

ScholarWorks@GSU

A Systematic Review and Meta-Analysis of the Effectiveness of Brexanolone as a Treatment for Postpartum Depression

Authors	Beharry, Sarah
Citation	Beharry, Sarah. "A Systematic Review and Meta-Analysis of the Effectiveness of Brexanolone as a Treatment for Postpartum Depression." Thesis, Georgia State University, 2024. https://doi.org/10.57709/36390255
DOI	https://doi.org/10.57709/36390255
Download date	2026-03-09 03:47:37
Link to Item	https://hdl.handle.net/20.500.14694/9860

ABSTRACT

A SYSTEMATIC REVIEW AND META-ANALYSIS OF THE EFFECTIVENESS OF BREXANOLONE AS A TREATMENT FOR POSTPARTUM DEPRESSION

by

SARAH MARIAM BEHARRY

November 17, 2023

Postpartum depression (PPD) is a debilitating disorder that affects the functioning of new mothers and their ability to care for their infants. The onset of postpartum depression can be attributed to genetic and social factors as well as fluctuations in hormones during and after pregnancy. Brexanolone is a new medication developed to treat moderate to severe PPD. The present study aimed to complete a systematic review and meta-analysis of the available literature to support the claim that brexanolone is an effective medication to treat PPD. Additional aims include determining the demographic features of participants and the generalizability of the results. The *Lancet*, *PubMed*, *PsycINFO*, *Embase*, and *Science Direct* were searched using a Boolean search string to find primary studies and trials that were evaluating the efficacy of brexanolone with an outcome of changes in depression scores measured using the HAM-D among adults within 12 months postpartum. The collected studies were screened for eligibility and coded for study characteristics. The statistical analysis used a fixed effects model due to the small number of studies to estimate the mean effect size (standardized mean difference) using the *metafor* and *tidyverse* packages in R. The search yielded 233 articles, of which five were eligible. The statistical analyses yielded an SMD of 3.52 (95% CI: [3.16, 3.88]) for randomized controlled trials and an SMD accounting for pre-post of 2.97 (95% CI: [2.21, 3.73]) for single-group studies comparing outcome results to baseline data. There was significant heterogeneity among both the randomized controlled trial studies and the pre-post studies in their effect size estimates. This was indicated by the results of the Q-test for heterogeneity, which yielded results of $Q = 12.8646$ ($p = 0.0049$) and $Q = 11.4964$ ($p = 0.0032$) for the randomized controlled trials and pre-post studies, respectively. The Egger's regression test yielded a result of $z = -8.4661$ (p -value: <0.0001), indicating potential publication bias. The results were statistically significant, indicating that brexanolone is an effective treatment for women who suffer from moderate to severe PPD. Further studies investigating the efficacy of brexanolone should be conducted with larger and more diverse sample sizes.

A SYSTEMATIC REVIEW AND META-ANALYSIS OF THE EFFECTIVENESS OF
BREXANOLONE AS A TREATMENT FOR POSTPARTUM DEPRESSION

by

SARAH MARIAM BEHARRY

B.S., GEORGIA STATE UNIVERSITY

A Thesis Submitted to the Graduate Faculty
of Georgia State University in Partial Fulfillment
of the
Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA

30303

APPROVAL PAGE

A SYSTEMATIC REVIEW AND META-ANALYSIS OF THE EFFECTIVENESS OF
BREXANOLONE AS A TREATMENT FOR POSTPARTUM DEPRESSION

by

SARAH MARIAM BEHARRY

Approved:

Dr. Therese Pigott

Committee Chair

Dr. Alexander Kirpich

Committee Member

November 17, 2023

Date

Acknowledgments

I want to begin by thanking Dr. Terri Pigott for her continuous support and advisement throughout this challenging journey. I was first introduced to the ideas of systematic reviews and meta-analyses from the PH8890 – Special Topics in Biostatistics: Systematic Review and Meta-Analysis course taught by her. Her resources from that course helped to significantly guide the path of the current systematic review. I cannot thank Dr. Pigott enough for her guidance and reassurance that I could complete this thesis. She is a wonderful mentor and was there with me every step of the way. I'm so grateful to have had her as my committee chair.

Not only do I thank Dr. Alexander Kirpich for agreeing to serve as my thesis committee member, but I also thank him for everything he has taught me in the courses I've taken with him. I was introduced to the programming language R in his PHPH7525 – Statistical Computing course, which helped me to familiarize myself with the program before using it to complete the statistical analyses in my systematic review. I thank Dr. Kirpich for his support and guidance throughout this process.

My attempt to search the literature could not have been done without the help of the School of Public Health librarian, Bethany Havas. Bethany helped me to figure out how to begin undertaking the task of developing search string terms and what databases to search. She also taught me about helpful tools and methods to use when searching the literature that I will continue to utilize in my education and career.

I would also like to thank Dr. Frank Castelli for helping to motivate me in my times of weakness and regularly checking in on my progress throughout the course of completing my thesis. He encouraged me to keep going and not fear what comes next.

Finally, I truly could not have completed this thesis without the love, support, and encouragement from my best friend, M'Ayesha, my sisters, Aaliyah and Hannah, my parents, Elaine and Dereck, and most importantly, both of my grandfathers, Rajah and Winston. My friends and family continued to remind me that I could do this when I was struggling the most. They were always there for me, telling me how proud they were of me and reminding me of the confidence they had in me.

Author's Statement Page

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this thesis may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission of the author.

Sarah Mariam Beharry

Signature of Author

Table of Contents

Title Page	i
Approval Page	ii
Acknowledgements	iii
Author’s Statement Page	iv
Table of Contents	v
List of Tables	vi
List of Figures	vii
Chapter I: Introduction	1
1.1. Postpartum Depression	1
1.2. Brexanolone.....	5
Chapter II: Review of the Literature	7
Chapter III: Methods & Procedures	9
3.1. Eligibility Criteria	9
3.2. Search Strategy	10
3.3. Title & Abstract Screening.....	10
3.4. Full Text Review & Data Extraction	11
3.5. Statistical Analysis.....	12
3.6. Publication Bias	13
Chapter IV: Results	15
4.1. Literature Search.....	15
4.2. Study Characteristics	15
4.3. Meta-Analysis.....	19
4.4. Publication Bias	20
Chapter V: Discussion & Conclusion	22
5.1. Interpretation of Results	22
5.2. Limitations.....	23
5.3. Future Research	24
5.4. Conclusion.....	24
References	26
Appendices	35
Appendix A – Screening Tool.....	35
Appendix B – Data Extraction Codebook	37
Appendix C – RStudio Data Analysis	39

List of Tables

Table 1: Characteristics of Included Studies	17
Table 2: Patient Demographics of Included Studies	18

List of Figures

Figure 1: PRISMA Flow Diagram.....	16
Figure 2: Forest Plot of Standardized Mean Differences for Randomized Controlled Trials.....	19
Figure 3: Forest Plot of Standardized Mean Changes for Pre-Post Studies.....	20
Figure 4: Contoured Funnel Plot of Standardized Mean Difference	21

Chapter I – Introduction

1.1. Postpartum Depression

Childbirth is a strenuous and exhausting process that millions of women endure at some point in their lifetime. During and after pregnancy, a female will experience a wide range of changes in her body, including physical, hormonal, and psychological shifts (Mughal et al., 2022). Many complications that lead to maternal morbidity often develop during the postpartum period (Zainur & Loh, 2006), which begins after the birth of the infant and will last approximately eight weeks, ultimately ending once a female’s body has returned to its pre-pregnancy state (Lopez-Gonzalez & Kopparapu, 2022). Examples of maternal morbidities include hemorrhage, hypertension, pulmonary embolism, sepsis, incontinence, and, most commonly, non-psychotic mental disorders such as depression (Zainur & Loh, 2006).

The first scientific study observing depression during the postpartum period was conducted by Robin (1962). However, reports of emotional difficulties after giving birth were described as early as 700 BC by Hippocrates and were initially termed “the baby blues” (Kroh & Lim, 2021). As more research was conducted surrounding these feelings of sadness and anxiety during and after pregnancy, scientists more appropriately named this phenomenon “postpartum depression,” leaving the terms of “baby blues” for feelings of depression and anxiety that develop within the first few days after giving birth and last for only a few weeks with no need for medical treatment (Rich, n.d.).

