Brexanolone in Postpartum Depression: *Post Hoc* Analyses to Help Inform Clinical Decision-Making

Margaret E. Gerbasi, PhD,[1] Samantha Meltzer-Brody, MD, MPH,[2] Sarah Acaster, MS,[3] Moshe Fridman, PhD,[4] Vijayveer Bonthapally, PhD,[1] Paul Hodgkins, PhD,[1] Stephen J. Kanes, MD, PhD,[1] and Adi Eldar-Lissai, PhD[1]

Abstract

Background: Brexanolone (BRX) injection was approved by the United States Food and Drug Administration in 2019 for the treatment of adults with postpartum depression (PPD) based on two Phase 3 clinical trials.
Materials and Methods: Data from the three trials were combined. PPD-specific 17-item Hamilton Rating Scale for Depression (HAMD-17) group-level minimal important difference (MID) and patient-level meaningful change (meaningful change threshold [MCT]) were estimated and applied to differences in BRX versus placebo (PBO) at hour 60 (primary endpoint) and day 30 (end of trial follow-up). Likelihood of HAMD-17 response and remission and Clinical Global Impression of Improvement (CGI-I) response for BRX versus PBO were assessed at hour 60 and as sustained through day 30 using relative risk. Associated number needed to treat (NNT) and number needed to harm (NNH) values were also estimated.
Results: Two-hundred nine patients were included. The average HAMD-17 MID estimate was −2.1; the least-squared mean difference between BRX and PBO exceeded this at hour 60 and day 30. Minimal, moderate, and large MCTs were estimated to be −9, −15, and −20 points, respectively. Significantly more BRX-treated than PBO-treated patients achieved minimal, moderate, and large change (all $ps < 0.05$) at hour 60 and large meaningful response at day 30 ($p < 0.05$). BRX-treated patients were more likely to sustain HAMD-17 remission and CGI-I response through day 30 versus PBO. NNTs ranged from 4 to 8, with NNH of 97.
Conclusions: BRX provided meaningful changes relative to PBO, rapid (hour 60), and sustained improvements (day 30) in PPD symptoms, low NNT, and large NNH. These results may help inform treatment decision-making. Clinicaltrials.gov registration numbers: NCT02614547, NCT02942004, and NCT02942017.

Keywords: postpartum depression, brexanolone, meaningful change, number needed to treat, number needed to harm

Introduction

POSTPARTUM DEPRESSION (PPD) IS one of the most common medical complications during and after pregnancy.[1–7] In the United States, an estimated 11.5% of women giving birth experience symptoms of PPD, with global estimates of 17.7%.[6,8] Suicide is strongly associated with depressive symptoms and is a leading cause of pregnancy-related mortality.[9–16] Mothers with PPD may experience difficulties with physical functioning and bonding with their infants.[17,18] Maternal PPD is associated with poorer cogni-

tive and physical development of the child as well as higher rates of depression and poorer academic performance as the child grows into adolescence.[19–22]

In 2019, the Food and Drug Administration approved brexanolone (BRX) injection, a neuroactive steroid chemically identical to allopregnanolone, for the treatment of PPD in adults based on two Phase 3 randomized clinical trials (RCTs) (NCT02942004 and NCT02942017) following Breakthrough Therapy Designation in 2016.[23] BRX is administered via a 60-hour intravenous infusion in a monitored health care setting, and is available only through a restricted

[1]Sage Therapeutics, Inc., Cambridge, Massachusetts, USA.
[2]Department of Psychiatry, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA.
[3]Acaster Lloyd Consulting Ltd., London, United Kingdom.
[4]AMF Consulting, Los Angeles, California, USA.

program under a Risk Evaluation and Mitigation Strategy (REMS) requiring patient enrolment, a restricted distribution program, and administration in a certified health care facility with monitoring by a health care provider for excessive sedation or sudden loss of consciousness.[24] In the Phase 3 trials, BRX achieved the primary endpoint of significantly greater reduction in depressive symptoms, as measured by least-squares mean reduction in 17-item Hamilton Rating Scale for Depression (HAMD-17) total score at the end of infusion (hour 60) compared with placebo (PBO).[23] Pooled analyses of two Phase 3 trials and a Phase 2 trial conducted under the same umbrella protocol showed a significant difference in HAMD-17 total score for BRX versus PBO as early as 24 hours and through the duration of the study period (day 30).[23] Nearly twice as many BRX-treated patients achieved HAMD-17 remission (score ≤7) at hour 60 (primary endpoint) compared to PBO-treated patients (50% vs. 26%, $p < 0.001$). The most common adverse reactions (incidence ≥5% and at least twice the rate of PBO) were sedation/somnolence, dry mouth, loss of consciousness, and flushing/hot flush.

While the HAMD-17 is considered the gold standard assessment in clinical trials for major depression and PPD, there is no PPD-specific estimate of meaningful change to facilitate patient-centered evaluation of the data.[25–28] Meaningful change thresholds (MCTs) relate to patient-level meaningful change, rather than group change, and are estimated using both distribution-based methods and anchor-based methodologies.[29] Distribution-based methods alone can be used to estimate group-level meaningful difference or change, known as a minimal important difference (MID). MCTs and MIDs are population- and disease-specific and should be established within the specific population of interest.[30] While a range of meaningful differences have been reported for HAMD-17 in populations of patients with major depressive episode (MDE), none has previously been reported based on a PPD population.[31,32]

In addition to the core trial endpoints, more applied, patient-centric secondary analysis of key trial data could further inform health care decision-making. For example, the evaluation of relative risk (RR), number needed to treat (NNT), and number needed to harm (NNH) specifically, can aid clinicians and health care decision-makers as they seek to understand the likelihood of favorable treatment outcomes.[33–36]

Thus, the current analyses aimed to explore outcomes relevant to clinicians, patients, and health care decision-makers in the evaluation of treatments for PPD. More specifically, the objectives of this study were to (1) estimate and apply PPD-specific HAMD-17 MCTs at hour 60 and day 30 and (2) further evaluate the proportion of patients treated with BRX versus PBO achieving HAMD-17 remission and response and Clinical Global Impression of Improvement (CGI-I) response by estimating the associated risk (RR) and NNT at hour 60 and through trial follow-up (day 30).

