Data supplement for Deligiannidis et al., Zuranolone for the Treatment of Postpartum Depression. Am J Psychiatry (doi: 10.1176/appi.ajp.20220785)

METHODS

Inclusion Criteria

Patients were included if they signed an informed consent form, were women 18–45 years of age, in good physical health, and with no clinically significant findings, as determined by the investigator on physical examination, 12-lead electrocardiogram, or clinical laboratory tests. Patients agreed to adhere to the study requirements, including not participating in night shift work, and had ceased lactating or agreed not to provide breastmilk to their infant(s) from just prior to receiving study drug on Day 1 until 7 days following the last dose of study drug. A negative pregnancy test at Screening and Day 1 prior to the start of study drug administration was required. Patients had a diagnosis of major depressive episode that began no earlier than the third trimester and no later than the first 4 weeks following childbirth, and met criteria for major depressive episode per Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), diagnosed by Structured clinical interview for DSM-5 disorders, clinical trials version. Eligible patients had a 17-item Hamilton Depression Rating Scale (HAM-D) total score ≥26 at Screening and Day 1 (prior to randomization) and patients were ≤12 months postpartum. Any patients taking antidepressants must have been on a stable dose for at least 30 days prior to Day 1. Patients who had stopped taking antidepressants within 30 days must have stopped for longer than 5 half-lives of the antidepressant before Day 1 and were stratified into the “no antidepressant” group. Patients receiving psychotherapy must have been receiving therapy on a regular schedule for at least 30 days prior to Day 1. Patients must have been willing to delay start of other antidepressant or antianxiety medications and any new pharmacotherapy

regimens, including as-needed benzodiazepine anxiolytics and sleep aids, until after completion of the Day 45 visit. Patients were included if they agreed to use at least 1 highly effective contraceptive method during participation in the study and for 30 days following the last dose of study drug, unless surgically sterile (bilateral oophorectomy, hysterectomy, and/or bilateral salpingectomy), or did not engage in sexual relations that carry a risk of pregnancy (not including abstinence). Finally, patients were required to agree to refrain from alcohol use and misuse of drugs for the duration of the study.

Exclusion Criteria

Patients were excluded from this study if currently at significant risk of suicide, as judged by the investigator, or had attempted suicide associated with the current episode of postpartum depression (PPD). Patients were excluded if they had a recent history or active clinically significant manifestations of metabolic, hepatic, renal, hematological, pulmonary, cardiovascular, gastrointestinal, musculoskeletal, dermatological, urogenital, or neurological disorders; eyes, ears, nose, and throat disorders; or any other acute or chronic condition that, in the investigator's opinion, would limit the patient's ability to complete or participate in this clinical study. Patients with a body mass index (BMI) ≤18 or ≥45 kg/m^2 were excluded. Patients with a BMI of 40 to 44.9 kg/m^2 at Screening underwent a broader evaluation of medical comorbidities, such as chronic obstructive pulmonary disease, concomitant medications, and prior tolerability of sedating agents. Patients having a clinically significant abnormal 12-lead electrocardiogram at Screening or Day 1 (mean QT corrected according to Fridericia's formula [QTcF] of >470 msec) were excluded from the study. Any known allergy to zuranolone, allopregnanolone, or related compounds was exclusionary.

Patients with index pregnancy resulting in miscarriage, still birth, or neonatal/infant death, or the patients who had terminated parental rights (i.e., child had been placed for adoption) were excluded. Patients were excluded if they had active psychosis per investigator assessment, a medical history of nonfebrile seizures, or a medical history of bipolar disorder,