Postpartum depression also referred to as postnatal or perinatal depression (National Institute of Mental Health, 2020), is a major depressive episode that typically develops within four weeks up to one year after a female gives birth (Mughal et al., 2022). Women with postpartum depression experience symptoms that include extreme sadness, anxiety, irritability,

feelings of guilt, worthlessness, or hopelessness, difficulty concentrating or making decisions, abnormal appetite, loss of interest in daily activities, doubts about their ability to care for their infant, thoughts of harming their infant, insomnia or hypersomnia, and suicidal ideation or attempt (Mughal et al., 2022; National Institute of Mental Health, 2020). When left untreated, postpartum depression can result in negative consequences for both the mother and their children. It has been observed in children up to five years of age of mothers who suffer from postpartum depression that they perform significantly less well on cognitive tasks, are more likely to be rated as behaviorally disturbed by teachers, have increased distractibility, experience antisocial or neurotic behavior, experience slower development in motor skills, are quicker to cry in response to stimuli, and cry louder and longer (Grace et al., 2003; Slomian et al., 2019). Mothers who experience postpartum depression often have adverse physical and psychological health, worse quality of life, difficulties in social relationships, engage in more risky behaviors, such as smoking, exhibit poor maternal care, experience bonding difficulties with their infant, and experience issues with breastfeeding their infant (Slomian et al., 2019).

It has proven difficult to assess the global prevalence of postpartum depression due to the lack of uniformity among the methodological parameters used to diagnose postpartum depression (Moraes et al., 2017). The global prevalence of postpartum depression is estimated to be anywhere between 6.5% and 12.9%, with estimates being even higher in low- and middle-income countries (Stewart & Vigod, 2016). Despite the variability amongst how and when postpartum depression is assessed, the most common method of identifying the presence of postpartum women is to screen for depression using a self-report depression symptom questionnaire within 12 months of the female giving birth (Levis et al., 2020). Recommended screening tests that have been validated to identify depression during the postpartum period

include the Edinburgh Postnatal Depression Scale (EPDS), the Postpartum Depression Screening Scale (PDSS), the Patient Health Questionnaire-9 (PHQ-9), the Beck Depression Inventory (BDI), and the Hamilton Rating Scale for Depression (HAM-D) (American College of Obstetricians, 2015). The HAM-D is the most common clinician-rated scale used to assess the severity of depression among patients (Carrozzino et al., 2020) and is considered to be the gold standard criterion to evaluate the efficacy of treatments for depression in clinical trials (Gerbasi et al., 2020). Although the EPDS is the most common depression self-reporting tool used to screen for postpartum depression around the world due to significant levels of sensitivity and specificity (Moraes et al., 2017), research has shown that clinician ratings from the HAM-D align with self-reports from the EPDS and PHQ-9 making the HAM-D an appropriate and accurate tool used to identify the presence of postpartum depression in clinical settings (Gerbasi et al., 2020). The scores indicating the severity of depression using the HAM-D is as follows: a score of 0-7 indicates no depression, 8-16 indicates mild depression, 17-23 indicates moderate depression, and a score of 24 or higher indicates severe depression (Zimmerman et al., 2013).

The specific pathogenesis of postpartum depression remains unknown. The onset of postpartum depression can be attributed to genetic factors, including familial history of mental illness. Social factors such as marital difficulties, intimate partner violence, low social support, and adverse life events are often significant contributing factors in the development of postpartum depression (Stewart & Vigod, 2016). In addition to these factors, research has long shown strong evidence supporting the involvement of the fluctuating levels of reproductive hormones, such as estrogen and progesterone, during and after pregnancy to lead to the onset of postpartum depression (Bloch et al., 2000).

One of the most recent hormones that has been evaluated for its influence on developing postpartum depression is allopregnanolone. This endogenous neuroactive steroid is a positive allosteric modulator of gamma-aminobutyric acid (GABA) receptors (Meltzer-Brody & Kaner, 2020). GABA receptors are inhibitory receptors that respond when GABA is released into the post-synaptic nerve terminal (Allen et al., 2023). Positive allosteric modulators increase the efficiency with which chloride channels open in neural communication when a chemical binds to a GABA receptor. This essentially results in an inhibition of a possible action potential. Therefore, positive allosteric modulators are used to manage medical conditions such as seizures, anxiety, and, more recently, postpartum depression (Edwards & Preuss, 2023). Allopregnanolone levels fluctuate throughout pregnancy, increasing during pregnancy and significantly decreasing after childbirth (Luisi et al., 2000; Nappi et al., 2001). Reduced levels of allopregnanolone in blood serum are associated with an increased risk of developing depression (Schüle et al., 2014). The GABA receptors fail to effectively adapt to the sudden changes in the levels of allopregnanolone, which may be linked to the development of postpartum depression (Maguire & Mody, 2008).

The appropriate method to treat postpartum depression depends on the severity of a mother's symptoms, her functional status, and her ability to care for her infant. Interventions effectively targeting women who experience mild symptoms of postpartum depression include psychosocial interventions such as peer support groups and nondirective counseling (Stewart & Vigod, 2016). Women who experience moderate symptoms of postpartum depression should be provided with psychological treatments such as cognitive behavioral therapy (CBT) and interpersonal therapy (IPT) (Stewart & Vigod, 2019). When the onset of postpartum depression is severe and psychosocial and psychological treatments are ineffective, it may then be required

to turn to pharmacological treatments (Stewart & Vigod, 2019). Specific pharmacological treatments, including antidepressants such as selective serotonin reuptake inhibitors (SSRIs), are validated as effective treatments to treat postpartum depression. However, they are often used as a last resort due to concerns about breastmilk transmission and potential negative side effects (Dennis & Hodnett, 2007). There have been recent pharmacological treatments, such as brexanolone, an aqueous formulation of allopregnanolone (Azhar & Din, 2022), that have been developed specifically to treat postpartum depression with the aim of targeting hormone fluctuations that may trigger the development of postpartum depression.

1.2. Brexanolone

Brexanolone, also known as Zulresso or SAGE-547, is the first pharmacological treatment approved by the Federal Drug and Administration in March 2019 to specifically treat moderate to severe postpartum depression in adults (Azhar & Din, 2022; Powell et al., 2019). The precise mechanism of action of brexanolone has yet to be entirely understood (Azhar & Din, 2022). Since brexanolone is a neuroactive steroid and is analogous to allopregnanolone, it is thought that brexanolone acts as a positive allosteric modulator for the GABA receptors (Powell et al., 2019).

Administration of brexanolone is done intravenously by a healthcare provider at a certified healthcare facility and is currently only available through the Risk Evaluation and Mitigation Strategy Program. The brexanolone infusion is administered over a 60-hour period with a dose starting with 30 mcg/kg/hour, increasing to 90 mcg/kg/hour, and then decreasing back to 30 mcg/kg/hour (Azhar & Din, 2022). Clinical trials have shown the potential of brexanolone as an effective treatment for moderate to severe postpartum depression. The present

study aims to complete a systematic review and meta-analysis of the available literature to support further the claim of the effectiveness of brexanolone by conducting statistical analyses to observe the significance of changes in depression scores using the HAM-D before and after the brexanolone treatment. Secondary aims include determining the demographic features of participants included in studies involving brexanolone and the generalizability of results from the current literature.

Chapter II – Review of the Literature

As noted above, postpartum depression is a significant public health issue that deserves meaningful recognition and research investments in order to improve the health outcomes of new mothers, their infants, and their families. Psychosocial and psychological treatments have proven to be effective at treating mild postpartum depression, but when the need for pharmacological treatments arises, the options are highly limited. Brexanolone may be a potentially effective treatment targeted towards moderate to severe postpartum depression, however, further research and analysis of the current literature needs to be done.

Since the development and approval of brexanolone to be used as a treatment for postpartum depression in 2019, there has only been one systematic review conducted to examine the effectiveness, tolerability, and safety of the brexanolone infusion. Zheng et al. (2019) systematically searched Chinese and English databases to identify studies that conducted randomized controlled trials involving patients diagnosed with postpartum depression. The primary outcome they were observing was HAM-D total scores and remission. Secondary outcomes included changes in depressive symptoms, discontinuation of the study by participants, and adverse reactions to the brexanolone infusion treatment. Their search yielded two articles (Meltzer-Brody et al., 2018; Kaner et al., 2017). In the statistical analysis, Zheng et al. (2019) analyzed treatment response and remission, discontinuation rates, and adverse reaction rates using a random effects model. They failed to analyze the changes in depressive symptoms. Ultimately Zheng et al. (2019) concluded that the response to brexanolone could last up to a week and is considered safe from their statistical analysis.

The previously conducted systematic review provided immense knowledge regarding the safety and remission rates of the brexanolone treatment, however, they were unable to address

the efficacy of brexanolone in terms of reducing depression scores. The current systematic review aims to fill in the gap left by Zheng et al. (2019) by including studies released after the systematic review and single-group pre-post design studies in addition to the randomized controlled trials. Although single-group pre-post design studies are considered weak, it is advantageous for this review since the changes in depression scores before and after treatment are measured in the same participant (Privitera & Ahlgrim-Delzell, 2019). A meta-analysis will then be conducted to determine the statistical significance of changes in depression scores. Despite the lack of evidence supporting the efficacy from Zheng et al (2019), their review provides a helpful guideline to develop the inclusion and exclusion criteria for the proposed systematic review.