Materials and Methods

Study design and participants

The current report presents a *post hoc* analysis of clinician-reported outcomes data pooled together from previously published phase 2 and phase 3 clinical trials (hereafter "combined dataset"), which examined the safety and efficacy of BRX injection compared with PBO injection in pa-

tients with moderate to severe PPD (Study A: NCT02614547; Study B: NCT02942004; and Study C: NCT02942017).[23,37]

Full descriptions of the trial designs, inclusion and exclusion criteria have been published previously.[23,37] In brief, one phase 2 (Study A) and two phase 3 (Studies B and C) multicenter, randomized, double-blind, PBO-controlled trials were conducted across the United States of America under an umbrella protocol (multiple, similar clinical trials conducted under one Institutional Review Board application). The clinical trials included women 18–45 years old, ≤6 months postpartum, with a diagnosis of moderate-to-severe PPD and a qualifying HAMD-17 (Studies A and B: ≥26; C: 20–25) who were either randomized 1:1:1 to receive 60 hour infusion of BRX 90 μg/kg/h, BRX 60 μg/kg/h, or PBO (Study B), or 1:1 to receive BRX or PBO (Studies A and C).

Once randomized, patients were treated in a medically supervised setting for 72 hours: 60 hours of continuous study drug infusion and 12 hours for completion of assessments. Patients were followed until day 30, with clinical and safety assessments on days 7 and 30. The primary endpoint in each trial was the least-squares mean difference in change from baseline in HAMD-17 total score at hour 60.

The studies were conducted with adherence to and compliance with the Declaration of Helsinki and Good Clinical Practice Guidelines. The study protocols were reviewed and approved by relevant Institutional Review Boards or independent Ethics Committees, and all patients provided written informed consent before study inclusion.

Outcomes

The HAMD-17 is a clinician-reported scale that evaluates core symptoms of depression.[25] Items are scored on a 0–4 (0 = none/absent and 4 = severe) or a 0–2 (0 = absent/none and 2 = clearly present) scale. Individual item scores are summed to compute the total score, which ranges from 0 to 52, with higher scores indicating more severe depression.[25] Clinical response on the HAMD-17 is defined as a ≥50% reduction from baseline; remission is defined as a HAMD-17 total score of ≤7.

The CGI-I is also a clinician-reported scale that uses a seven-point Likert scale to measure the overall improvement in the patient's condition.[38] Response choices include the following: 1 = very much improved, 2 = much improved, 3 = minimally improved, 4 = no change, 5 = minimally worse, 6 = much worse, and 7 = very much worse.[38] CGI-I responders are defined as patients receiving a rating of 1 (very much improved) or 2 (much improved).

Statistical analyses

The analyses were conducted on the BRX and PBO combined efficacy dataset. The dataset included all randomized patients who started infusion and had a valid baseline HAMD-17 and at least one postbaseline HAMD-17.

To estimate HAMD-17 total score patient-level MCT, distribution and anchor-based approaches were used. The distribution-based methods were as follows: ½ standard deviation (SD) at baseline and one standard error of measurement (SEM; SD × sqrt of [1 − reliability coefficient]). The reliability coefficient used to calculate the SEM was the intraclass correlation coefficient (ICC) between baseline and hour 2 HAMD-17 total scores for those subjects rated as "no change" using the

CGI-I. The anchor-based approach used the CGI-I: the mean HAMD-17 total change from baseline was calculated for each CGI-I response at hour 60 and day 30, independent of the treatment arm. A level of minimal, moderate, and large meaningful change was estimated based on minimal, much, and very much improved CGI-I response categories, respectively. Mean data calculated across both timepoints for each CGI-I response was rounded to the nearest integer to yield a level of change achievable by an individual patient. The anchor-based values were compared to the distribution-based values to ensure that patient-level MCTs were greater than distribution-based MID estimates. Fisher's exact test was used to explore the difference in the proportion of subjects reporting minimal, moderate, and large meaningful change at hour 60 and day 30 across treatment groups.

To evaluate the proportion of patients demonstrating sustained HAMD-17 response and remission and CGI-I response, "sustained" was defined as the continued categorization of a responder and remitter at hour 60, day 7, and day 30. The 95% confidence intervals (CIs), and corresponding *p*-values, for BRX versus PBO were estimated using Fisher's exact test at hour 60 and for sustained responders and remitters at day 30. The RR and NNT were calculated based on the probability of

being a responder and remitter (RR; P_{tx}/P_{pbo}) or the proportion of responders and remitters (NNT; $1/[P_{tx} - P_{pbo}]$) in each treatment arm at hour 60, and sustained responders and remitters at day 30. The NNH was estimated based on the proportion of subjects in each treatment arm discontinuing the study drug due to an adverse event (AE).[34] Detailed safety and discontinuation information has been previously reported by Meltzer-Brody et al.[23] Asymptotic confidence limits were calculated for the RR; 95% CI were calculated for NNT and NNH based upon Wilson score intervals.[33]

Due to the exploratory nature of the analysis, all statistical tests were two-sided and *p*-values ≤0.05 were considered significant. Analyses were *post hoc* in nature and were not adjusted for multiplicity. All analyses were performed using SAS® software Version 9.4.

