Bioidentical Estrogen for Menopausal Depressive Symptoms: A Systematic Review and Meta-Analysis

James M. Whedon, DC, MS,1 Anupama KizhakkeVeettil, BAMS, MAOM, PhD2
Nancy A. Rugo, MSN, ARNP3 and Kelly A. Kieffer, MD, MS4

Abstract

Background: Proponents of bioidentical estrogens claim that they are superior for treating menopausal symptoms, including depressive symptoms. Small trials examining the effects of bioidentical estrogens on depressive symptoms show conflicting results. We conducted a systematic review to assess the effectiveness and safety of bioidentical estrogens for treatment of depressive symptoms in peri- and postmenopausal women.

Methods: We searched the scientific literature for randomized controlled trials of at least 4 weeks duration, comparing bioidentical estrogen with placebo for depressive symptoms in menopausal women. The main outcome measure was improvement in depressive symptoms on a validated scale.

Results: We found 12 clinical trials that met inclusion criteria, two of which contained insufficient data for quantitative analysis. In the 10 studies (inclusive of 1208 subjects) for which complete data were available for inclusion in the meta-analysis, bioidentical estrogen had no clinically significant effect on depressive symptoms (standardized mean difference [SMD] = -0.02; confidence interval [95% CI] = -0.41 to +0.38). Pooled studies were highly heterogeneous, and numerous approaches to reducing heterogeneity were unsuccessful. Subgroup analyses showed no significant difference in effect for women treated with adjunctive progestogen, women treated with unopposed estrogen, perimenopausal, or postmenopausal and mixed populations. A possible benefit in perimenopausal women treated with unopposed estradiol may have been diluted by studies including older postmenopausal women whose depressive symptoms were unrelated to menopause.

Conclusions: In this first systematic review of bioidentical hormone replacement therapy, we found that bioidentical estrogen has no clear benefit in treating depressive symptoms in menopausal women, but heterogeneity of available studies limits the potential for definitive conclusions. Future research should compare bioidentical estrogen with nonbioidentical estrogen for treatment of depressive symptoms in perimenopausal women.

Introduction

Depressive symptoms are common in menopause,1 but the extent to which declining ovarian function contributes to these symptoms and estrogen replacement relieves them is unclear.2 Observational studies suggest that depressive symptoms are more frequent in the perimenopause, particularly in the late menopausal transition,3,4 with a peak around the last menstrual period. In a minority of women, more severe depressive symptoms may persist for the first 2–3 years after menopause before declining.5 The frequency of depressive symptoms increases again in later life, beginning in the mid-60s,6 however, this occurs in both men and women and has no apparent relationship to menopause.

The results of large randomized trials of menopausal hormone therapy (MHT) indicate that for most women the benefits of long-term treatment do not exceed the harms with respect to major medical conditions.7,8 Some experts have recommended short-term MHT to relieve symptoms related to the menopausal transition,6,9 including depressive symptoms,10 but the efficacy of MHT in improving menopausal depressive symptoms is uncertain. A meta-analysis published in 1997 that included nonrandomized and uncontrolled studies of various hormonal regimens concluded that treatment with estrogen improved depressive symptoms in menopause.11 Conversely, data from the Women’s Health Initiative (WHI), the largest randomized controlled trial of MHT conducted to date, showed no improvement in depressive symptoms for women taking either conjugated equine estrogens (CEE) alone12,13 or a combination of CEE and medroxyprogesterone.14,15 In a 2010 scientific statement on postmenopausal MHT, the Endocrine Society concluded that level B evidence

1University Health System, Southern California University of Health Sciences, Whittier, California.
2Department of Oriental Medicine, Southern California University of Health Sciences, Whittier, California.
3Department of Anesthesiology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire.
4Department of Internal Medicine, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire.
supports an antidepressant effect of estradiol in perimenopausal but not postmenopausal women. These discrepancies in the literature could be related to overestimates of benefit in studies of low to moderate methodological quality, or to a clinically meaningful difference in the effects of different forms of estrogen, or to differences in study populations.