Another major factor of studies involving postpartum depression that is often overlooked is the demographic distribution of individuals taking part in the studies. The lack of diversity among racial and ethnic groups of study participants can greatly limit the generalizability of results to a broader population (Shea et al., 2022). The systematic review aims to identify the demographic distribution of included studies and determine how this component can affect how the results can be adapted to other populations.

Chapter III – Methods & Procedures

3.1. Eligibility Criteria

The PICOS acronym was used to outline the inclusion criteria. The acronym stands for the following: population/participants/problem, intervention, comparisons, outcomes, and study designs. The framework was developed to help create guidelines for inclusion and exclusion criteria for systematic reviews (Richardson et al., 1995). The PICO acronym used is as follows:

- Population: mothers <1 year postpartum
- Intervention: brexanolone intravenous treatment
- Comparison: placebo or baseline data
- Outcome: reduced depression scores measured using HAMD
- Study design: experimental studies, either two-group comparison or single-group pre-post designs.

Studies were eligible for the proposed review if they meet the following requirements: the study is a primary experimental study, participants are mothers within 12 months postpartum, participants are diagnosed with PPD by a medical professional, the researchers are analyzing PPD, the researchers of the study are measuring changes in PPD using HAM-D, and there is a comparison to a placebo group or baseline data. Randomized controlled trials and single-group pre-post designs were eligible for the review. Studies were also eligible if the researchers observed other mental illnesses or other biological interactions in addition to depression. Studies that used the longer HAM-D questionnaire and tested other pharmacological interventions along with brexanolone were also deemed eligible. The described eligibility criteria ensured that studies were included that allow for a systematic analysis of the effect of brexanolone on women suffering from PPD, using specifically the HAM-D questionnaire to measure the outcome.

3.2. Search Strategy

In order to collect data, a thorough search of the available literature was conducted to find eligible studies. Databases including *The Lancet*, *PubMed*, *PsycINFO*, *Embase*, and *Science Direct* were searched using the following Boolean search string: “(brexanolone OR Zulresso OR SAGE-547) AND (postpartum depression OR postnatal depression OR perinatal depression)” and filtering for articles, research articles, and clinical trials. These terms were chosen for the comprehensive search to find studies in which researchers analyzed the use of brexanolone as a treatment for postpartum depression. There are multiple variations to reference postpartum depression and brexanolone; therefore, the interchangeable terms were included in the search string. The filters were applied to exclude publications that did not yield primary data results. Using multiple databases ensured a proper search of all available articles, generating many potential studies for the systematic review and meta-analysis. The strategy of searching numerous databases using Boolean search strings is best for the purposes of the proposed review since the approach reduces biases and random errors while synthesizing a large quantity of specific primary studies (Gopalakrishnan & Ganeshkumar, 2013). Development of the Boolean search string was done with guidance from the Georgia State University librarian, Bethany Havas. The literature search was completed solely by the author.

3.3. Title & Abstract Screening

The titles and abstracts of studies identified from searching the literature were imported into Rayyan, an online tool used to detect duplicate articles imported from multiple sources and screen potentially eligible articles (Ouzzani et al., 2016). Once Rayyan detected the duplicated articles, they were either resolved or deleted. In addition to the highlighted “keywords for

include” added by Rayyan, terms including “brexanolone” and “depression” were added to aid the screening process. Articles were excluded for being unrelated, testing biological interactions of brexanolone, evaluating a different drug, being the wrong publication type, having a wrong study design, evaluating the wrong population, or being a background article. Articles were included if they were observing the effects of brexanolone, evaluating postpartum depression and if the HAM-D was used to assess changes in postpartum depression. If there was any skepticism about an article and its contents, the article was marked as maybe. The level one title and abstract screening tool can be found in Appendix A. Articles marked as “include” or “maybe” moved on to the full-text review process. Only one screening was done by the author.

3.4. Full-Text Review & Data Extraction

The full text of articles was collected by going back to the database source of each respective article. Each file was imported into Rayyan and added to the respective article. The full-text review was conducted to ultimately confirm the eligibility of the articles to be included in the systematic review. Articles were included if they reported sample statistics, observed a population of women aged 18-45, the population was within 12 months postpartum, and the study included a comparison group. Articles initially marked as “maybe” were excluded after the full-text review because they were pooled analysis studies. Pooled analysis studies are not favorable to include in the meta-analysis since they do not include primary data and simply provide a summary of the data by combining the data without it being weighted (Bravata & Olkin, 2001). The level two eligibility decision from the full-text screening tool can be found in Appendix A. The full-text review and eligibility decision were completed by the author.

The eligible studies were then assigned a study ID number and coded for characteristics including what the treatment outcome data was compared to, data collection intervals, the location of the study, the sample size of the studies, the year the study was conducted, if there were adverse effects reported, if statistics were reported, the dosing schedule of brexanolone, and the demographics of the sample. The studies are coded with these characteristics to organize them and further analyze other similarities between the studies. Coding the studies also helps to summarize the information and highlight aspects of the study design that were considered during the screening process (Levett, 2023). The characteristics retrieved from the studies can be found in Table 1.

3.5. Statistical Analysis

The studies were grouped into randomized controlled trials that had a placebo comparison group and those that compared depression outcome HAM-D scores to baseline HAM-D scores. In order to determine the magnitude of the relationship between the effect of brexanolone on postpartum depression, effect sizes were calculated appropriate to the study design (Littell et al., 2022). The standardized mean difference between the placebo group and the treatment group was calculated for the randomized controlled trials. When using patient-reported outcome measures such as the Hamilton Rating Scale for Depression, which is used in this case, the standardized mean difference is preferred because the results of the analysis are generalizable (Takeshima et al., 2014). For studies that had multiple arms comparing different doses of brexanolone to the same placebo group, the effect sizes were calculated separately using the same data from the placebo group for each effect size, inducing dependency among effect sizes for these studies since they used the same placebo group.

Since the remaining studies were not randomized controlled trials and instead compared changes in postpartum depression scores from the treatment to initial depression scores at baseline, the effect sizes were calculated in terms of standardized mean change using raw score standardization with heteroscedastic population variances. This effect size was chosen because the studies observed and reported raw changes in the depression score on two measurement occasions (Viechtbauer, n.d.).

The outcomes for both groups of studies were analyzed using the fixed effects model. A fixed effects model is used for the meta-analysis because of the limited number of studies included. This model is appropriate because the analysis estimates a single underlying effect, and it is assumed that there is one common effect of brexanolone on postpartum depression (Tufanaru et al., 2015). The meta-analysis was completed using RStudio 2023.09.0 and the tidyverse and metafor packages (Viechtbauer, 2010). The coding for the analysis using RStudio can be found in Appendix C.

3.6. Publication Bias

Due to the limited search strategy, there is likely publication bias among the selected studies. Publication bias arises from researchers not publishing their findings due to the belief that their study does not contribute to the necessary literature due to the statistical insignificance of their analyses. Because of this, effect sizes may be overrepresented and exaggerated in the published literature (Littell et al., 2022). To analyze the presence of publication bias, a funnel plot was generated, and Egger's test was conducted using RStudio (Egger et al., 1997). The funnel plot includes the calculated standardized mean differences of all the included studies. The studies were not separated based on study design in order to effectively analyze the publication

bias among all included studies. Finally, the statistical analyses of effect sizes and Egger's Regression were completed solely by the author.

Chapter IV – Results

4.1. Literature Search

The search from five databases yielded 233 articles. After removing duplicates and screening the articles, five studies were included in the meta-analysis. The PRISMA Flow Diagram in Figure 1 depicts the flow of information throughout the stages of the systematic review and highlights the number of studies identified, included, and excluded (Page et al., 2021). Articles assessed for eligibility were excluded for not having the appropriate study design or not being the correct publication type. The specific improper publication type was a conference proceeding which did not include primary data results.

4.2. Study Characteristics

Relevant information extracted from the selected studies can be found in Table 1. The study conducted by Meltzer-Brody et al. (2018) comprised of two studies with three active arms. The first study compared two arms to one placebo, and the second study compared the third arm to a separate placebo (Meltzer-Brody et al., 2018). Of the five studies, two of them were randomized controlled trials (Kanes et al., 2017; Meltzer-Brody et al., 2018), one was a proof-of-concept study (Kanes et al., 2016), one was a clinical treatment program (Patterson et al., 2022), and the final study was an experimental study observing the inflammatory response of brexanolone as a primary outcome and changes in postpartum depression score as a secondary outcome (Balan et al., 2023). Participants who had severe depression and were six months or less postpartum were considered eligible for the selected studies. All of the studies tested the participants for postpartum depression using the Hamilton Rating Scale for Depression before and 60 hours after the brexanolone infusion. There were differences in the dosing schedule for brexanolone between the studies; however, the researchers of each study administered a solution

of 5 mg/ml allopregnanolone in 250 mg/ml sulfobutylether- β -cyclodextrin to their participants. Three of the five studies reported adverse effects from the brexanolone infusion, and all reported statistics that were extracted and used for the meta-analysis.