Results

The combined efficacy data set included 209 patients: 102 treated with BRX and 107 treated with PBO. Demographics and baseline characteristics were generally well-balanced across treatment groups (Table 1). The mean age in each group was 27 and the mean HAMD-17 baseline score

TABLE 1. BASELINE DEMOGRAPHICS AND CHARACTERISTICS FOR PATIENTS TREATED WITH BREXANOLONE OR PLACEBO

	PBO n = 107	*BRX* n = 102	*Overall* n = 209
Characteristics			
Age, years, mean (SD)	27.42 (5.85)	27.83 (5.96)	27.62 (5.89)
Ethnicity			
Hispanic or Latino, *n* (%)	21 (20%)	17 (17%)	38 (18%)
Not Hispanic or Latino, *n* (%)	86 (80%)	85 (83%)	171 (82%)
Race			
Black or African American, *n* (%)	40 (37%)	37 (36%)	77 (37%)
White, *n* (%)	65 (61%)	61 (60%)	126 (60%)
Other, *n* (%)	2 (2%)	4 (4%)	6 (3%)
Height, cm, mean (SD)[a]	163.67 (8.16)	164.13 (6.44)	163.89 (7.36)
Weight, kg, mean (SD)	83.70 (23.85)	84.56 (23.58)	84.12 (23.66)
Body mass index, kg/m^2, mean (SD)	31.16 (8.20)	31.29 (8.13)	31.22 (8.15)
Personal history of psychiatric disorders			
Depression (non-PPD), *n* (%)	43 (40%)	37 (36%)	80 (38%)
Anxiety, *n* (%)	34 (32%)	38 (37%)	72 (34%)
Premenstrual dysphoric disorder, *n* (%)	3 (3%)	6 (6%)	9 (4%)
Substance abuse, *n* (%)	0 (0%)	1 (1%)	1 (0%)
Schizophrenia, *n* (%)	5 (5%)	2 (2%)	7 (3%)
Previous PPD episodes, *n* (%)	37 (35%)	27 (26%)	64 (31%)
Family history of PPD			
Yes, *n* (%)	24 (22%)	34 (33%)	58 (28%)
No, *n* (%)	83 (78%)	68 (67%)	151 (72%)
Onset of PPD			
Third trimester, *n* (%)	30 (28%)	22 (22%)	52 (25%)
Within 4 weeks of delivery, *n* (%)	77 (72%)	80 (78%)	157 (75%)
Antidepressant use at baseline			
Yes, *n* (%)	25 (23%)	22 (22%)	47 (22%)
No, *n* (%)	82 (77%)	80 (78%)	162 (78%)
Baseline HAMD-17[b]			
HAMD-17 total, mean (SD)	25.68 (3.64)	25.50 (3.46)	25.59 (3.55)

[a]Weight, height, and body mass index data were measured at screening.
[b]Assessed before injection on day 1.
BRX, brexanolone 90 μg/kg/h injection; HAMD-17, 17-item Hamilton Rating Scale for Depression; PBO, placebo injection; PPD, postpartum depression; SD, standard deviation.

(preinfusion) was 25. Of the 209 patients included, 126 were white (60%), 77 (37%) were black or African American, and 38 (18%) were Hispanic or Latino. Across both groups, a higher proportion of patients had onset of PPD during 4 weeks postpartum than during the third trimester. A similar proportion of patients were taking antidepressants at baseline in both groups.

PPD-specific meaningful patient level change

The distribution-based MID estimates based on ½ SD at baseline and 1 SEM were 1.8 and 2.4, respectively, with an average MID of 2.1. The HAMD-17 ICC used to evaluate the SEM was $r = 0.55$. Significant least-squares mean difference in HAMD-17 change scores between BRX and PBO at the end of hour 60 (-4.1 [95% CI: -6.0 to -2.3], $p < 0.0001$), as well as hours 24 (-3.0 [95% CI: -4.8 to -1.2], $p = 0.0012$), day 30 (-2.6 [95% CI: -4.7 to -0.4], $p = 0.0213$) have previously been reported.[23] Figure 1 applies the ½ SD at baseline and 1 SEM MID estimates to the combined BRX clinical trial data. The HAMD-17 group differences between BRX and PBO remain meaningful when applying estimates for the ½ SD MID, mean MID, and 1 SEM MID to the mean differences from hour 24 onward. The 95% CI estimates for hour 60 are also greater than the ½ SD MID and mean MID estimates.

The patient-level HAMD-17 meaningful change estimates for minimal, moderate, and large change were -9, -15, and -20, respectively. These were based on the CGI-I mean change scores at hour 60 and day 30 for those classified as minimally improved (mean [M] hour $60 = -9.65$; M day $30 = -8.73$), much improved (M hour $60 = -15.31$; M day $30 = -15.56$), or very much improved, (M hour $60 = -20.12$; M day $30 = -20.89$), respectively. The highest change score reported for patients showing no change on the CGI-I was -4. As the MCT estimates all exceeded both distribution-based MID estimates and the highest level of improvement observed in patients considered stable, these were considered valid and applied to the hour 60 and day 30 data.