schizophrenia, and/or schizoaffective disorder. History of mild, moderate, or severe substance use disorder (including benzodiazepines) diagnosed using DSM-5 criteria in the 12 months prior to Screening was exclusionary. Patients with exposure to another investigational medication or device within 30 days prior to Screening were excluded. Any patients who previously participated in clinical trials evaluating zuranolone, brexanolone, ganaxalone, or any other compound containing allopregnanolone, or who had been previously treated with ZULRESSO™ (brexanolone) were excluded. A history of sleep apnea, gastric bypass surgery, gastric sleeve or lap band, or any procedures that interfere with gastrointestinal transit was also exclusionary. Patients who had vagus nerve stimulation, electroconvulsive therapy, or had taken ketamine with the current PPD episode were excluded. Patients were excluded if taking benzodiazepines, barbiturates, or other γ-aminobutyric acid type-A ($GABA_A$) modulators (e.g., eszopiclone, zopiclone, zaleplon, and zolpidem) at Day −28; had been using these agents daily or near daily (≥4 times per week) for more than 1 year; or were taking any benzodiazepine or GABA modulator with a half-life of ≥48 hours (e.g., diazepam) from 30 days prior to Day 1. Patients were also excluded if taking non-GABA anti-insomnia medications (e.g., melatonin, Benadryl [antihistamines], or trazodone) or first- or second-generation (typical/atypical) antipsychotics within 14 days prior to Day 1. Known use of any known strong inhibitors of cytochrome P450 (CYP)3A4 within 14 days or 5 half-lives (whichever was longer) or prior consumption of grapefruit juice, grapefruit, or Seville oranges or products containing these within 14 days prior to the first dose of study drug was exclusionary. Strong CYP3A inducers, such as rifampin, carbamazepine, enzalutamide, mitotane, phenytoin, or St John's Wort, within 14 days or 5 half-lives (whichever was longer) prior to the first dose of study drug was also exclusionary. Patients taking psychostimulants (e.g., methylphenidate, amphetamine) or opioids, regularly or as needed, within 28 days prior to Day 1 were excluded. Prior diagnosis or treatment for any type of cancer (excluding basal cell carcinoma and melanoma in situ) within the past year prior to Screening and planned elective surgeries during the study were exclusionary. Patients with a

positive urine drug test and/or alcohol screen at Screening or on Day 1 prior to dosing were also excluded. Lastly, patients who were a dependent of a sponsor, investigator, investigator's deputy, or study site staff were excluded, as were those involuntarily confined or detained in a penal institution or other institution under court order.

Additional Statistical Analyses

Multiplicity adjustment to statistical testing of hypotheses of the key secondary endpoints were conducted by using fixed sequence strategy. Only if the primary endpoint was statistically significant at 0.05 level, the key secondary endpoints were tested sequentially, testing each endpoint at 5% level of significance only if the previous endpoint in the below sequence had been significant at 5% level. If an endpoint was not significant at 5% level, the next endpoint in the sequence was to be interpreted only with nominal p-value. The sequence of testing key secondary endpoints is as follows:

- Change from baseline in HAM-D score at Day 3
- Change from baseline in HAM-D score at Day 28
- Change from baseline in HAM-D score at Day 45
- Change from baseline in Clinical Global Impressions-Severity (CGI-S) at Day 15

Mixed effects models for repeated measures were used for all change from baseline outcome measures i.e. HAM-D score, HAM-D subscale scores (Core, Anxiety, Bech-6, and Maier), HAM-D individual item scores, CGI-S score, Hamilton Rating Scale for Anxiety (HAM-A) total score, HAM-A individual item scores, Montgomery-Åsberg Depression Rating Scale (MADRS) total score, MADRS individual item scores, Edinburgh Postnatal Depression Scale (EPDS) total score, and 9-item Patient Health Questionnaire (PHQ-9) total score. Unstructured covariance structures were used to model within-patient errors. Generalized estimating equation (GEE) methods were used for the analysis of HAM-D response (defined as ≥50% reduction from baseline in HAM-D score) and HAM-D remission (defined as HAM-D score ≤7.0). GEE models used the logit link function and included terms for treatment, baseline score, antidepressant use

at baseline, assessment timepoint, and timepoint-by-treatment as explanatory variables. A GEE method was also used for the analysis of CGI-I response (missing response not accounted or was counted as no response). The comparison of interest was the difference between zuranolone and matching placebo at Day 15. Unstructured covariance structures were used to model the within-patient errors for GEE analyses. Model-based point estimates (i.e., odds ratios), 95% confidence intervals, and p-values were reported.

