	37 (16)
Withdrawn from antidepressant at trial entry — no. (%)	53 (67)	51 (68)	52 (66)	156 (67)
Failure of treatment trial of augmentation agent during current depressive episode — no. (%)	5 (6)	3 (4)	6 (8)	14 (6)
Depression scores				
MADRS total score§				
Mean	31.9±5.4	33.0±6.3	32.7±6.2	32.5±6.0
Moderate: 20–30 — no. (%)	33 (42)	19 (25)	18 (23)	70 (30)
Severe: ≥31 — no. (%)	46 (58)	54 (72)	59 (75)	159 (68)
HAM-D-17 total score¶				
Mean	21.8±3.0	22.4±2.8	22.2±2.9	22.2±2.9
Moderate: 18–23 — no. (%)	57 (72)	49 (65)	59 (75)	165 (71)
Severe: ≥24 — no. (%)	22 (28)	26 (35)	20 (25)	68 (29)

* Plus–minus values are means ±SD. Randomly assigned participants received a single dose of a proprietary, synthetic formulation of psilocybin, which was administered together with psychological support. Percentages may not total 100 because of rounding. MDD denotes major depressive disorder.

† Race was reported by the participants.

‡ The body-mass index is the weight in kilograms divided by the square of the height in meters.

§ Total scores on the Montgomery–Åsberg Depression Rating Scale (MADRS) range from 0 to 60, with higher scores indicating greater severity of depression. Two participants in the 10-mg group and two participants in the 1-mg group had an MADRS total score of less than 20 at baseline.

¶ Total scores on the 17-item Hamilton Depression Rating Scale (HAM-D-17) range from 0 to 52, with higher scores indicating greater severity of depression.

EFFICACY

The least-squares mean change from baseline to week 3 in the MADRS total score was –12.0 points in the 25-mg group, –7.9 in the 10-mg group, and –5.4 in the 1-mg group. The difference in the least-squares mean change between the 25-mg group and the 1-mg group was –6.6 (95% confidence interval [CI], –10.2 to –2.9;

Table 2. Primary and Secondary Efficacy End Points (Modified Intention-to-Treat Population).*

End Point	Psilocybin, 25 mg (N=79)	Psilocybin, 10 mg (N=75)	Psilocybin, 1 mg (N=79)
Primary efficacy end point			
Change from baseline to wk 3 in MADRS total score			
Least-squares mean	−12.0±1.3	−7.9±1.4	−5.4±1.4
95% CI of the least-squares mean	−14.6 to −9.3	−10.6 to −5.2	−8.1 to −2.7
Least-squares mean difference vs. 1 mg	−6.6±1.9	−2.5±1.9	—
95% CI of the least-squares mean difference	−10.2 to −2.9	−6.2 to 1.2	
P value vs. 1 mg	<0.001	0.18†	—
Secondary efficacy end points			
Response at wk 3‡			
No. of participants (%)	29 (37)	14 (19)	14 (18)
Odds ratio vs. 1 mg (95% CI)	2.9 (1.2 to 6.6)	1.2 (0.5 to 3.0)	—
Remission at wk 3§			
No. of participants (%)	23 (29)	7 (9)	6 (8)
Odds ratio vs. 1 mg (95% CI)	4.8 (1.8 to 12.8)	1.2 (0.4 to 3.9)	—
Sustained response at wk 12¶			
No. of participants (%)	16 (20)	4 (5)	8 (10)
Odds ratio vs. 1 mg (95% CI)	2.2 (0.9 to 5.4)	0.7 (0.2 to 2.0)	—

* Plus–minus values are standard errors.
† This nonsignificant P value terminated significance testing on the basis of the prespecified hierarchical test procedure, and all the subsequent secondary efficacy end points are considered to be not significantly different between the 25-mg group or 10-mg group and the 1-mg group.
‡ A response was defined as a decrease of at least 50% from baseline in the MADRS total score.
§ Remission was defined as an MADRS total score of 10 or less.
¶ A sustained response was defined as a week 3 response sustained through week 12.