Figure 2 applies patient-level meaningful change values to hour 60 and day 30 and demonstrates that compared to PBO, BRX had a consistently higher level of response with a significantly greater proportion of patients demonstrating minimal (87% vs. 68%, Δ19%, $p < 0.01$), moderate (67% vs. 42%, Δ25%, $p < 0.001$), and large meaningful change (30% vs. 15%, Δ15%, $p < 0.05$) at hour 60, and large meaningful change at day 30 (41% vs. 26%, Δ15%, $p < 0.05$).

Response and remission: RR, NNT, and NNH

There were significantly higher proportions of HAMD-17 (74% vs. 56%, $p < 0.01$) and CGI-I (81% vs. 54%, $p < 0.001$) responders, and HAMD-17 remitters (50% vs. 26%, $p < 0.001$) in the BRX arm compared with PBO at hour 60 (Table 2). The proportion of sustained HAMD-17 remitters and CGI-I responders from hour 60 through day 30 was also significantly higher in the BRX arm compared with PBO (Table 2). The likelihood of being a responder or remitter, as measured by RR, was significantly higher in the BRX arm relative to PBO for both HAMD-17 and CGI-I responders and HAMD-17 remitters at hour 60, and for sustained CGI-I responders and HAMD-17 remitters at day 30. The proportion of sustained HAMD-17 responders was numerically, but not statistically higher at day 30 ($p = 0.065$, Table 3). As shown in Table 3, the RR in the BRX arm for responders was 1.34–1.50 times higher than PBO at hour 60 and 1.37–1.72 times higher than PBO at day 30. Depending on scale used, at hour 60, the NNT for BRX ranged from 4 to 6 and was 4–8 as sustained at day 30 (Table 3). The proportion of subjects who discontinued the studies due to an AE was $n = 2/102$ (2%) 2/102 in the BRX arm and $n = 1/107$ (0.9%) in the PBO arm, with a nonsignificant NNH of 97 (95% CI: 17 to -30).

Discussion

Clinicians, patients, and health care decision-makers (*e.g.*, payers, policy makers) evaluating treatment options for PPD
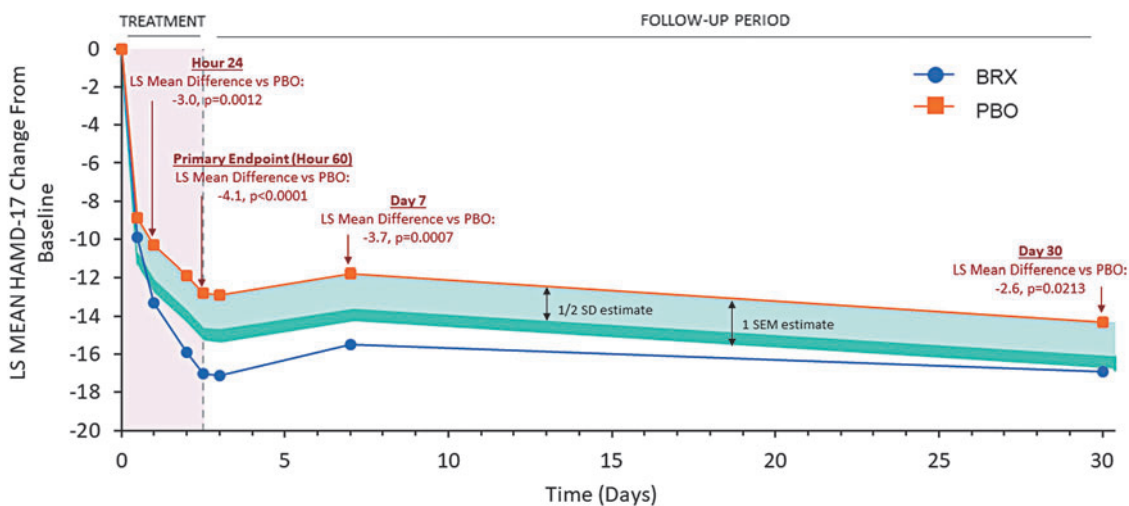

FIG. 1. Application of ½ SD and 1 SEM minimal important difference estimates to BRX versus PBO clinical trial data. The HAMD-17 group differences between BRX and PBO remain meaningful when applying estimates for the ½ SD estimate of minimal important difference and 1 SEM estimate of minimal important difference to the mean differences from hour 24 onward. LS, least-squares; BRX, brexanolone injection 90 μg/kg/h; HAMD-17, 17-item Hamilton Rating Scale for Depression; PBO, placebo; SD, standard deviation; SEM, standard error of measurement.