Table 2 provides information regarding the demographic data collected by the selected studies. The study conducted by Kanés et al. (2016) did not report any demographic data on their patients. This may be due to the study including only four individuals. All of the studies included adult women within the reproductive ages of 18 - 45. The studies also included many patients with diverse racial and ethnic backgrounds. Women were categorized by ethnicity as being Hispanic/Latino or not. The women were also categorized by race in terms of being African American, White, or belonging to another racial group. All of the studies were conducted in the United States.

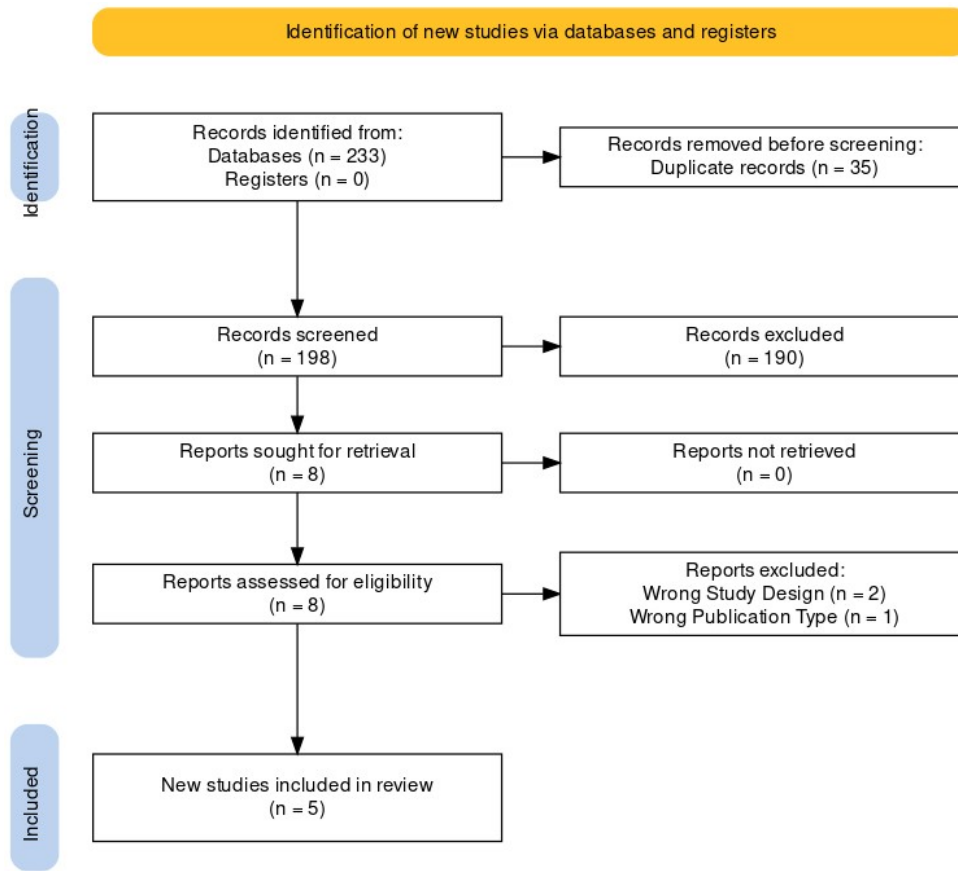


Figure 1. PRISMA Flow Diagram

Table 1: Characteristics of Included Studies

Study ID	Study	Study Year	Study Location	Study Design	Comparison Group	Time of Data Collection	Brexanolone Dosing	Adverse Effects Reported	Control Sample Size	Experimental Sample Size	Total Sample Size	Effect Size	Standard Error
1	Kanes et al. (2017)	2015-2016	US	RCT	placebo	60 hr post treatment	30 µg/kg per h (0–4 h); 60 µg/kg per h (4–24 h); 90 µg/kg per h (24–52 h); 60 µg/kg per h (52–56 h); 30 µg/kg per h (56–60 h)	dizziness, somnolence, sedation, sinus tachycardia, insomnia, infusion site pain, tension headache	11	10	21	4.112	0.594
2	*Meltzer-Brody et al. (A.1)	2016-2017	US	RCT	placebo	60 hr post treatment	60 µg/kg per h	headache, dizziness, somnolence, social ideation, intentional overdose attempt, loss of consciousness	46	47	93	4.736	0.164
3	*Meltzer-Brody et al. (A.2)	2016-2018	US	RCT	placebo	60 hr post treatment	90 µg/kg per h	headache, dizziness, somnolence	46	45	91	3.189	0.1
4	Meltzer-Brody et al. (B)	2016-2018	US	RCT	placebo	60 hr post treatment	60 µg/kg per h	headache, dizziness, somnolence, altered state of consciousness, syncope, fatigue, presyncope	54	54	108	3.103	0.082
5	Kanes et al. (2016)	2015	US	open-label, proof of concept	baseline	12, 24, 36, 48, 60, 84 hr post treatment	21.5 µg/kg per h (0–4 h); 43 µg/kg per h (5–8 h); 64.5 µg/kg per h (9–12 h); 86 µg/kg per h (13–48 h); 64.5 µg/kg per h (49–52 h); 43 µg/kg per h (53–56 h); 21.5 µg/kg per h (57–60 h)	sedation; infusion site discomfort, erythema, pain, and rash; TSH increase; dizziness; flushing; oropharyngeal pain			4	4.359	1.767
6	Patterson et al.	2019-2020	US	clinical treatment program	baseline	60 hr post treatment	unclear	none reported			16	5.95	1.104
7	Balan et al.		US	experimental study	baseline	60 hr post treatment	30 µg/kg per h (0–4 h); 60 µg/kg per h (4–24 h); 90 µg/kg per h (24–52 h); 60 µg/kg per h (52–56 h); 30 µg/kg per h (56–60 h)	none reported			18	2.292	0.195

*study used the same control group

Table 2: Patient Demographics of Included Studies

		Ethnicity		Race			Mean Age	Age Range
		Hispanic /Latino	Not Hispanic/Latino	African American	White	Other		
1	Kanes et al.	0	21	13	8	0	28.1	20-40
2	Meltzer-Brody et al. (A)	11	82	32	59	2	27.2	18-45
3	Meltzer-Brody et al. (A)	15	76	29	57	5	27.4	18-45
4	Meltzer-Brody et al. (B)	25	83	42	65	1	27.9	18-45
5	Kanes et al.	not reported	not reported	not reported	not reported	not reported	not reported	27-42
6	Patterson et al.	15	1	0	16	0	31	23-37
7	Balan et al.	1	17	0	17	1	31	24-41

4.3. Meta-Analysis

The five included studies yielded seven effect sizes. The standardized mean differences of the randomized controlled trials were estimated using the least square means reduction scores reported by each study, and a forest plot was generated as shown in Figure 2. The fixed effects analysis yielded a mean estimate of 3.52 with a confidence interval of (3.16, 3.88). The effect sizes of Meltzer-Brody et al. (A.1) and Meltzer-Brody et al. (A.2) are correlated with each other since they used the same control group. The Q-test for heterogeneity was significant ($Q(3) = 12.86, p = 0.005$).

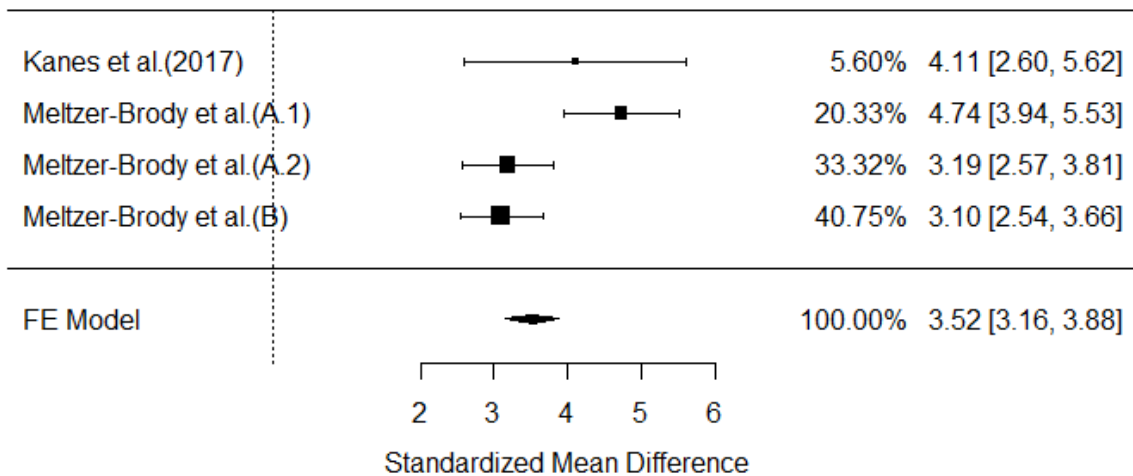


Figure 2. Forest Plot of Standardized Mean Differences for Randomized Controlled Trials

The standardized mean change effect size was estimated for the single-group studies that compared baseline scores to post-treatment scores, and a forest plot was generated as shown in Figure 3 below. The fixed effects analysis yielded a mean estimate of 2.97 with a confidence interval of (2.21, 3.73). The Q-test for heterogeneity was significant ($Q(2) = 11.4964, p = 0.0032$).