Kaplan-Meier methods were used to estimate the median time-to-first HAM-D response or remission. All patients who did not respond or remit to zuranolone were censored at the last available postbaseline HAM-D evaluation.

TABLE S1. Summary of baseline observed means, postbaseline LS means, and placebo-adjusted LS means of outcome measures over time (full analysis set)

	Zuranolone 50 mg n=98		Placebo n=97		LS mean difference			p-value
	Mean	SE	Mean	SE	Mean	95% CI Lower	95% CI Upper	
HAM-D score								
Baseline[a]	28.6	2.49	28.8	2.34				
Day 3[b]	−9.5	0.70	−6.1	0.71	−3.4	−5.4	−1.4	0.001
Day 8	−13.2	0.76	−9.5	0.76	−3.7	−5.8	−1.6	0.001
Day 15[c]	−15.6	0.82	−11.6	0.82	−4.0	−6.3	−1.7	0.001
Day 21	−15.7	0.87	−13.3	0.87	−2.5	−4.9	−0.1	0.045
Day 28[b]	−16.3	0.88	−13.4	0.88	−2.9	−5.4	−0.5	0.020
Day 45[b]	−17.9	0.90	−14.4	0.90	−3.5	−6.0	−1.0	0.007
CGI-S score								
Baseline[a]	5.0	0.66	4.9	0.58				
Day 3	−1.0	0.10	−0.7	0.10	−0.2	−0.5	0	0.093
Day 8	−1.7	0.12	−1.1	0.12	−0.6	−0.9	−0.3	0.0004
Day 15[b]	−2.2	0.14	−1.6	0.14	−0.6	−0.9	−0.2	0.005
Day 21	−2.3	0.14	−1.9	0.14	−0.5	−0.9	−0.1	0.022
Day 28	−2.3	0.14	−1.9	0.14	−0.4	−0.8	0	0.044
Day 45	−2.6	0.14	−2.1	0.14	−0.5	−0.9	−0.1	0.012
HAM-A total score								
Baseline[a]	24.4	6.01	24.7	5.96				
Day 3	−7.1	0.61	−5.3	0.61	−1.8	−3.5	−0.1	0.037
Day 8	−10.5	0.70	−8.5	0.70	−2.0	−4.0	0	0.045
Day 15	−12.8	0.69	−10.6	0.70	−2.2	−4.2	−0.3	0.024
Day 21	−13.6	0.74	−12.1	0.74	−1.4	−3.5	0.6	0.171
Day 28	−13.7	0.76	−12.5	0.75	−1.2	−3.3	0.9	0.248
Day 45	−14.8	0.81	−12.5	0.81	−2.3	−4.5	0	0.050
MADRS total score								

Baseline[a]	35.5	5.37	35.0	4.81				
Day 3	−16.4	1.11	−11.8	1.10	−4.6	−7.7	−1.5	0.004
Day 15	−19.7	1.20	−14.6	1.21	−5.1	−8.4	−1.7	0.003
Day 28	−20.2	1.23	−16.8	1.22	−3.4	−6.8	0	0.051
Day 45	−22.5	1.27	−17.8	1.28	−4.7	−8.3	−1.1	0.010

Results are presented as LS mean and SE CFB unless otherwise noted.

[a]Baseline values are reported as mean ± SD. [b]Key secondary endpoint. [c]Primary endpoint.

Primary and key secondary endpoints were adjusted for multiplicity. All other secondary endpoints were not adjusted for multiplicity and are to be interpreted with nominal p-values.

CFB = change from baseline; CGI-S = Clinical Global Impressions-Severity; CI = confidence interval; HAM-A = Hamilton Rating Scale for Anxiety; HAM-D = 17-item Hamilton Depression Rating Scale; LS = least squares; MADRS = Montgomery-Åsberg Depression Rating Scale; SD = standard deviation; SE = standard error.