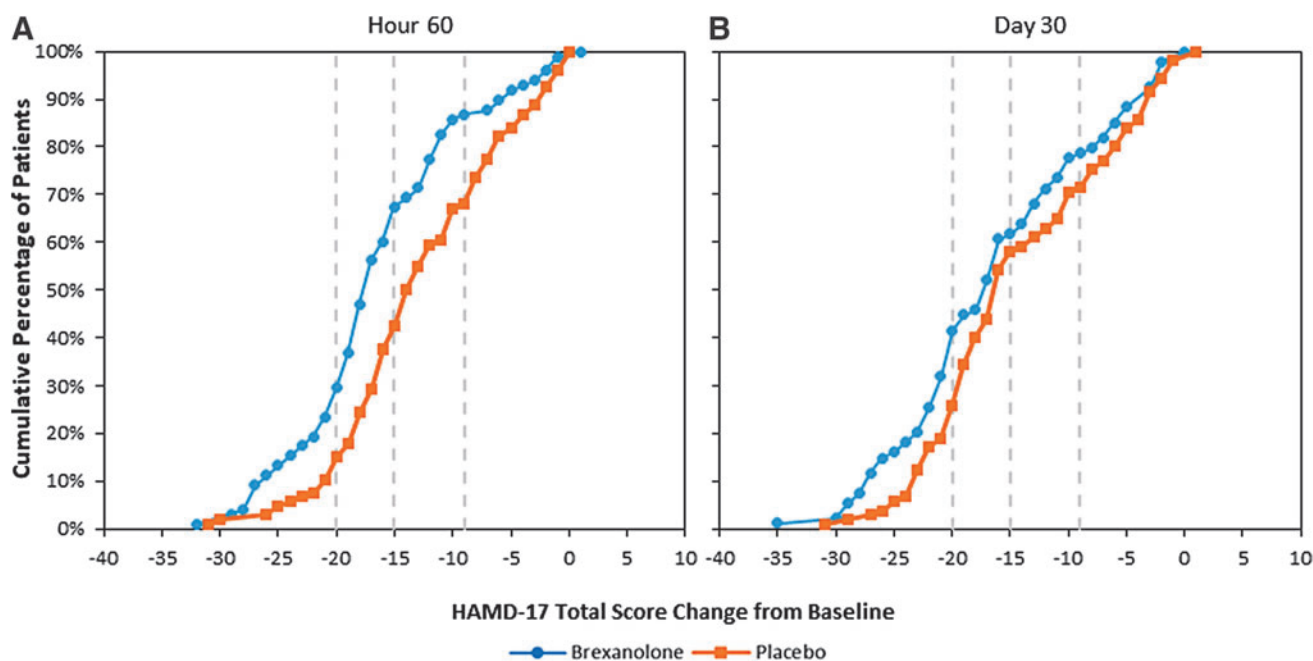**FIG. 2.** HAMD-17 patient-level meaningful change from baseline at hour 60 (**A**) and day 30 (**B**) with BRX versus PBO. Improvement cutoffs were −9, −15, and −20 for minimal, moderate, and large meaningful change, respectively. BRX, brexanolone injection 90 μg/kg/h; PBO, placebo injection.

need a clear understanding of possible treatment options and their benefit profiles. Management of patients with PPD generally follows a stepped-care approach. Patients with mild symptoms are treated through low-intensity interventions (*e.g.*, group therapy). However, for patients with symptoms that do not respond, are more severe, or present acutely, pharmacologic interventions, either alone or adjunctive to low-intensity interventions, are typically offered.[39] Before the approval of BRX, antidepressants, such as selective serotonin reuptake inhibitors (SSRIs), were the primary pharmacologic treatment option available for patients with PPD despite not being indicated specifically for PPD. Although there are many published RCTs and analyses evaluating SSRIs for the treatment of patients with MDE,[40,41] the paucity of evidence within PPD makes it difficult to evaluate these pharmacological treatment options in the PPD population.[42] Barriers to effective treatment with SSRIs include delay of 6–12 weeks to achieve optimal improvement, frequent subtherapeutic dosing, poor patient adherence due to

side effects, withdrawal effects experienced by the majority of patients, and inadequate clinician follow-up.[43–49] While the rapidity of effect, hypothesized mechanism of action, and modality of BRX may allow it to circumvent many of these barriers, its intravenous administration, associated with excessive sedation and loss of consciousness requiring a REMS, and the evolving reimbursement landscape surrounding its administration are sizeable obstacles that will likely initially confine its real-world placement in the treatment cascade for PPD to be primarily among patients requiring rapid resolution of symptoms, or those who have not responded to other treatments.

To our knowledge, this is the first article to propose PPD-specific estimates of HAMD-17 meaningful differences and patient-level meaningful change. While the primary objective of our analysis was to estimate and apply PPD-specific MCTs, the distribution-based analysis also provides an estimate of PPD-specific meaningful group difference (MID). Two different distribution-based approaches were

TABLE 2. PROPORTION OF PATIENTS ACHIEVING RESPONSE AND REMISSION AT HOUR 60 AND SUSTAINED RESPONSE
AND REMISSION FROM HOUR 60 TO DAY 30

Parameters	Hour 60				Sustained hour 60 to day 30			
	PBO n = 106–107	*BRX* n = 98	*Delta*	p *Value*	*PBO* n = 104–106	*BRX* n = 95–101	*Delta*	p *Value*
HAMD-17 response[a]	59 (56%)	73 (74%)	18%	<0.01	38 (37%)	49 (50%)	13%	0.065
HAMD-17 remission[b]	28 (26%)	49 (50%)	24%	<0.001	17 (16%)	32 (32%)	16%	0.014
CGI-I response[c]	57 (54%)	79 (81%)	27%	<0.001	37 (36%)	58 (61%)	25%	<0.001

[a]HAMD-17 response: ≥50% reduction from baseline.
[b]HAMD-17 remission: score ≤7.
[c]CGI response: very much/much improved.
BRX, brexanolone 90 μg/kg/h injection; PBO, placebo injection; CGI-I, Clinical Global Impression of Improvement.