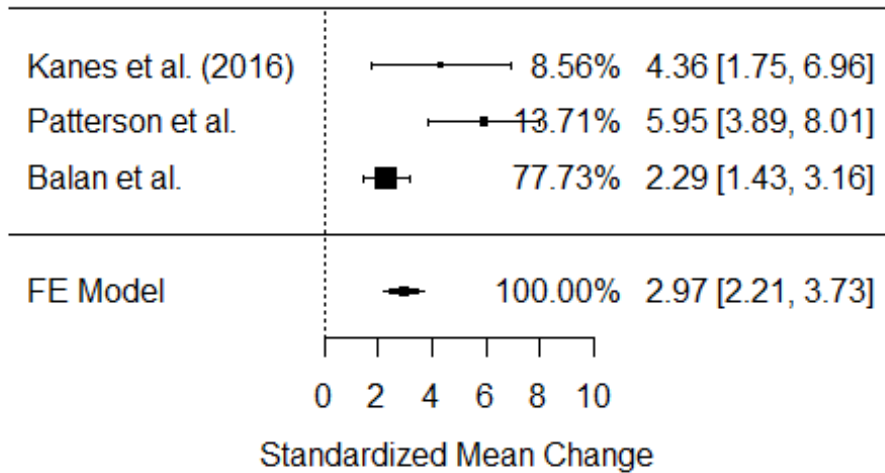


Figure 3: Forest Plot of Standardized Mean Change for Single-Group Pre-Post Studies

4.4. Publication Bias

The Egger's Regression Test for Funnel Plot Asymmetry yielded a result of $z = -8.4661$, $p\text{-value} < 0.0001$, and a confidence interval of (4.2927, 5.7529). Figure 4 below shows the contoured funnel plot generated from the effect sizes of the included studies. All of the effect sizes fall outside of the funnel. The four data points on the right of the funnel are the standardized mean differences of the randomized controlled trials, which are positive. The negative data points represent the standardized mean differences of the studies analyzing the pre-treatment and post-treatment data. Since none of the effect sizes fall within the white area of the funnel, the effect sizes are statistically significant. The results are indicative of the presence of publication bias.

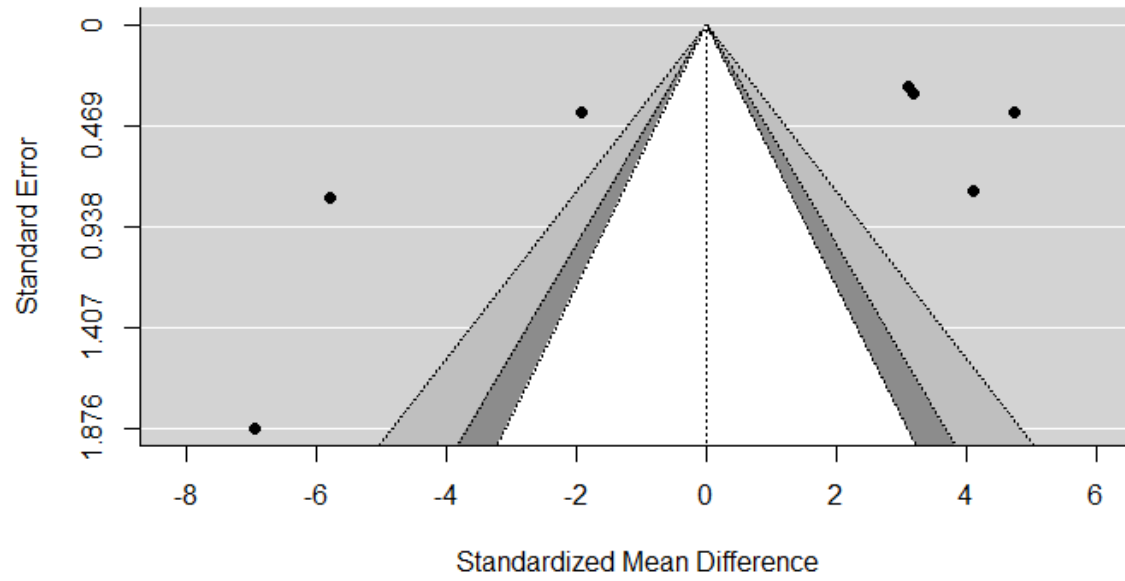


Figure 4. Contoured Funnel Plot of Standardized Mean Difference

Chapter V – Discussion & Conclusion

5.1. Interpretation of Results

The positive estimate for the forest plot of the randomized controlled trials indicates treatment has led to improvements in depression scores. This positive estimate is a result of the studies reporting changes in depression scores in terms of least square mean reduction scores. The p-value of the estimate from the fixed effects model is <0.0001 , which is less than the alpha value of 0.05. This indicates that the results are statistically significant. The calculated confidence interval does not include zero. Because of the statistically significant p-value and confidence interval, the null hypothesis of a zero effect size is rejected. The significance of the Q-test for heterogeneity indicates that there is a large variation across the studies (West et al., 2011).

The forest plot of the studies comparing depression scores after brexanolone treatment to baseline scores yielded a positive estimate from the fixed effect model. Based on the way RStudio calculates the standardized mean change using raw score standardization with heteroscedastic population variances, the positive estimate indicates that the brexanolone treatment has led to improvements in depression scores. The p-value of the estimate is <0.0001 , which is less than the alpha value of 0.05. Because of this, it can be concluded that the results are statistically significant. In addition to this, the calculated confidence interval does not include zero. The null hypothesis of a zero effect size is rejected for the pre-post studies. The Q-test for heterogeneity was also significant for the pre-post studies which indicates large variation across the studies (West et al., 2011).

The Egger's Regression Test yielded a p-value of <0.0001 . This value is less than the alpha value of 0.05; therefore, the results of the regression are statistically significant. Because of

this, there is evidence that publication bias is present among the studies. However, the capacity of Egger's test to detect bias is limited since there are only five included studies, so the results of this analysis should be treated with caution (Egger et al., 1997).

Upon initial inspection, the data points of the contoured funnel plot in Figure 4 appear to be distributed asymmetrically. The data points of the funnel plot all fall outside of the funnel plot, which suggests that the calculated effect sizes are statistically significant (Littell et al., 2022). The appearance of the funnel plot and distribution of the data points indicate evidence of publication bias among the studies. This interpretation is supported by Egger's Regression Test.

5.2. Limitations

Due to the nature of the systematic review and meta-analysis, there were many methodological limitations. The first limitation is that the methodology was carried out by only one researcher. This may have resulted in missed eligible studies during the screening process (Waffenschmidt, 2019). Another consideration is that grey literature was not searched for potentially eligible studies. This includes databases in which researchers released relevant data but have not yet published an article. Including data from grey literature would most likely have impacted the statistical analysis by lessening the influence of publication bias. These limitations affected the potential number of articles that could have been included in the study. Had more studies been included in the review, a meta-regression to examine heterogeneity could have been calculated along with the other statistical analyses. Since a meta-regression was not conducted, it is uncertain how effect sizes varied across the studies (Littell et al., 2022). Lastly, the lack of studies also impacts the generalizability of the results of the review. All of the included studies

were conducted in the US with patients mainly from African American and White racial groups. Therefore, the generalizability is limited to postpartum women within the US.

5.3. Future Research

There is still much research that needs to be done regarding postpartum depression in general, and specifically about the effects of brexanolone on women who suffer from postpartum depression. Some of the studies included in this review utilized different dosing schedules and concentrations of brexanolone. Future systematic reviews should address the implications, if any, of utilizing different dosing schedules and concentrations and if one is more effective than another. The mechanism of action and biological interactions of brexanolone also needs to be established. Further randomized controlled trials should be conducted utilizing larger sample sizes with patients from countries other than the US and with different cultural, socioeconomic, and racial backgrounds to improve the generalizability of the results.