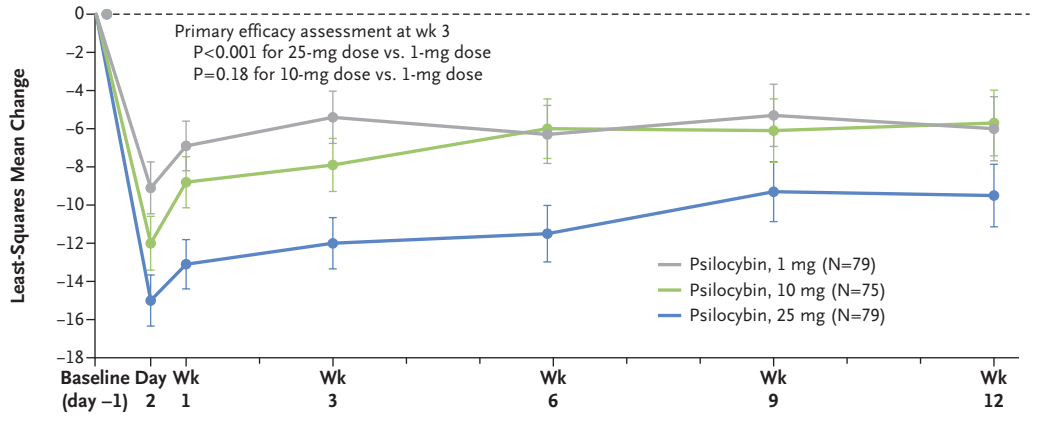

Figure 2. Change from Baseline in MADRS Total Score (Modified Intention-to-Treat Population).
Total scores on the Montgomery–Åsberg Depression Rating Scale (MADRS) range from 0 to 60, with higher scores indicating greater severity of depression. I bars represent standard errors.

P<0.001), and the difference between the 10-mg group and the 1-mg group was −2.5 (95% CI, −6.2 to 1.2; P=0.18) (Table 2 and Fig. 2). The nonsignificant finding for the comparison between the 10-mg group and the 1-mg group terminated significance testing on the basis of the prespecified hierarchical test procedure, and all the subsequent key secondary efficacy end points are considered to be not significantly different between the 25-mg group or the 10-mg group and the 1-mg group. Additional analyses for the primary efficacy end point are shown in Figure S2. These alternative data-handling strategies and analysis models provided results that were consistent with the findings for the primary efficacy end point.

The incidence of response at week 3 was 37% in the 25-mg group, 19% in the 10-mg group, and 18% in the 1-mg group (odds ratio in the 25-mg group vs. the 1-mg group, 2.9 [95% CI, 1.2 to 6.6]; odds ratio in the 10-mg group vs. the 1-mg group, 1.2 [95% CI, 0.5 to 3.0]) (Table 2). The incidence of remission at week 3 was 29% in the 25-mg group, 9% in the 10-mg group, and 8% in the 1-mg group (odds ratio in the 25-mg group vs. the 1-mg group, 4.8 [95% CI, 1.8 to 12.8]; odds ratio in the 10-mg group vs. the 1-mg group, 1.2 [95% CI, 0.4 to 3.9]). The incidence of sustained response at week 12 was 20% in the 25-mg group, 5% in the 10-mg group, and 10% in the 1-mg group (odds ratio in the 25-mg group vs. the 1-mg group, 2.2 [95% CI, 0.9 to 5.4]; odds ratio in the 10-mg group vs. the 1-mg group, 0.7 [95% CI, 0.2 to 2.0]). Because of the failure of hierarchical testing, no definite conclusions can be drawn from secondary end-point results. The confidence interval for the odds ratio for sustained response at week 12 for both the 25-mg dose and the 10-mg dose as compared with the 1-mg dose included 1. A post hoc analysis of the primary end point that included sex or the number of lifetime episodes of depression showed results similar to those for the primary analysis. The results from per-protocol analysis of the primary end point were also consistent with the modified intention-to-treat population (Fig. S2). Additional efficacy results are included in Tables S3 through S6 and Figures S1 and S2.