Table 3. Relative Risks and Number Needed to Treat for Brexanolone Versus Placebo at Hour 60 for Patients Achieving Response and Remission and Hour 60 to Day 30 for Sustained Response and Remission

	Hour 60		Hour 60 to day 30	
Parameters	RR (95% CI)	NNT (95% CI)	RR (95% CI)	NNT (95% CI)
HAMD-17 response[a]	1.34 (1.09 to 1.64)	6 (4 to 18)	1.37 (0.99 to 1.89)	8 (4 to 557)
HAMD-17 remission[b]	1.89 (1.3 to 2.75)	5 (3 to 10)	1.96 (1.16 to 3.30)	7 (4 to 27)
CGI-I response[c]	1.50 (1.23 to 1.83)	4 (3 to 8)	1.72 (1.27 to 2.33)	4 (3 to 9)

[a]HAMD-17 response: ≥50% reduction from baseline.
[b]HAMD-17 remission: score ≤7.
[c]CGI response: very much/much improved.
PBO: $n = 106–107$ at hour 60 and $n = 104–106$ at hour 60 to day 30; BRX: $n = 98$ at hour 60; $n = 95–101$ at hour 60 to day 30.
CI, confidence interval; NNT, number needed to treat; RR, relative risk.

taken: the ½ SD (MID 1.8) and 1 SEM (MID 2.4). However, given the instability of the sample in an active treatment setting, as illustrated by the relatively low HAMD-17 ICC of 0.55, the ½ SD MID of 1.8 may represent the best estimate of a HAMD-17 MID in a moderate to severe PPD population. The previously reported significant least-squares mean differences between BRX and PBO are meaningful when all MID estimates are applied to the means, as well as when the ½ SD MID and mean MID are applied to the 95% CI estimates for hour 60.[23]

The MID estimates reported here appear lower than other published MID associated with the HAMD-17 (MID = 11).[31,32] However, previous estimates were calculated in MDE populations using CGI-I anchor-based methods for patient-level meaningful change and are thus inappropriate for interpreting group differences, being instead more comparable to the reported MCT estimates (no meaningful change ≤4; minimal change −9; moderate −15; large −20). In this context the previously reported values are reasonably well-aligned with those reported here, but the PPD-specific MCTs provide more granularity across different levels of response.

Once applied, a significantly higher proportion of patients demonstrated minimal, moderate, and large levels of meaningful improvement with BRX relative to PBO as rapidly as 60 hours after treatment. In addition, BRX provided a significantly higher proportion of patients with a large meaningful change (≥−20) at day 30. All these estimates provide valuable additions to the interpretation of HAMD-17 (symptom) improvement in PPD, permitting population-specific exploration of HAMD-17 data with greater granularity regarding levels of improvement than the universal response and remission definitions.

In line with previous published findings, in these *post hoc* analyses, BRX demonstrated rapid (by hour 60) and high rates of response and remission (HAMD-17 and CGI-I response: 74% and 81% respectively, HAMD-17 remission: 50%) and sustained to day 30 (HAMD-17 and CGI-I response: 50% and 61%, respectively, HAMD-17 remission: 32%) compared with PBO.[23,37] Notably, this rapid effect was also observed despite the variable PBO response, which is well documented in depression and can result in high PBO responses.[50–52]

The RR, NNT, and NNH data presented here can be used by health care decision-makers to understand the magnitude and the relevance of BRX treatment effects.[34–36] Patients treated with BRX showed significantly higher probabilities of

response and remission at 60 hours (RR: 1.34–1.50) and sustained response and remission at day 30 (RR: 1.37–1.72) compared with PBO. Single-digit NNTs for both the rapid (hour 60) and sustained (through day 30) outcomes support the robustness of the observed BRX effect in PPD. To contextualize these results beyond the current dataset, the NNT estimates observed for BRX at hour 60 and through day 30 were lower than, or equal to, the NNTs (6–10) reported by Citrome et al.[34] in an indirect comparison of 34 antidepressant clinical trials in MDE, with response measured after 6–10 weeks of treatment. In addition, the data presented herein demonstrate that BRX had low discontinuation rates due to an AE with an associated NNH estimate of 97. The NNH for BRX was nonsignificant but higher than most trials included in the same indirect comparison (NNH: 7–43).[34]

Although no head-to-head trials of BRX versus SSRIs for the treatment of PPD have been conducted, a recent evaluation of the efficacy of these treatments using match-adjusted indirect comparisons found that BRX demonstrated larger improvement in PPD symptoms compared to SSRIs.[53] A recent cost-effectiveness analysis also indicates that treatment of adults with PPD using BRX is cost-effective compared to SSRIs over an 11-year time horizon, particularly for patients with severe symptoms, based on a United States health care payer perspective.[54]

There are limitations to this study. As this analysis was conducted on clinical trial data, estimating MCT values was not an objective at study onset, therefore the estimates reported here should be validated on an external dataset, and ideally additional anchors, including patient-reported anchors, should also be included. The MID values may be improved in a nonclinical trial sample, as a noninterventional study may provide a more accurate estimate of reliability (ICC), but the ½ SD value provides a good initial estimate to be further validated. The maximum follow-up of 30 days after the end of study treatment may be considered a limitation, however, this follow-up duration is similar to that used after primary endpoint assessment (from 4–24 weeks) in clinical trials of antidepressants in PPD.[42] Additional research should evaluate the real-world efficacy of BRX compared to SSRIs to provide the optimal inputs to inform clinical decision-making.

Conclusions

The results of these analyses provide estimates of meaningful group differences and patient-level change specific to

PPD populations. Applying these estimates to BRX, the first treatment indicated specifically for women with PPD, using combined trial data indicates that both group- and patient-level improvements for BRX-treated patients relative to PBO-treated patients were clinically meaningful as early as 60 hours and were sustained to day 30. These estimates may also be used by future researchers and health care decision-makers in application to other PPD treatments. Patients treated with BRX demonstrated a higher likelihood of rapid and sustained HAMD-17 and CGI-I improvements. The data reported herein further support the positive clinical profile of BRX and provide valuable insights for health care decision-makers evaluating its placement within the PPD treatment cascade.