5.4. Conclusion

Postpartum depression is a debilitating mental illness that millions of women suffer from around the world. Data from the statistical analyses of the current systematic review suggest that a treatment of brexanolone infusion is effective at reducing the incidences of postpartum depression, as indicated by the reduction in depression scores measured using the Hamilton Depression Rating Scale. However, this treatment should only be administered to women within 12 months postpartum who are suffering from moderate to severe depression, as encouraged by the current literature. There is still much to be learned about the mechanism of action of

brexanolone and how it should effectively be used in terms of dosing regimen to treat women suffering from postpartum depression.

References

- Azhar Y, Din AU. Brexanolone. In: StatPearls. StatPearls Publishing, Treasure Island (FL); 2022. PMID: 31082098.
- Balan, I., Patterson, R., Boero, G., Krohn, H., O'Buckley, T. K., Meltzer-Brody, S., & Morrow, A. L. (2023). Brexanolone Therapeutics in post-partum depression involves inhibition of systemic inflammatory pathways. *eBioMedicine*, 89, 104473.
<https://doi.org/10.1016/j.ebiom.2023.104473>
- Bloch, M., Schmidt, P. J., Danaceau, M., Murphy, J., Nieman, L., & Rubinow, D. R. (2000). Effects of gonadal steroids in women with a history of postpartum depression. *American Journal of Psychiatry*, 157(6), 924–930. <https://doi.org/10.1176/appi.ajp.157.6.924>
- Bravata, D. M., & Olkin, I. (2001). Simple pooling versus combining in meta-analysis. *Evaluation & the Health Professions*, 24(2), 218–230.
<https://doi.org/10.1177/01632780122034885>
- Carrozzino, D., Patierno, C., Fava, G. A., & Guidi, J. (2020). The hamilton rating scales for depression: A critical review of clinimetric properties of different versions. *Psychotherapy and Psychosomatics*, 89(3), 133–150. <https://doi.org/10.1159/000506879>
- Dennis, C.-L., & Hodnett, E. D. (2007). Psychosocial and psychological interventions for treating postpartum depression. *Cochrane Database of Systematic Reviews*.
<https://doi.org/10.1002/14651858.cd006116.pub2>

Edwards Z, Preuss CV. GABA Receptor Positive Allosteric Modulators. In: StatPearls.

StatPearls Publishing, Treasure Island (FL); 2022. PMID: 32119330.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634.

<https://doi.org/10.1136/bmj.315.7109.629>

Ganeshkumar, P., & Gopalakrishnan, S. (2013). Systematic reviews and meta-analysis:

Understanding the best evidence in Primary Healthcare. *Journal of Family Medicine and Primary Care*, *2*(1), 9. <https://doi.org/10.4103/2249-4863.109934>

Gerbası, M. E., Eldar-Lissai, A., Acaster, S., Fridman, M., Bonthapally, V., Hodgkins, P., Kanesh, S. J., & Meltzer-Brody, S. (2020). Associations between commonly used patient-reported outcome tools in postpartum depression clinical practice and the Hamilton Rating Scale for Depression. *Archives of Women's Mental Health*, *23*(5), 727–735.

<https://doi.org/10.1007/s00737-020-01042-y>

Grace, S. L., Evindar, A., & Stewart, D. E. (2003). The effect of postpartum depression on child cognitive development and behavior: A review and critical analysis of the literature.

Archives of Women's Mental Health, *6*(4), 263–274. <https://doi.org/10.1007/s00737-003-0024-6>

Haddaway, Neal R., Matthew J. Page, Chris C. Pritchard, and Luke A. McGuinness.

“*PRISMA2020*: An R Package and Shiny App for Producing PRISMA 2020-compliant Flow Diagrams, with Interactivity for Optimised Digital Transparency and Open

Synthesis.” *Campbell Systematic Reviews* 18, no. 2 (June 2022): e1230.

<https://doi.org/10.1002/cl2.1230>.

Huntington-Klein, N. (2022). *The effect: An introduction to research design and causality*. CRC Press.

Kanes, Stephen, Helen Colquhoun, Handan Gunduz-Bruce, Shane Raines, Ryan Arnold, Amy Schacterle, James Doherty, et al. “Brexanolone (SAGE-547 Injection) in Post-Partum Depression: A Randomised Controlled Trial.” *The Lancet* 390, no. 10093 (July 29, 2017): 480–89. [https://doi.org/10.1016/S0140-6736\(17\)31264-3](https://doi.org/10.1016/S0140-6736(17)31264-3).

Kanes, Stephen J., Helen Colquhoun, James Doherty, Shane Raines, Ethan Hoffmann, David R. Rubinow, and Samantha Meltzer-Brody. “Open-Label, Proof-of-Concept Study of Brexanolone in the Treatment of Severe Postpartum Depression.” *Human Psychopharmacology: Clinical and Experimental* 32, no. 2 (2017): e2576. <https://doi.org/10.1002/hup.2576>.

Kroh, S. J., & Lim, G. (2021). Perinatal depression. *International Anesthesiology Clinics*, 59(3), 45–51. <https://doi.org/10.1097/aia.0000000000000325>

Levett, P. (2023, October 13). *Research guides: Systematic reviews: Data Extraction/coding/study characteristics/results*. Data Extraction/Coding/Study characteristics/Results - Systematic Reviews - Research Guides at George Washington University. https://guides.himmelfarb.gwu.edu/systematic_review/data-extraction

Levis, B., Negeri, Z., Sun, Y., Benedetti, A., & Thombs, B. D. (2020). Accuracy of the Edinburgh Postnatal Depression Scale (EPDS) for screening to detect major depression

among pregnant and postpartum women: Systematic review and meta-analysis of individual participant data. *BMJ*, m4022. <https://doi.org/10.1136/bmj.m4022>

Littell, J. H., Valentine, J. C., & Pigott, T. D. Unit 6: Introduction to Effect Sizes. In J. C. Valentine, J. H. Littell, & S. Young (Eds.), *Systematic reviews and meta-analysis: A Campbell Collaboration online course*. Open Learning Initiative, 2022. Available from <https://oli.cmu.edu/courses/systematic-reviews-and-meta-analysis/>

Lopez-Gonzalez D.M., Kopparapu A.K. Postpartum Care Of The New Mother. In: StatPearls. StatPearls Publishing, Treasure Island (FL); 2022. PMID: 33351433.

Luisi, S., Petraglia, F., Benedetto, C., Nappi, R. E., Bernardi, F., Fadalti, M., ... & Genazzani, A. R. (2000). Serum allopregnanolone levels in pregnant women: changes during pregnancy, at delivery, and in hypertensive patients. *The Journal of Clinical Endocrinology & Metabolism*, 85(7), 2429-2433.

Maguire, J., & Mody, I. (2008). GABAAR plasticity during pregnancy: relevance to postpartum depression. *Neuron*, 59(2), 207-213.

Meltzer-Brody, S., & Kanes, S. J. (2020). Allopregnanolone in postpartum depression: Role in pathophysiology and treatment. *Neurobiology of Stress*, 12, 100212. <https://doi.org/10.1016/j.ynstr.2020.100212>

Meltzer-Brody, S., Colquhoun, H., Riesenber, R., Epperson, C. N., Deligiannidis, K. M., Rubinow, D. R., Li, H., Sankoh, A. J., Clemson, C., Schacterle, A., Jonas, J., & Kanes, S. (2018). Brexanolone injection in post-partum depression: Two multicentre, double-blind,

randomised, placebo-controlled, phase 3 trials. *The Lancet*, 392(10152), 1058–1070.

[https://doi.org/10.1016/s0140-6736\(18\)31551-4](https://doi.org/10.1016/s0140-6736(18)31551-4)

Moraes, G. P., Lorenzo, L., Pontes, G. A., Montenegro, M. C., & Cantilino, A. (2017). Screening and diagnosing postpartum depression: When and how? *Trends in Psychiatry and Psychotherapy*, 39(1), 54–61. <https://doi.org/10.1590/2237-6089-2016-0034>

Mughal S, Azhar Y, Siddiqui W. Postpartum Depression. In: StatPearls. StatPearls Publishing, Treasure Island (FL); 2022. PMID: 30085612.

Nappi, R. E., Petraglia, F., Luisi, S., Polatti, F., Farina, C., & Genazzani, A. R. (2001). Serum allopregnanolone in women with postpartum “blues”. *Obstetrics & Gynecology*, 97(1), 77-80.