TABLE S2. Effect size of the observed mean difference in CFB in HAM-D score between zuranolone and placebo by study visit (full analysis set)

	Mean difference[a]	Pooled SD	Cohen's d
Day 3	3.2	7.10	0.45
Day 8	3.8	7.58	0.50
Day 15	4.2	8.07	0.52
Day 21	2.4	8.21	0.29
Day 28	2.8	8.57	0.33
Day 45	2.9	8.75	0.33

[a]Reflects the difference in the observed mean CFB in HAM-D score between zuranolone and placebo.

CFB = change from baseline; d = Cohen's effect size; HAM-D = 17-item Hamilton Depression Rating Scale; SD = standard deviation.

TABLE S3. Summary of HAM-D response and remission over time (full analysis set)

	Zuranolone 50 mg n=98			Placebo n=97			Odds ratio			p-value
	n	N	%	n	N	%	Mean	95% CI Lower	95% CI Upper	
HAM-D response										
Day 3	26	98	26.5	12	96	12.5	2.64	1.24	5.61	0.012
Day 8	47	93	50.5	24	95	25.3	2.99	1.61	5.54	0.001
Day 15	53	93	57.0	35	90	38.9	2.02	1.11	3.67	0.021
Day 21	50	84	59.5	35	83	42.2	2.15	1.17	3.96	0.014
Day 28	48	77	62.3	35	85	41.2	2.39	1.29	4.43	0.006
Day 45	52	84	61.9	46	85	54.1	1.53	0.84	2.81	0.166
HAM-D remission										
Day 3	8	98	8.2	5	96	5.2	1.64	0.51	5.25	0.406
Day 8	17	93	18.3	11	95	11.6	1.68	0.74	3.81	0.218
Day 15	25	93	26.9	15	90	16.7	1.78	0.88	3.62	0.111
Day 21	25	84	29.8	17	83	20.5	1.55	0.78	3.06	0.210
Day 28	26	77	33.8	20	85	23.5	1.50	0.76	2.99	0.247
Day 45	37	84	44.0	25	85	29.4	2.08	1.11	3.92	0.023

HAM-D response was defined as a ≥50% reduction from baseline in HAM-D score. HAM-D remission was defined as a HAM-D score ≤7.

HAM-D response and remission data were not adjusted for multiplicity and are to be interpreted with nominal p-values.

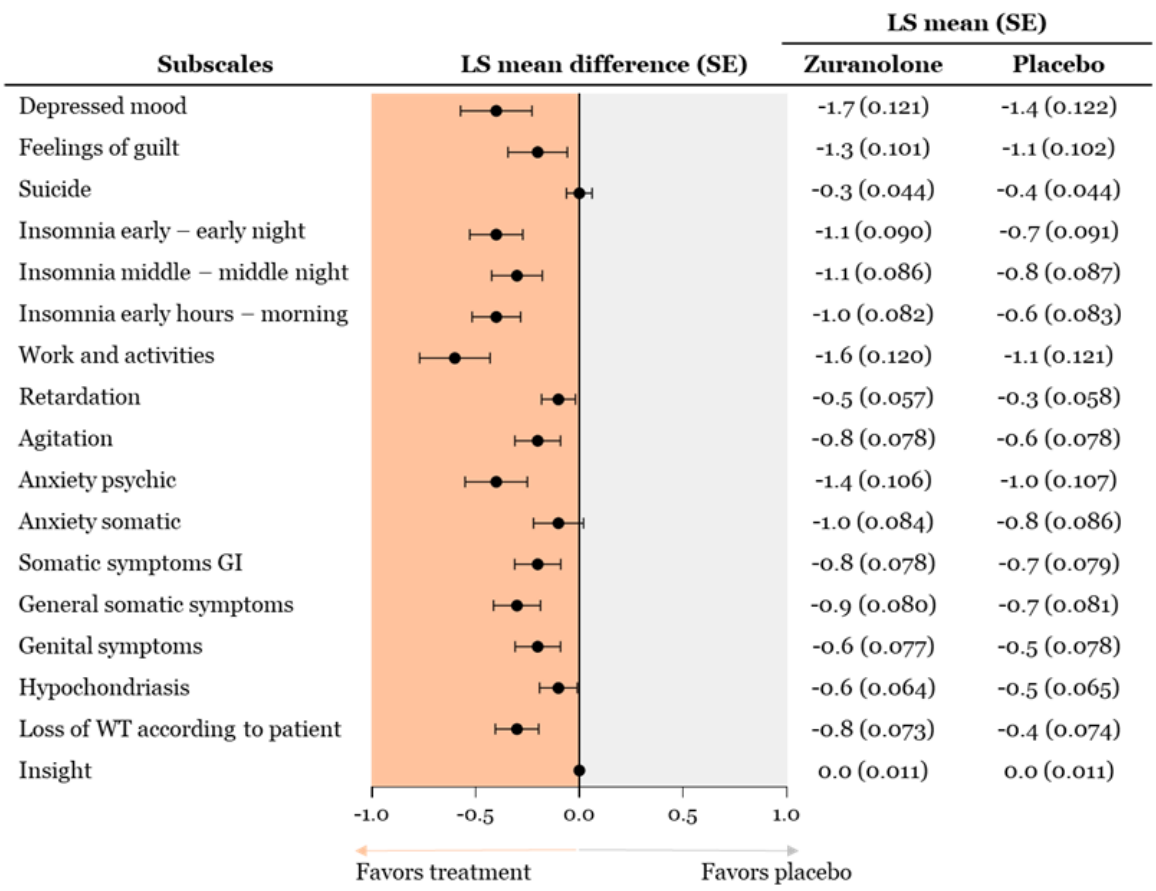
CI = confidence interval; HAM-D = 17-item Hamilton Depression Rating Scale.