Some women and some physicians express concerns regarding CEE, which was used in the major MHT trials, because it includes estrogens that do not occur naturally in humans. Treatment with bioidentical estrogens has been proposed as an alternative to the use of CEE for symptoms occurring during menopause, including depressive symptoms. Bioidentical estrogens are molecularly identical to estrogens made in the human body, and are available both in customized form through a compounding pharmacist and in standardized FDA-approved formulations. There is some disagreement among experts regarding the specific forms of estrogen that are bioidentical. In a 2007 review of bioidentical MHT, Cirigliano described the injectable estrogens estradiol valerate and estradiol cypionate as a combination of CEE and MHT, and reported superior relief of sexual symptoms with bioidentical hormones. This definition is consistent with the lists of bioidentical estrogens identified in recent reviews by Conaway and Pattimakiel. The relative efficacy and safety of individually compounded formulations of MHT, while a controversial and important topic, was beyond the scope of this study.

Advocates of bioidentical estrogens suggest that they are safer, more amenable to individualization of treatment, and more supportive of health than CEE. In a survey of women with menopausal symptoms, a majority believed that bioidentical hormones were safer than conventional MHT, and reported superior relief of sexual symptoms with bioidentical hormones. Two recent reviews favor bioidentical hormone replacement over conventional MHT, based upon the accumulated evidence for efficacy and safety. Other authors express concern regarding the risks associated with compounded formulations that are not FDA approved, and maintain that there is a paucity of evidence regarding bioidentical MHT in general. There is insufficient evidence to make conclusions regarding the use of compounded hormones as compared with FDA-approved formulations.

Results of large randomized trials comparable to the WHI are not available to provide definitive answers regarding the efficacy and safety of bioidentical MHT. Small randomized clinical trials evaluating the efficacy of bioidentical estrogens in the treatment of depressive symptoms have been conducted since the 1997 meta-analysis, but until now a systematic review of higher quality studies has not been undertaken, and no systematic review limited to bioidentical estrogen formulations has been performed. We report the results of a systematic review of double-blind, randomized controlled trials evaluating the effects of bioidentical estrogens compared to placebo on depressive symptoms in menopausal women.

Methods

Review protocol

We developed a protocol outlining a systematic approach to identification and selection of relevant studies, using standard methodology for analysis as described in the Cochrane Handbook, and applying the PRISMA guidelines for reporting of findings. The original protocol is available from the authors upon request.

Study eligibility criteria

We included randomized controlled trials reporting the effect on depressive symptoms of bioidentical estrogens compared to placebo in menopausal women. Bioidentical estrogens included in the review were 17-beta estradiol, estradiol acetate, estradiol hemihydrates, estril, estrone, and estropipate, administered orally, transdermally, or vaginally. We included studies of bioidentical estrogen alone or combined with a progestogen, compared to placebo delivered by the same mechanism as the study drug. We defined the population of peri- and postmenopausal women to include women without hysterectomy who had no menses for at least 12 months before enrollment, women with bilateral oophorectomy with or without hysterectomy, women over 40 with any of the following: irregular menses, self-reported vasomotor symptoms, or serum follicle stimulating hormone level >25 µIU/mL, and all women over age 60. We included only studies that used a previously validated depression scale, and specified a minimum duration of follow-up of at least 4 weeks since depressive symptoms may require a minimum of 4–6 weeks to demonstrate a measurable response to pharmacologic therapy.

Outcome measures

Primary outcome. The primary outcome analyzed was improvement in depressive symptoms as measured by a validated depression scale. Accepted scales included the following: the Hamilton Rating Scale for Depression, Beck Depression Inventory, Center for Epidemiologic Studies Depression scale, Montgomery-Asberg Depression Rating Scale, Geriatric Depression Scale, General Health Questionnaire, Hospital Anxiety and Depression Scale, Profile of Mood States, Patient Health Questionnaire, Depression Anxiety Stress Scales, Mental Health Inventory, Brief Assessment Schedule Depression Cards, General Well Being Index, and Women’s Health Questionnaire.