National Institute of Mental Health. (2020, March). Perinatal Depression Infographic (NIMH Identifier No. 20-MH-8116). U.S. Department of Health and Human Services, National Institutes of Health. Retrieved November 4, 2023, from <https://www.nimh.nih.gov/sites/default/files/documents/health/publications/perinatal-depression/perinatal-depression.pdf>.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews*, 5:210. DOI: 10.1186/s13643-016-0384-4.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J.,

Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The Prisma 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. <https://doi.org/10.1136/bmj.n71>

Patterson, R., Krohn, H., Richardson, E., Kimmel, M., & Meltzer-Brody, S. (2022). A Brexanolone Treatment Program at an academic medical center: Patient selection, 90-day posttreatment outcomes, and lessons learned. *Journal of the Academy of Consultation-Liaison Psychiatry*, *63*(1), 14–22. <https://doi.org/10.1016/j.jaclp.2021.08.001>

Powell, J. G., Garland, S., Preston, K., & Piszczatoski, C. (2019). Brexanolone (zulresso): Finally, an FDA-approved treatment for postpartum depression. *Annals of Pharmacotherapy*, *54*(2), 157–163. <https://doi.org/10.1177/1060028019873320>

Privitera, G. J., & Ahlgrim-Delzell, L. (2019). 13. Quasi-Experimental and Single-Case Experimental Designs. In *Research methods for education* (1st ed., pp. 333–370). essay, SAGE.

Rich, M. (n.d.). *What is postpartum depression?*. UNICEF Parenting. <https://www.unicef.org/parenting/mental-health/what-postpartum-depression#postpartum-depression>

Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. A. (1995). The well-built clinical question: A key to evidence-based decisions. *ACP Journal Club*, *123*(3). <https://doi.org/10.7326/acpjc-1995-123-3-a12>

- Robin, A. A. (1962). The psychological changes of normal parturition. *The Psychiatric Quarterly*, 36(1–4), 129–150. <https://doi.org/10.1007/bf01586106>
- Schüle, C., Nothdurfter, C., & Rupprecht, R. (2014). The role of allopregnanolone in depression and anxiety. *Progress in Neurobiology*, 113, 79–87.
<https://doi.org/10.1016/j.pneurobio.2013.09.003>
- Shea, L., Pesa, J., Geonnotti, G., Powell, V., Kahn, C., & Peters, W. (2022). Improving diversity in study participation: Patient perspectives on barriers, racial differences and the role of Communities. *Health Expectations*, 25(4), 1979–1987. <https://doi.org/10.1111/hex.13554>
- Slomian, J., Honvo, G., Emonts, P., Reginster, J.-Y., & Bruyère, O. (2019). Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes. *Women's Health*, 15, 174550651984404. <https://doi.org/10.1177/1745506519844044>
- Stewart, D. E., & Vigod, S. (2016). Postpartum depression. *New England Journal of Medicine*, 375(22), 2177–2186. <https://doi.org/10.1056/nejmcp1607649>
- Stewart, D. E., & Vigod, S. N. (2019). Postpartum depression: Pathophysiology, treatment, and emerging therapeutics. *Annual Review of Medicine*, 70(1), 183–196.
<https://doi.org/10.1146/annurev-med-041217-011106>
- Takeshima, N., Sozu, T., Tajika, A., Ogawa, Y., Hayasaka, Y., & Furukawa, T. A. (2014). Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? *BMC Medical Research Methodology*, 14(1).
<https://doi.org/10.1186/1471-2288-14-30>

The American College of Obstetricians and Gynecologists Committee Opinion no. 630.

Screening for perinatal depression. (2015). *Obstetrics and gynecology*, 125(5), 1268–1271.

<https://doi.org/10.1097/01.AOG.0000465192.34779.dc>

Tufanaru, C., Munn, Z., Stephenson, M., & Aromataris, E. (2015). Fixed or random effects meta-analysis? common methodological issues in systematic reviews of effectiveness.

International Journal of Evidence-Based Healthcare, 13(3), 196–207.

<https://doi.org/10.1097/xeb.0000000000000065>

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(1), Article 1. <https://doi.org/10.18637/jss.v036.i03>

Viechtbauer, W. (n.d.). *Calculate effect sizes and outcome measures - ESCALC*. Calculate Effect Sizes and Outcome Measures.

<https://wviechtb.github.io/metafor/reference/escalc.html#author-1>

Waffenschmidt, S., Knellingen, M., Sieben, W., Bühn, S., & Pieper, D. (2019). Single screening versus conventional double screening for study selection in systematic reviews: A methodological systematic review. *BMC Medical Research Methodology*, 19(1).

<https://doi.org/10.1186/s12874-019-0782-0>

West, S. L., Gartlehner, G., Mansfield, A. J., Poole, C., Tant, E., Lenfestey, N., ... & Lohr, K. N. (2011). Comparative effectiveness review methods: clinical heterogeneity.

Zainur, R. Z., & Loh, K. Y. (2006). Postpartum morbidity-what we can do. *Medical journal of Malaysia*, 61(5), 651.

Zheng, W., Cai, D.-B., Zheng, W., Sim, K., Ungvari, G. S., Peng, X.-J., Ning, Y.-P., Wang, G., & Xiang, Y.-T. (2019). Brexanolone for postpartum depression: A meta-analysis of

randomized controlled studies. *Psychiatry Research*, *279*, 83–89.

<https://doi.org/10.1016/j.psychres.2019.07.006>

Zimmerman, M., Martinez, J. H., Young, D., Chelminski, I., & Dalrymple, K. (2013). Severity classification on the hamilton depression rating scale. *Journal of Affective Disorders*,

150(2), 384–388. <https://doi.org/10.1016/j.jad.2013.04.028>

Appendices

Appendix A – Screening Tool

Level 1: Initial Screening (from titles & abstracts)

1.1. Is this document about brexanolone?

- 1 Yes
- 0 No (STOP)
- 9 Can't tell (retrieve full text)

1.2. Is this document about postpartum depression?

- 1 Yes
- 0 No (STOP)
- 9 Can't tell (retrieve full text)

1.3. Does the study utilize the Hamilton Rating Scale for Depression (HAM-D) as a measure of outcome?

- 1 Yes
- 0 No (STOP)
- 9 Can't tell (retrieve full text)

Level 2: Eligibility Decision (from full text)

2.1 Does the study include a comparison group of placebo or comparing to baseline data?

- 1 Yes
- 0 No (STOP)
- 9 Can't tell

2.2 Were sample statistics reported?

- 1 Yes
- 0 No (STOP)
- 9 Can't tell

2.3 Does the study use a population of women between the ages of 18 – 45?

- 1 Yes
- 0 No (STOP)
- 9 Can't tell

2.4 Is the study population <1 year postpartum?

- 1 Yes
- 0 No (STOP)
- 9 Can't tell

Appendix B – Data Extraction Codebook

1.1 What was the comparison group?

- 1 Placebo
- 2 Baseline data
- 0 Unclear

1.2 Were adverse effects reported?

- 1 Yes
- 2 No
- 0 Unclear

1.3 Were statistics reported?

- 1 Yes
- 2 No
- 0 Unclear

1.4 What was the study size?

- Report number of participants in treatment and comparison group.

1.5 What year was the study conducted?

- Report the year(s) in which the study took place.

1.6 When was the data collected?

- Report the point in time of the brexanolone infusion in which the scientists collected HAM-D scores.

1.7 What were the sample demographics?

- Report number of participants belonging to each ethnicity: Hispanic/Latino and not Hispanic/Latino.
- Report number of participants belonging to each race: African American, White, and Other.
- Report the average age of participants.

1.8 Where was the study conducted?

- Report the country where the study took place.

1.9 What was the study design?

- Report the design of each study.

1.10 What was the dosing regimen of brexanolone?

- Report the concentration of brexanolone administered to each patient and the hours each concentration was administered.

Appendix C – RStudio Data Analysis

Activating Necessary Programs

```
# activate the programs needed
library(tidyverse)

## — Attaching core tidyverse packages ————— tidyverse 2.
0.0 —
## ✓ dplyr      1.1.0      ✓ readr      2.1.4
## ✓ forcats   1.0.0      ✓ stringr    1.5.0
## ✓ ggplot2   3.4.1      ✓ tibble     3.1.8
## ✓ lubridate 1.9.2      ✓ tidyr      1.3.0
## ✓ purrr     1.0.1
## — Conflicts ————— tidyverse_conflict
s() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()     masks stats::lag()
## i Use the http://conflicted.r-lib.org/conflicted package to force
all conflicts to become errors

library(metafor)

## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##   expand, pack, unpack
##
## Loading required package: metadat
##
## Loading the 'metafor' package (version 3.8-1). For an
## introduction to the package please type: help(metafor)
```