TABLE S4. Summary of CFB in HAM-D score over time by ADT use at baseline (full analysis set)

	Zuranolone 50 mg				Placebo			
	ADT=Yes n=15		ADT=No n=83		ADT=Yes n=15		ADT=No n=82	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
HAM-D score, CFB								
Day 3	−8.2	6.88	−9.8	7.50	−7.4	6.02	−6.1	6.92
Day 8	−14.1	8.67	−13.2	7.37	−10.5	7.67	−9.4	7.64
Day 15	−13.5	8.34	−16.0	7.47	−15.2	8.46	−10.6	8.36
Day 21	−13.6	8.71	−15.9	7.58	−15.2	9.20	−12.7	8.52
Day 28	−14.9	10.73	−16.7	7.77	−15.3	9.29	−13.2	8.70
Day 45	−15.8	10.05	−18.1	8.03	−15.2	9.37	−14.7	9.10

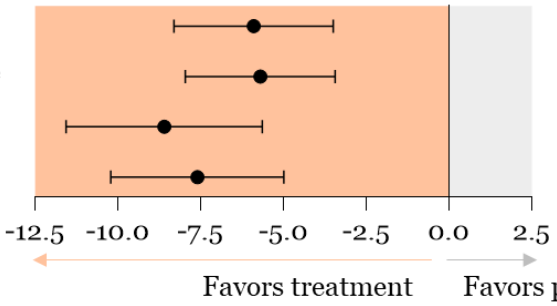
ADT = antidepressant therapy; CFB = change from baseline; HAM-D= 17-item Hamilton Depression Rating Scale; SD = standard deviation.

FIGURE S1. Change from baseline in HAM-D individual items at Day 15 (full analysis set, n=183)



CI = confidence interval; GI = gastrointestinal; HAM-D = 17-item Hamilton Rating Scale for Depression; LS = least squares; SE = standard error; WT = weight.