SAFETY

Adverse events occurred in 66 participants (84%) in the 25-mg group, 56 (75%) in the 10-mg group, and 57 (72%) in the 1-mg group. The most frequent adverse events reported in the 25-mg group with onset on the day of psilocybin administration (day 1) were headache (in 24% of the participants), nausea (in 22%), and dizziness and fatigue (in 6% each) (Table 3). Adverse events that were rated as severe on day 1 were reported by 4% of the participants in the 25-mg group, 8% of those in the 10-mg group, and 1% of those in the 1-mg group. Just one participant (in the 25-mg group) was treated with adjunctive medication (lorazepam for acute anxiety) on day 1. There were no serious adverse events reported on day 1.

From day 2 up to week 3 (primary end-point assessment), severe adverse events were reported by 9% of the participants in the 25-mg group, 7% of those in the 10-mg group, and 1% of those in the 1-mg group. The serious adverse events in the 25-mg group were suicidal ideation (in two participants) and intentional self-injury (nonsuicidal self-injurious behavior) (in two participants) and in the 10-mg group were suicidal ideation (in two participants), intentional self-injury (in one participant), and hospitalization (for severe depression, in one participant). No serious adverse events were reported from day 2 up to week 3 in the 1-mg group.

After week 3 and up to week 12 (end of trial), severe adverse events were reported by 3% of the participants in the 25-mg group, 4% of those in the 10-mg group, and no participants in the 1-mg group. Serious adverse events in the 25-mg group were suicidal behavior (in three participants), codeine withdrawal syndrome (in one participant), and adjustment disorder with anxiety and depressed mood (in one participant); in the 10-mg group were intentional self-injury (in one participant), depression (in one participant), and suicidal ideation (in one participant); and in the 1-mg group were intentional self-injury (in one participant). Severe adverse events during the trial period according to MedDRA system organ class and preferred term are shown in Table S7.

At the baseline visit, suicidal ideation (passive or active but with no intent or plan) was reported by 21 participants (27%) in the 25-mg group, 27 (36%) in the 10-mg group, and 19 (24%) in the 1-mg group. The number of participants who showed worsening of suicidal state from baseline to week 3 were 11 (14%) in the 25-mg group, 13 (17%) in the 10-mg group, and 7 (9%) in the 1-mg group (Table S8). Three

participants in the 25-mg group reported suicidal behavior after week 3. All three had a history of suicidal behavior or nonsuicidal self-injury before the trial and did not have a treatment response at week 3. No clinically significant changes in vital signs, clinical laboratory tests, or 12-lead ECGs were observed during the trial (see the Supplementary Appendix).

DISCUSSION

This phase 2 clinical trial showed the feasibility of psilocybin monotherapy for up to 12 weeks in patients with a treatment-resistant episode of major depression. The change from baseline to week 3 in the MADRS total score (primary end point) was significantly better with a 25-mg dose than with a 1-mg dose; there was not significant difference between the 10-mg dose and the 1-mg dose. In addition to headache, nausea, dizziness, and fatigue, some participants had suicidal ideation or self-injurious behavior, and the proportions of these participants were numerically higher in the 25-mg and 10-mg groups than in the 1-mg group. In view of the participants who showed worsening of suicidal state, suicidality demands clinical vigilance in future trials of psilocybin for depression. The incidences of response and remission at 3 weeks were generally in the same direction as the primary end-point results; however, the analyses of these end points were ordered in the prespecified hierarchical test procedure after the significance testing had terminated, and no definite conclusions can be drawn from these results. The confidence interval for the odds ratio for sustained response at week 12 for the 25-mg-group as compared with the 1-mg group included 1.