Author Disclosure Statement

Funding Information

References

1. Martin JA, Hamilton BE, Osterman MJK, Driscoll AK. Births: Final data for 2018. National Vital Statistics Reports, 2019;68. Available at: https://cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_13–508.pdf Accessed March 31, 2020.

2. Callaghan WM, Kuklina EV, Berg CJ. Trends in postpartum hemorrhage: United States, 1994–2006. Am J Obstet Gynecol 2010;202:353.e1–353.e6.

3. Centers for Disease Control and Prevention. Diabetes during pregnancy, 2018. Available at: https://cdc.gov/reproductivehealth/maternalinfanthealth/diabetes-during-pregnancy.htm Accessed March 31, 2020.

4. Centers for Disease Control and Prevention. Data on selected pregnancy complications in the United States, 2019. Available at: https://cdc.gov/reproductivehealth/maternalinfanthealth/pregnancy-complications-data.htm Accessed March 31, 2020.

5. DeSisto Cl, Kim SY, Sharma AJ. Prevalence estimates of gestational diabetes mellitus in the United States, pregnancy risk assessment monitoring system (PRAMS), 2007–2010. Prev Chronic Dis 2014;11:E104.

6. Ko JY, Rockhill KM, Tong VT, Morrow B, Farr SL. Trends in postpartum depressive symptoms-27 states, 2004, 2008, and 2012. MMWR Morb Mortal Wkly Rep 2017;66:153–158.

7. Roberts JM, August PA, Bakris G, et al. Hypertension in pregnancy. Report of the American College of Ob-

8. stetricians and Gynecologists' task force on hypertension in pregnancy. Obstet Gynecol 2013;122:1122–1131.

8. Hahn-Holbrook J, Cornwell-Hinrichs T, Anaya I. Economic and health predictors of national postpartum depression prevalence: A systematic review, meta-analysis, and meta-regression of 291 studies from 56 countries. Front Psychiatry 2017;8:248.

9. Appleby L, Mortensen PB, Faragher EB. Suicide and other causes of mortality after post-partum psychiatric admission. Br J Psychiatry 1998;173:209–211.

10. Bodnar-Deren S, Klipstein K, Fersh M, Shemesh E, Howell EA. Suicidal ideation during the postpartum period. J Womens Health 2016;25:1219–1224.

11. Gressier F, Guillard V, Cazas O, Falissard B, Glangeaud-Freudenthal NM, Sutter-Dallay AL. Risk factors for suicide attempt in pregnancy and the post-partum period in women with serious mental illnesses. J Psychiatr Res 2017;84:284–291.

12. Lindahl V, Pearson JL, Colpe L. Prevalence of suicidality during pregnancy and the postpartum. Arch Womens Ment Health 2005;8:77–87.

13. Mangla K, Hoffman MC, Trumpff C, O'Grady S, Monk C. Maternal self-harm deaths: An unrecognized and preventable outcome. Am J Obstet Gynecol 2019;221:295–303.

14. Moses-Kolko EL, Hipwell AE. First-onset postpartum psychiatric disorders portend high 1-year unnatural-cause mortality risk. Am J Psychiatry 2016;173:559–561.

15. Palladino CL, Singh V, Campbell J, Flynn H, Gold KJ. Homicide and suicide during the perinatal period: Findings from the national violent death reporting system. Obstet Gynecol 2011;118:1056–1063.

16. Review to Action. Report from nine maternal mortality review committees. 2018. Available at: https://reviewtoaction.org Accessed March 31, 2020.

17. American Psychiatric Association. Diagnostic and statistcal manual of mental disorders, 5th ed. Washington, DC: American Psychiatric Association, 2013.

18. National Institute of Mental Health. Postpartum depression facts. 2020. Available at: https://nimh.nih.gov/health/publications/postpartum-depression-facts/index.shtml Accessed March 31, 2020.

19. Koutra K, Chatzi L, Bagkeris M, Vassilaki M, Bitsios P, Kogevinas M. Antenatal and postnatal maternal mental health as determinants of infant neurodevelopment at 18 months of age in a mother-child cohort (Rhea Study) in Crete, Greece. Soc Psychiatry Psychiatr Epidemiol 2013;48:1335–1345.

20. Moore Simas TA, Huang M-Y, Patton C, et al. The humanistic burden of postpartum depression: A systematic literature review. Curr Med Res Opin 2019;35:383–393.

21. Netsi E, Pearson RM, Murray L, Cooper P, Craske MG, Stein A. Association of persistent and severe postnatal depression with child outcomes. JAMA Psychiatry 2018;75:247–253.

22. Surkan PJ, Ettinger AK, Hock RS, Ahmed S, Strobino DM, Minkovitz CS. Early maternal depressive symptoms and child growth trajectories: A longitudinal analysis of a nationally representative US birth cohort. BMC Pediatr 2014;14:185.

23. Meltzer-Brody S, Colquhoun H, Riesenberg R, et al. Brexanolone injection in post-partum depression: Two multicentre, double-blind, randomised, placebo-controlled, phase 3 trials. Lancet 2018;392:1058–1070.

24. Zulresso [package insert]. Cambridge, MA: Sage Therapeutics, Inc., 2019.

25. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62.

26. Hantsoo L, Ward-O'Brien D, Czarkowski KA, Gueorguieva R, Price LH, Epperson CN. A randomized, placebo-controlled, double-blind trial of sertraline for postpartum depression. Psychopharmacology 2014;231:939–948.