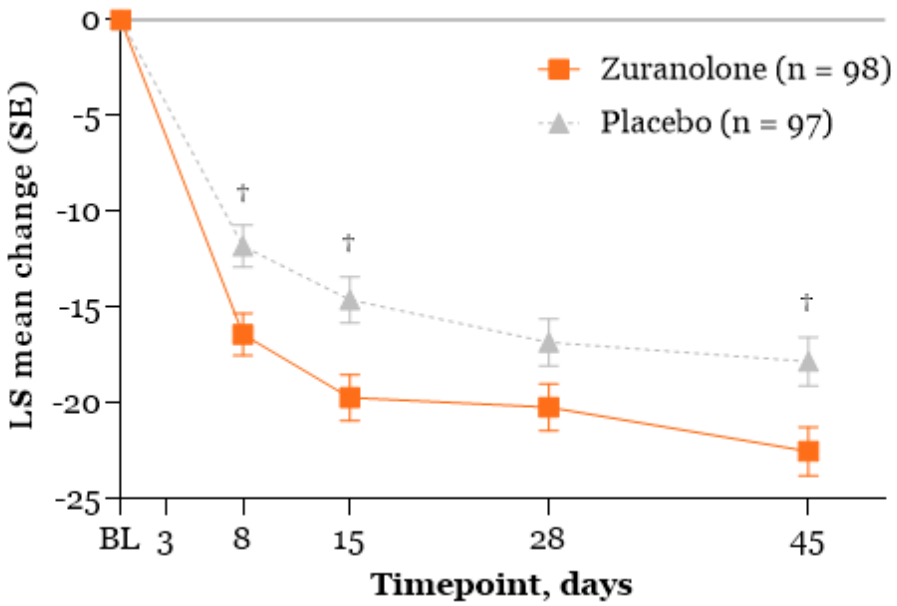
FIGURE S2. Change from baseline in HAM-D subscales at Day 15 (full analysis set)



Subscales	LS mean difference (SE)	LS mean (SE)	
		Zuranolone	Placebo
Core subscale score		−27.2 (1.689)	−21.3 (1.702)
Anxiety subscale score		−26.2 (1.594)	−20.5 (1.607)
Bech-6 subscale score		−34.1 (2.085)	−25.5 (2.100)
Maier subscale score		−30.7 (1.838)	−23.1 (1.852)

HAM-D = 17-item Hamilton Rating Scale for Depression; LS = least squares; SE = standard error.

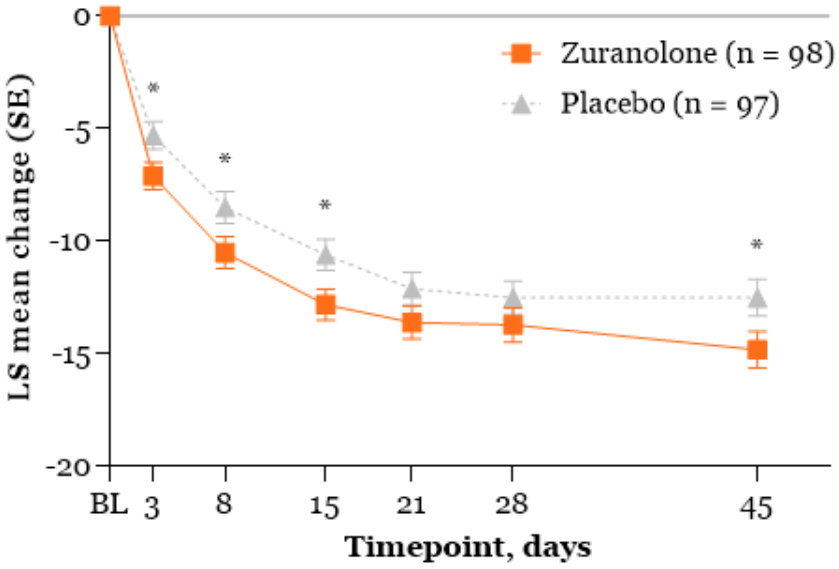
FIGURE S3. Change from baseline in MADRS total score (full analysis set)



p-values designated as: †<0.01. Secondary endpoints were not adjusted for multiplicity and are to be interpreted with nominal p-values.

BL = baseline; LS = least squares; MADRS = Montgomery-Åsberg Depression Rating Scale; SE = standard error.