Secondary outcome. As a secondary outcome, we evaluated the effect of treatment with bioidentical estrogen on improvement of vasomotor symptoms. Conventional MHT is beneficial for treatment of vasomotor symptoms associated with menopause, and estradiol is commonly prescribed to relieve vasomotor symptoms, but to date no systematic review has exclusively focused on bioidentical estrogen for vasomotor symptoms. Studies were included in the analyses of effects on vasomotor symptoms only if a validated scale was used to measure the outcome.

Adverse outcomes. Adverse outcomes studied included new or worsening irregular or heavy vaginal bleeding; endometrial hyperplasia or cancer; weight gain; and any serious adverse event including vascular events, thromboembolic events, new diagnosis of breast cancer, new diagnosis of any other cancer, or death.

Search methods

We performed an initial literature search in March 2012 and updated it in May 2015. We searched Medline, the
**BIOIDENTICAL ESTROGEN FOR MENOPAUSAL SYMPTOMS**

**Table 1. Search Terms**

<table>
<thead>
<tr>
<th>Domain</th>
<th>MeSH headings</th>
<th>Entry terms and keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bioidentical estrogen</td>
<td>17 beta estradiol</td>
<td>17 beta estradiol</td>
</tr>
<tr>
<td></td>
<td>17 beta estradiol</td>
<td>Oestradiol</td>
</tr>
<tr>
<td>Estradiol</td>
<td>17 beta oestradiol</td>
<td>17 beta oestradiol</td>
</tr>
<tr>
<td>Estrone</td>
<td>17 beta oestradiol</td>
<td>17 beta oestradiol</td>
</tr>
<tr>
<td>Hormone replacement</td>
<td>Natural hormone replacement</td>
<td>Estropipate</td>
</tr>
<tr>
<td>therapy</td>
<td>therapy</td>
<td>Bioidentical HRT</td>
</tr>
<tr>
<td>Menopausal</td>
<td>Postmenopausal</td>
<td>Postmenopausal</td>
</tr>
<tr>
<td></td>
<td>Perimenopausal</td>
<td>Perimenopausal</td>
</tr>
<tr>
<td>Menopause</td>
<td>Perimenopausal</td>
<td>Perimenopausal</td>
</tr>
<tr>
<td>Depressive</td>
<td>Depression</td>
<td>Depressive disorder</td>
</tr>
<tr>
<td>symptoms</td>
<td>Disorder</td>
<td>Well being</td>
</tr>
<tr>
<td>Quality of life</td>
<td>Health</td>
<td>Mood</td>
</tr>
<tr>
<td>Mental health</td>
<td>Affect</td>
<td></td>
</tr>
</tbody>
</table>

Note: HRT, hormone replacement therapy.

We combined results of all searches and removed duplicates. One investigator (A.K., K.K., or J.W.) screened the titles and abstracts of our initial search results to remove studies that did not meet inclusion criteria. A second reviewer evaluated 10% of all abstracts to confirm the validity of this single screener approach. Two investigators examined the full text of all remaining studies; consensus of at least two reviewers was required to make a final determination regarding study inclusion.

**Data extraction and quality assessment**

We collected data regarding study methodology and outcomes from published studies that met inclusion criteria using a piloted, standardized data collection form. Two independent reviewers extracted data from each eligible study; discrepancies were resolved by consensus. For each study, we collected data related to study design, characteristics of the study sample, interventions, methodology, and all outcomes of interest. We attempted to obtain any missing or incomplete data by contacting study authors, and excluded from meta-analysis any studies for which we were unable to obtain complete outcomes data. When we identified studies with duplicate or overlapping outcomes data, we included only the study reporting the most complete data for each outcome, or if data were equally complete we included the most recent report. We analyzed the methodologies of our included studies using the Cochrane Collaboration’s tool for assessing the risk of bias in included studies. Two independent reviewers assessed the risk of bias for each study, and discrepancies were resolved by consensus.