The current trial was designed to address some limitations of previous pilot studies and trials, including limited power, short-duration crossover design, reliance on single-site recruitment of participants, and interpretation of treatment effects that may be confounded by intensive concurrent psychological therapy. The current trial had a primary end point at 3 weeks but observed participants over 12 weeks of follow-up in a parallel-group design, included a trial population in which more than 90% of the participants did not have previous exposure to psilocybin, and used remote raters who were unaware of the details of the trial and the trial-group assignments to determine the primary end-point measure (MADRS total score). The manualized, time-limited approach to preparation, support, and integration of the psychedelic experience ensured safety and is not a stand-alone psychotherapy.

For participants in this trial, psilocybin therapy represented a third-, fourth-, or fifth-line treatment. The incidence of response at week 3 of 37% in the 25-mg group in our trial was numerically lower than that described for first-line treatment of major depressive disorder in several large trials of citalopram,[1] nefazodone, and escitalopram, sertraline, or venlafaxine[17] but was higher than the incidences of response reported in the STAR*D trial for second-line treatments and beyond. Pharmacokinetic research has shown dose-dependent increases in receptor occupancy and subjective effects of psilocybin across the dose range of 3 to 30 mg.[18] These findings may explain the differences in efficacy between the groups in the current trial.

Limitations of the current trial include the lack of an active comparator, the lack of an ethnically diverse participant sample, and the exclusion of persons judged to be at a clinically significant risk for suicide. The intensity of the acute subjective effect of the 25-mg and 10-mg doses, as compared with the 1-mg dose, reduces the effectiveness of the double-blind structure of the trial. We did not assess participants' ability to guess their dose assignment, and ensuring blinding is an inherent limitation of studies of drugs that produce psychedelic subjective effects. Whether other preparations of psilocybin than the proprietary one used in this trial would show the same effects cannot be determined.

In this trial of psilocybin administered in a single session with psychological support, a 25-mg dose but not a 10-mg dose resulted in a significantly greater reduction (improvement) in MADRS total scores than a 1-mg dose at 3 weeks in participants with treatment-resistant depression but was associated with adverse events. Secondary end-point results generally supported the primary analysis with the exception of 12-week sustained response, at which time point the observed numerical difference was not considered to be statistically significant. Longer and larger trials, including comparison with existing treatments for depression, are required to deter-

Table 3. Adverse Events Reported on Day 1, from Day 2 up to Week 3, and after Week 3 up to Week 12 (Safety Population).*

Adverse Event	Psilocybin, 25 mg (N=79)	Psilocybin, 10 mg (N=75)	Psilocybin, 1 mg (N=79)
		number (percent)	
Day 1			
Any adverse event	48 (61)	35 (47)	30 (38)
Any severe adverse event	3 (4)	6 (8)	1 (1)
Adverse events occurring in ≥5% of participants in any group			
Headache	19 (24)	11 (15)	13 (16)
Nausea	17 (22)	5 (7)	1 (1)
Euphoric mood	4 (5)	5 (7)	3 (4)
Fatigue	5 (6)	2 (3)	4 (5)
Insomnia	2 (3)	3 (4)	5 (6)
Anxiety	3 (4)	6 (8)	0
Mood altered	4 (5)	3 (4)	0
Dizziness	5 (6)	1 (1)	0
Paresthesia	2 (3)	4 (5)	0
Abnormal thinking	0	4 (5)	0
Any serious adverse event	0	0	0
Day 2 up to wk 3			
Any adverse event	44 (56)	36 (48)	35 (44)
Any severe adverse event	7 (9)	5 (7)	1 (1)
Adverse events occurring in ≥5% of participants in any group			
Headache	9 (11)	5 (7)	9 (11)
Insomnia	4 (5)	5 (7)	8 (10)
Anxiety	4 (5)	6 (8)	3 (4)
Fatigue	6 (8)	2 (3)	3 (4)
Suicidal ideation	5 (6)	4 (5)	2 (3)
Depression	3 (4)	3 (4)	4 (5)
Mood altered	4 (5)	0	1 (1)
Any serious adverse event	4 (5)	4 (5)	0
Suicidal ideation	2 (3)	2 (3)	0
Intentional self-injury	2 (3)	1 (1)	0
Hospitalization	0	1 (1)	0
After wk 3 up to wk 12			
Any adverse event	23 (29)	24 (32)	24 (30)
Any severe adverse event	2 (3)	3 (4)	0
Adverse events occurring in ≥5% of participants in any group			
Headache	3 (4)	2 (3)	6 (8)
Any serious adverse event	4 (5)	3 (4)	1 (1)
Suicidal behavior	3 (4)	0	0