27. Khan A, Khan SR, Shankles EB, Polissar NL. Relative sensitivity of the Montgomery-Asberg Depression Rating Scale, the Hamilton Depression rating scale and the Clinical Global Impressions rating scale in antidepressant clinical trials. Int Clin Psychopharmacol 2002;17:281–285.

28. Kim DR, Epperson CN, Weiss AR, Wisner KL. Pharmacotherapy of postpartum depression: An update. Expert Opin Pharmacother 2014;15:1223–1234.

29. Wyrwich KW, Norquist JM, Lenderking WR, Acaster S. Methods for interpreting change over time in patient-reported outcome measures. Qual Life Res 2013;22:475–483.

30. United States Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims, 2009. Available at: https://fda.gov/media/77832/download Accessed March 31, 2020.

31. Furukawa TA, Akechi T, Azuma H, Okuyama T, Higuchi T. Evidence-based guidelines for interpretation of the Hamilton Rating Scale for Depression. J Clin Psychopharmacol 2007;27:531–534.

32. Moncrieff J, Kirsch I. Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. Contemp Clin Trials 2015;43:60–62.

33. Bender R. Calculating confidence intervals for the number needed to treat. Control Clin Trials 2001;22:102–110.

34. Citrome L. Vortioxetine for major depressive disorder: An indirect comparison with duloxetine, escitalopram, levomilnacipran, sertraline, venlafaxine, and vilazodone, using number needed to treat, number needed to harm, and likelihood to be helped or harmed. J Affect Disord 2016;196:225–233.

35. McGough JJ, Faraone SV. Estimating the size of treatment effects: Moving beyond p-values. Psychiatry (Edgmont) 2009;6:21–29.

36. Schechtman E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat—which of these should we use? Value Health 2002;5:431–436.

37. Kanes S, Colquhoun H, Gunduz-Bruce H, et al. Brexanolone (SAGE-547 injection) in post-partum depression: A randomised controlled trial. Lancet 2017;390:480–489.

38. Guy W. ECDEU assessment manual for psychopharmacology. Rockville, Maryland: U.S. Dept. of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institute of Mental Health, Psychopharmacology Research Branch, Division of Extramural Research Programs, 1976.

39. Meltzer-Brody S, Howard LM, Bergink V, et al. Postpartum psychiatric disorders. Nat Rev Dis Primers 2018;4:18022.

40. Borges S, Chen YF, Laughren TP, et al. Review of maintenance trials for major depressive disorder: A 25-year perspective from the US Food and Drug Administration. J Clin Psychiatry 2014;75:205–214.

41. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet 2018;391:1357–1366.

42. Molyneaux E, Howard LM, McGeown HR, Karia AM, Trevillion K. Antidepressant treatment for postnatal depression. Cochrane Database Syst Rev 2014:CD002018.

43. Bushnell GA, Sturmer T, Swanson SA, et al. Dosing of selective serotonin reuptake inhibitors among children and adults before and after the FDA black-box warning. Psychiatr Serv 2016;67:302–309.

44. Byatt N, Simas TA, Lundquist RS, Johnson JV, Ziedonis DM. Strategies for improving perinatal depression treatment in North American outpatient obstetric settings. J Psychosom Obstet Gynaecol 2012;33:143–161.

45. Frazer A, Benmansour S. Delayed pharmacological effects of antidepressants. Mol Psychiatry 2002;7(Suppl 1):S23–S28.

46. Jones LE, Turvey C, Carney-Doebbeling C. Inadequate follow-up care for depression and its impact on antidepressant treatment duration among veterans with and without diabetes mellitus in the Veterans Health Administration. Gen Hosp Psychiatry 2006;28:465–474.

47. Kelly K, Posternak M, Alpert JE. Toward achieving optimal response: Understanding and managing antidepressant side effects. Dialogues Clin Neurosci 2008;10:409–418.

48. Lingam R, Scott J. Treatment non-adherence in affective disorders. Acta Psychiatr Scand 2002;105:164–172.

49. Davies J, Read J. A systematic review into the incidence, severity and duration of antidepressant withdrawal effects: Are guidelines evidence-based? Addict Behav 2019;97:111–121.

50. Khan A, Fahl Mar K, Faucett J, Khan Schilling S, Brown WA. Has the rising placebo response impacted antidepressant clinical trial outcome? Data from the US Food and Drug Administration 1987–2013. World Psychiatry 2017;16:181–192.

51. Undurraga J, Baldessarini RJ. Randomized, placebo-controlled trials of antidepressants for acute major depression: Thirty-year meta-analytic review. Neuropsychopharmacology 2012;37:851–864.

52. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: Variable, substantial, and growing. JAMA 2002;287:1840–1847.

53. Cooper MC, Kilvert HS, Hodgkins P, Roskell NS, Eldar-Lissai A. Using matching-adjusted indirect comparisons and network meta-analyses to compare efficacy of brexanolone injection with selective serotonin reuptake inhibitors for treating postpartum depression. CNS Drugs 2019;33:1039–1052.

54. Eldar-Lissai A, Cohen JT, Meltzer-Brody S, et al. Cost-effectiveness of brexanolone versus selective serotonin reuptake inhibitors for the treatment of postpartum depression in the United States. J Manag Care Spec Pharm 2020;26:627–638.

Address correspondence to:
Adi Eldar-Lissai, PhD
Sage Therapeutics, Inc.
215 First Street
Cambridge, MA 02142
USA

E-mail: adi.eldar-lissai@sagerx.com