**Analysis**

Primary analysis. We summarized the findings for our primary outcome by calculating a standardized mean difference (SMD). To interpret the summary SMD, we multiplied each boundary of the 95% confidence interval (CI) by the pooled baseline mean standard deviation on the Center for Epidemiologic Studies Depression scale from our largest included study. This provides an approximate range of the possible clinical effects of intervention compared to placebo. For the one study that reported outcomes by two different depression scales, we analyzed the results of the Hamilton Rating Scale for Depression, which was the most commonly used scale among the included studies. We qualitatively summarized results for the secondary outcome of vasomotor symptoms, which were reported in a minority of studies and with a variety of scales. All harms outcomes were reported as dichotomous variables, which we summarized qualitatively due to the inconsistent measurement of harms in the included studies. For the primary outcome, we conducted data synthesis using a random effects model, which estimates the distribution of treatment effects across studies. We reported directional trends in individual studies for the secondary outcome. We assessed heterogeneity for each summary estimate using the p-value for the χ² test and an I² percentage. We considered heterogeneity to be present among the findings of the included studies when either the p-value was <0.10 or the I² exceeded 50%. If heterogeneity was identified, apparent outlier studies were removed to determine their effect on heterogeneity and to identify a subset of studies that could be pooled for analysis. We informally assessed for heterogeneity of findings for vasomotor symptoms. We assessed for the possible presence of reporting biases in our primary outcome (depressive symptoms) by visually inspecting a funnel plot. We used RevMan 5 (The Nordic Cochrane Centre, Copenhagen) to perform the meta-analysis.
Subgroup analyses. To reduce the effect of heterogeneity among the included studies, we prespecified three subgroup analyses: (1) type of estrogen (estradiol or another form of bioidentical estrogen); (2) progestogen use (unopposed estrogen or estrogen with adjunctive progestogen as the study intervention); and (3) phase of menopause of the subjects at enrollment (perimenopausal or postmenopausal). Because of the variability of inclusion criteria across the different studies, it was not possible to choose a single criterion for menopausal phase that would reliably and exclusively include studies of definitively perimenopausal women. Given the higher biological plausibility that hormonal treatment might benefit perimenopausal women, we felt it was important to attempt to analyze the treatment effect in this subgroup. To minimize the chance of missing a clinically significant effect, we included studies in the perimenopausal subgroup if they included women who met one or both of the following criteria: (1) <60 years of age with ongoing vaso-motor symptoms, or (2) last menstrual period within 2 years. Women were postmenopausal if they were over 60 or had surgical menopause or their last period > 2 years before study enrollment. To avoid masking a potential effect of hormonal treatment in perimenopausal women, we labeled studies that included women meeting the criteria for both subgroups or in which the stage of menopause was unclear as “mixed” with regards to menopausal status, and grouped them with studies including only postmenopausal women for analysis.

Results

Description of studies

Results of search. Searches of online databases yielded 610 articles from Medline, 201 articles from the Cochrane Library, 398 articles from CINAHL, and 192 articles from IBIDS. Searching and screening of additional sources yielded two additional articles, both from Clinicaltrials.gov, for a total of 1403 records. We identified and excluded 241 duplicate records. Of the remaining 1162 unique studies, we excluded 1125 based on title and abstract review. We reviewed 37 full text articles; 12 articles met all inclusion criteria and were included in the qualitative review. Ten of the included articles reported adequate outcomes data for inclusion in the meta-analysis. See Figure 1 for the complete study selection flow diagram.