Generating Forest Plot for RCTs

```
# reading in data
brexrct <- read.csv("rctdata.csv")

# computing SMD
brexrct <- escalc(measure = "SMD", # the measure is the type of effect size
                 m1i = brexmean, sd1i = brexsd, n1i = brexn, # info for ou
                 r Trt group
                 m2i = cntmean, sd2i = cntsd, n2i = cntn, # info for contr
                 ol
                 data = brexrct, # name of our data set
                 var.names = c("smd", "varsmd")) # names of our ES and SE

View(brexrct)
```

```

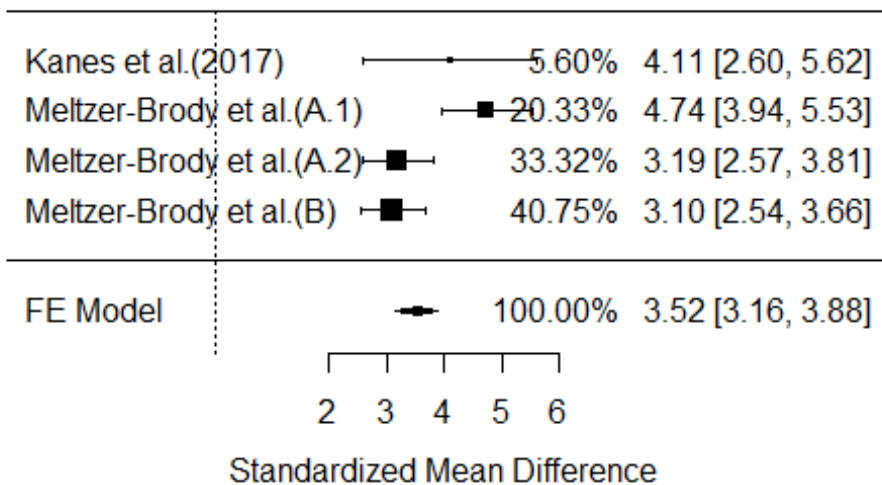
# calculating mean effect size, CI, heterogeneity stats
brexrct_fe <- rma(yi = smd, #effect size
                vi = varsm, # variance of effect size
                data = brexrct, # data set
                method = "FE") # method for doing MA

summary(brexrct_fe)

##
## Fixed-Effects Model (k = 4)
##
##   logLik  deviance      AIC      BIC      AICc
## -6.5371  12.8646  15.0742  14.4605  17.0742
##
## I^2 (total heterogeneity / total variability): 76.68%
## H^2 (total variability / sampling variability): 4.29
##
## Test for Heterogeneity:
## Q(df = 3) = 12.8646, p-val = 0.0049
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## 3.5200  0.1824  19.3016  <.0001  3.1626  3.8775  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# forest plot
studynames <- c("Kanes et al.(2017)", "Meltzer-Brody et al.(A.1)", "Meltzer-Bro
dy et al.(A.2)", "Meltzer-Brody et al.(B)")
forest(brexrct_fe, showweights = T, slab = studynames)

```



Generating Forest Plots for Pre-Post Studies

```
# read in the pre-post data
brexprepost <- read.csv("prepostdata.csv")

# computing SMCR
brexprepost <- escalc(measure = "SMCRH", # SMD with pooled sds
  m1i = premn, # pretest mean
  sd1i = presd, # pretest sd
  m2i = postmn, # posttest mean
  sd2i = postsd, # posttest sd
  ni = totn, # sample size
  ri = corr, # correlation between pre and post
  data = brexprepost, # data set
  var.names = c("smdw", "varsmdw"))

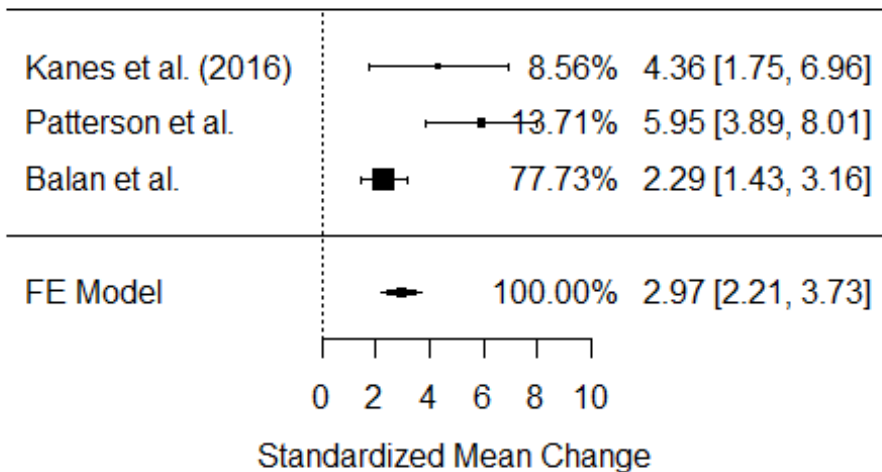
view(brexprepost)

# calculating mean effect size, CI, heterogeneity stats
brexprepost_fe <- rma(yi = smdw, #effect size
  vi = varsmdw, # variance of effect size
  data = brexprepost, # data set
  method = "FE") # method for doing MA

summary(brexprepost_fe)

##
## Fixed-Effects Model (k = 3)
##
##   logLik  deviance      AIC      BIC     AICc
## -8.0206  11.4964  18.0412  17.1398  22.0412
##
## I^2 (total heterogeneity / total variability):  82.60%
## H^2 (total variability / sampling variability):  5.75
##
## Test for Heterogeneity:
## Q(df = 2) = 11.4964, p-val = 0.0032
##
## Model Results:
##
## estimate      se    zval    pval   ci.lb   ci.ub
##  2.9706  0.3890  7.6373 <.0001  2.2083  3.7330 ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# forest plot
study_names <- c("Kanes et al. (2016)", "Patterson et al.", "Balan et al.")
forest(brexprepost_fe, showweights = T, slab=study_names)
```



Publication Bias (Funnel Plot & Egger's Regression)

```
pubbias <- read.csv("combineddata.csv")

pubbias <- escalc(measure = "SMD", # the measure is the type of effect size
  m1i = trtmean, sd1i = trtsd, n1i = trtn, # info for our T
  rt group
  m2i = cntmean, sd2i = cntsd, n2i = cntn, # info for contr
  ol
  data = pubbias, # name of our data set
  var.names = c("smd", "varsmc")) # names of our ES and SE

# funnel plot
pubbias_mean <- rma(yi = smd,
  vi = varsmc,
  data = pubbias,
  method = "FE")

summary(pubbias_mean)

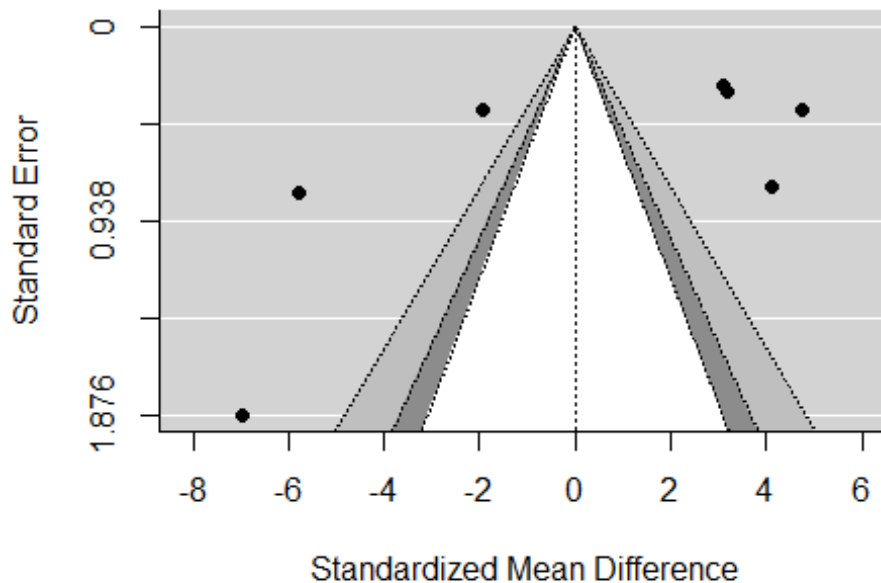
##
## Fixed-Effects Model (k = 7)
##
##   logLik   deviance      AIC      BIC      AICc
## -148.4726  292.2176  298.9452  298.8911  299.7452
##
## I^2 (total heterogeneity / total variability):  97.95%
```

```

## H^2 (total variability / sampling variability): 48.70
##
## Test for Heterogeneity:
## Q(df = 6) = 292.2176, p-val < .0001
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## 2.1835 0.1621 13.4677 <.0001 1.8657 2.5012 ***
##
## ---
## Signif. codes:  0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1

funnel(pubbias_mean, # RE results
       level=c(90, 95, 99), #significance levels
       shade = c("white", "gray55", "gray75"), # colors in plot
       refline = 0, # reference line for the 0 effect
       legend = F)

```



```

# Egger's regression
regtest(pubbias_mean)

##
## Regression Test for Funnel Plot Asymmetry
##
## Model:      fixed-effects meta-regression model
## Predictor:  standard error

```

```
##  
## Test for Funnel Plot Asymmetry: z = -8.4661, p < .0001  
## Limit Estimate (as sei -> 0): b = 5.0228 (CI: 4.2927, 5.7529)
```