Table 3. (Continued.)

Adverse Event	Psilocybin, 25 mg (N=79)	Psilocybin, 10 mg (N=75)	Psilocybin, 1 mg (N=79)
		number (percent)	
Intentional self-injury	0	1 (1)	1 (1)
Adjustment disorder with anxiety and depressed mood	1 (1)	0	0
Depression	0	1 (1)	0
Drug withdrawal syndrome†	1 (1)	0	0
Suicidal ideation	0	1 (1)	0

* Shown are adverse events that emerged or worsened after trial-drug administration.
† The event involved codeine withdrawal.

mine the efficacy and safety of psilocybin for treatment-resistant depression.

The views expressed are those of the authors and not necessarily those of the NHS, the National Institute for Health or Care Research, or the Department of Health and Social Care in the United Kingdom.

Supported by COMPASS Pathfinder.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

We thank the participants, without whom this research would not have been possible; the following staff members at COMPASS Pathfinder and Worldwide Clinical Trials for their contributions: Molly Lennard-Jones, Rachel Winzer, Batya Septimus, Merve Atli, Ozlem Redjep, Hannah Tadley, Nisha Thiara, Hollie Simmons, Julia Forte, Niccolo Bassani, and Alexandra Novikova; Edward Schweizer for helping with the first draft of the manuscript; Megan Croal for helping revise subsequent drafts; the staff at all the sites for help with recruitment of participants, including the Champalimaud Foundation (Portugal), Hospital del Mar (Spain), Leiden UMC (the Netherlands), and St. Pancras Clinical Research (United Kingdom); and Bill Richards, Mike Emmanuel, and Mary Johnson for their advice in designing the trial originally.