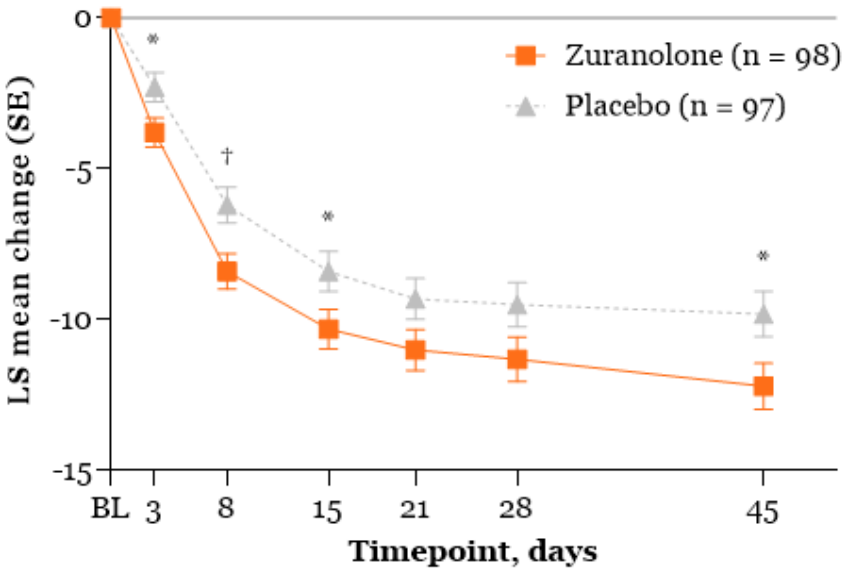
FIGURE S4. Change from baseline in HAM-A total score (full analysis set)



p-values designated as: *<0.05. Secondary endpoints were not adjusted for multiplicity and are to be interpreted with nominal p-values.

BL = baseline; HAM-A =Hamilton Rating Scale for Anxiety total score; LS = least squares; SE = standard error.

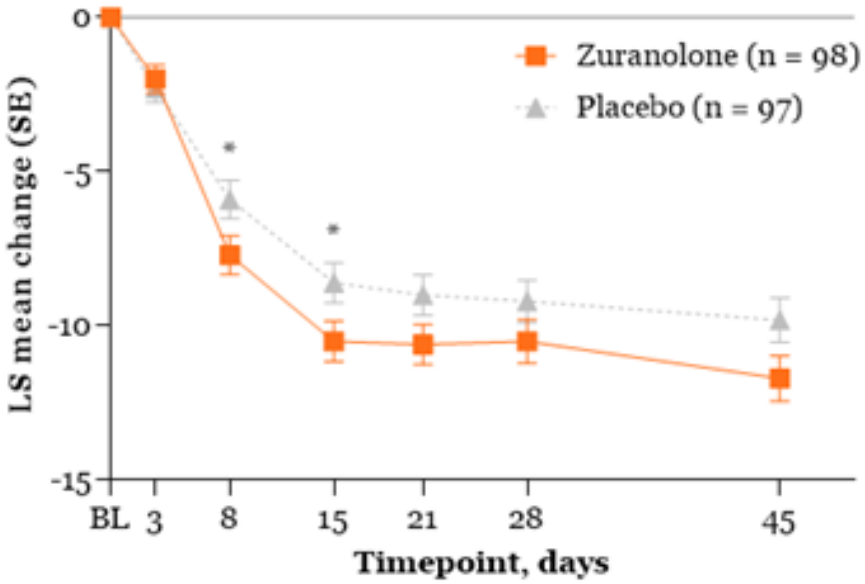
FIGURE S5. Change from baseline in patient-reported measure of depressive symptoms as assessed by EPDS total score



p-values designated as: *<0.05, †<0.01. Secondary endpoints were not adjusted for multiplicity and are to be interpreted with nominal p-values.

BL = baseline; EPDS = Edinburgh Postnatal Depression Scale; LS = least squares; SE = standard error.

FIGURE S6. Change from baseline in patient-reported measure of depressive symptoms as assessed by PHQ-9 total score



p-values designated as: *<0.05. Secondary endpoints were not adjusted for multiplicity and are to be interpreted with nominal p-values.

BL = baseline; LS = least squares; PHQ-9 = 9-item Patient Health Questionnaire total score; SE = standard error.