APPENDIX

The authors' full names and academic degrees are as follows: Guy M. Goodwin, F.Med.Sci., Scott T. Aaronson, M.D., Oscar Alvarez, M.R.C.Psych., Peter C. Arden, M.P.H., Annie Baker, R.G.N., James C. Bennett, M.Sc., Catherine Bird, M.Sc., Renske E. Blom, M.D., Christine Brennan, M.Sc., Donna Brusch, C.C.R.C., Lisa Burke, M.Sc., Kete Campbell-Coker, R.G.N., Robin Carhart-Harris, Ph.D., Joseph Cattell, R.G.N., Aster Daniel, R.G.N., Charles DeBattista, M.D., Boadie W. Dunlop, M.D., Katherine Eisen, Ph.D., David Feifel, M.D., Ph.D., MacKenzie Forbes, M.S., Hannah M. Haumann, M.D., David J. Hellerstein, M.D., Astrid I. Hoppe, M.Sc., Muhammad I. Husain, M.R.C.Psych., Luke A. Jelen, M.R.C.Psych., Jeanine Kamphuis, M.D., Julie Kawasaki, M.S.W., John R. Kelly, M.D., Richard E. Key, M.S.W., Ronit Kishon, Ph.D., Stephanie Knatz Peck, Ph.D., Gemma Knight, Clin.Psy.D., Martijn H.B. Koolen, M.D., Melanie Lean, D.Clin.Psy., Rasmus W. Licht, Ph.D., Jessica L. Maples-Keller, Ph.D., Jan Mars, M.Sc., Lindsey Marwood, Ph.D., Martin C. McElhiney, Ph.D., Tammy L. Miller, Psy.D., Arvin Mirow, M.D., Sunil Mistry, M.Sc., Tanja Mletzko-Crowe, M.A., Liam N. Modlin, M.B.A.C.P., René E. Nielsen, M.D., Elizabeth M. Nielson, Ph.D., Sjoerd R. Offerhaus, M.D., Veronica O'Keane, M.D., Tomáš Páleníček, Ph.D., David Printz, M.D., Marleen C. Rademaker, M.Sc., Aumer van Reemst, M.Sc., Frederick Reinholdt, M.A., Dimitris Repantis, M.D., James Rucker, M.D., Samuel Rudow, B.S., Simon Ruffell, M.D., A. John Rush, M.D., Robert A. Schoevers, M.D., Mathieu Seynaeve, M.R.C.Psych., Samantha Shao, B.S., Jair C. Soares, M.D., Metten Somers, Ph.D., Susan C. Stansfield, Ph.D., Diane Sterling, Ph.D., Aaron Strockis, B.A., Joyce Tsai, Ph.D., Lucy Visser, M.Sc., Mourad Wahba, M.R.C.Psych., Samuel Williams, M.Sc., Allan H. Young, F.R.C.Psych., Paula Ywema, B.Sc., Sidney Zisook, M.D., and Ekaterina Malievskaia, M.D.

The authors' affiliations are as follows: COMPASS Pathfinder (G.M.G., J.C.B., L.M., S.M., S.C.S., J.T., S.W., E.M.), the Department of Psychological Medicine, Institute of Psychiatry, Psychology, and Neuroscience, King's College London (C. Bird, L.A.J., G.K., L.N.M., F.R., J.R., S. Ruffell, M. Seynaeve, A.H.Y.), the National Institute for Health and Care Research Clinical Research Facility, King's College Hospital NHS Foundation Trust (K.C.-C., J.C., A.D.), and South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital (L.A.J., L.N.M., J.R., A.H.Y.), London, and the Cumbria, Northumberland, Tyne and Wear Foundation Trust and Newcastle University, Newcastle (M.W.) — all in the United Kingdom; the Institute for Advanced Diagnostics and Therapeutics, Sheppard Pratt, Baltimore (S.T.A., M.F., T.L.M., S. Rudow); Sant Joan de Déu Hospital and the Sant Joan de Déu Research Foundation, Barcelona (O.A.); SUNY Downstate College of Medicine (P.C.A.), the New York State Psychiatric Institute (D.J.H., R.E.K., R.K., M.C.M., E.M.N.), and the Department of Psychiatry, Columbia University (D.J.H., R.K., M.C.M., E.M.N.) — all in New York; the Department of Psychiatry, Trinity Centre for Health Sciences, Tallaght University Hospital, Dublin (A.B., C. Brennan, L.B., J.R.K., V.O.); the Department of Psychiatry, University Medical Center (UMC) Utrecht Brain Center, University Medical Center Utrecht, Utrecht (R.E.B., H.M.H., A.I.H., M.H.B.K., S.R.O., M.C.R., A.R., M. Somers, L.V., P.Y.), the Research Department, GGz Centraal Innova, Amersfoort (R.E.B.), and the Department of Psychiatry, UMC Groningen, Groningen (J. Kamphuis, J.M., R.A.S.) — all in the Netherlands; the Department of Psychiatry, University of California San Diego (D.B., J. Kawasaki, S.K.P., D.P., S.S., A.S., S.Z.), and Kadima Neuropsychiatry Institute (D.F., S.K.P., A.M., D.S.), La Jolla, the Weill Institute for Neurosciences, University of California San Francisco, San Francisco (R.C.-H.), and the Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford (C.D., K.E., M.L.) — all in California; the Department of

Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta (B.W.D., J.L.M.-K., T.M.-C.); the Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, and the Department of Psychiatry, University of Toronto, Toronto (M.I.H.); the Department of Psychiatry, Aalborg University Hospital, and the Department of Clinical Medicine, Aalborg University, Aalborg, Denmark (R.W.L., R.E.N.); the National Institute of Mental Health, Klecany, Czech Republic (T.P.); Charité–Universitätsmedizin Berlin, Department of Psychiatry and Psychotherapy, Campus Benjamin Franklin, Berlin (D.R.); Duke University School of Medicine, Durham, NC (A.J.R.); and the University of Texas (UT) Harris County Psychiatric Center and the UT Center of Excellence on Mood Disorders, Department of Psychiatry and Behavioral Sciences, UT Houston Medical School, Houston (J.C.S.).

REFERENCES

1. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 2006;163:1905-17.
2. Warden D, Rush AJ, Trivedi MH, Fava M, Wisniewski SR. The STAR*D project results: a comprehensive review of findings. Curr Psychiatry Rep 2007;9:449-59.
3. Pigott HE. The STAR*D trial: it is time to reexamine the clinical beliefs that guide the treatment of major depression. Can J Psychiatry 2015;60:9-13.
4. Passie T, Seifert J, Schneider U, Emrich HM. The pharmacology of psilocybin. Addict Biol 2002;7:357-64.
5. Grob CS, Danforth AL, Chopra GS, et al. Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer. Arch Gen Psychiatry 2011;68:71-8.
6. Ross S, Bossis A, Guss J, et al. Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: a randomized controlled trial. J Psychopharmacol 2016;30:1165-80.
7. Griffiths RR, Johnson MW, Carducci MA, et al. Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: a randomized double-blind trial. J Psychopharmacol 2016;30:1181-97.
8. Davis AK, Barrett FS, May DG, et al. Effects of psilocybin-assisted therapy on major depressive disorder: a randomized clinical trial. JAMA Psychiatry 2021;78:481-9.
9. Carhart-Harris R, Giribaldi B, Watts R, et al. Trial of psilocybin versus escitalopram for depression. N Engl J Med 2021;384:1402-11.
10. Carhart-Harris RL, Bolstridge M, Rucker J, et al. Psilocybin with psychological support for treatment-resistant depression: an open-label feasibility study. Lancet Psychiatry 2016;3:619-27.
11. Tai SJ, Nielson EM, Lennard-Jones M, et al. Development and evaluation of a therapist training program for psilocybin therapy for treatment-resistant depression in clinical research. Front Psychiatry 2021;12:586682.
12. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382-9.
13. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59:Suppl 20:22-33.
14. Chandler GM, Iosifescu DV, Pollack MH, Targum SD, Fava M. Validation of the Massachusetts General Hospital Antidepressant Treatment History Questionnaire (ATRQ). CNS Neurosci Ther 2010;16:322-5.
15. Williams JB, Kobak KA. Development and reliability of a structured interview guide for the Montgomery Asberg Depression Rating Scale (SIGMA). Br J Psychiatry 2008;192:52-8.
16. Posner K, Brent D, Lucas C, et al. Columbia-suicide severity rating scale (C-SSRS). New York: Research Foundation for Mental Hygiene, 2008.
17. Saveanu R, Etkin A, Duchemin A-M, et al. The international Study to Predict Optimized Treatment in Depression (iSPOT-D): outcomes from the acute phase of antidepressant treatment. J Psychiatr Res 2015;61:1-12.
18. Madsen MK, Fisher PM, Burmester D, et al. Psychedelic effects of psilocybin correlate with serotonin 2A receptor occupancy and plasma psilocin levels. Neuropsychopharmacology 2019;44:1328-34.