Included studies. Key characteristics of the 12 studies that met all inclusion criteria for the systematic review can be found in Table 2. Total enrollment was 1268 subjects in the

FIG. 1. Study selection flow chart.
## Table 2. Characteristics of Included Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Mean age (years)</th>
<th>Menopause status</th>
<th>Depression scale</th>
<th>Mean baseline depression score</th>
<th>Estrogen type</th>
<th>Estrogen dose, route</th>
<th>Progestogen</th>
<th>Weeks of treatment</th>
<th>Completeness of follow-up, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thomson and Oswald</td>
<td>42</td>
<td>49.1</td>
<td>Peri</td>
<td>Hamilton</td>
<td>Mild</td>
<td>Estrone</td>
<td>1.5 mg twice daily, oral</td>
<td>None</td>
<td>8</td>
<td>81</td>
</tr>
<tr>
<td>Wiklund et al.</td>
<td>223</td>
<td>52.65</td>
<td>Mixed</td>
<td>General Well Being</td>
<td>Unclear</td>
<td>Estradiol</td>
<td>50 µg/day, transdermal</td>
<td>None</td>
<td>12</td>
<td>94</td>
</tr>
<tr>
<td>Derman et al.</td>
<td>82</td>
<td>50</td>
<td>Peri</td>
<td>Beck Short Form</td>
<td>Mild</td>
<td>Estradiol</td>
<td>1–2 mg/day, oral†</td>
<td>Cyclic Norethindrone</td>
<td>16</td>
<td>57</td>
</tr>
<tr>
<td>Saletu et al.</td>
<td>64</td>
<td>51.2</td>
<td>Peri</td>
<td>Hamilton</td>
<td>Moderate</td>
<td>Estradiol</td>
<td>50 µg/day, transdermal</td>
<td>None</td>
<td>13</td>
<td>83</td>
</tr>
<tr>
<td>Klaiber et al.</td>
<td>38</td>
<td>53.5</td>
<td>Mixed</td>
<td>Hamilton &amp; Beck</td>
<td>NR</td>
<td>Estrone</td>
<td>1.5 mg/day, oral</td>
<td>Cyclic Norethindrone</td>
<td>8</td>
<td>NR</td>
</tr>
<tr>
<td>Bech et al.</td>
<td>151</td>
<td>NR</td>
<td>Peri</td>
<td>Beck</td>
<td>Normal</td>
<td>Estradiol</td>
<td>1–2 mg/day, oral†</td>
<td>Daily or Cyclic Norethindrone</td>
<td>52</td>
<td>70</td>
</tr>
<tr>
<td>Soares et al.</td>
<td>50</td>
<td>49.8</td>
<td>Peri</td>
<td>Montgomery-Asberg</td>
<td>Moderate</td>
<td>Estradiol</td>
<td>100 µg/day, transdermal</td>
<td>None</td>
<td>12</td>
<td>90</td>
</tr>
<tr>
<td>Morrison et al.</td>
<td>57</td>
<td>62.3</td>
<td>Post</td>
<td>Hamilton &amp; Epidemiologic Studies</td>
<td>Moderate</td>
<td>Estradiol</td>
<td>0.1 mg/day transdermal</td>
<td>None</td>
<td>8</td>
<td>96</td>
</tr>
<tr>
<td>Yaffe et al.</td>
<td>417</td>
<td>66.75</td>
<td>Post</td>
<td>Epidemiologic Studies</td>
<td>Normal</td>
<td>Estradiol</td>
<td>14 µg/day, transdermal</td>
<td>None</td>
<td>104</td>
<td>90</td>
</tr>
<tr>
<td>Pefanco et al.</td>
<td>57</td>
<td>76</td>
<td>Post</td>
<td>Geriatric Depression Scale</td>
<td>Normal</td>
<td>Estradiol</td>
<td>0.25 mg/day, NR</td>
<td>Daily Micronized Progesterone</td>
<td>156</td>
<td>79</td>
</tr>
<tr>
<td>Marinho et al.</td>
<td>65</td>
<td>54.1</td>
<td>Mixed</td>
<td>Hamilton</td>
<td>Normal</td>
<td>Estradiol</td>
<td>2 mg/day, oral</td>
<td>None</td>
<td>12</td>
<td>88</td>
</tr>
<tr>
<td>Newhouse et al.</td>
<td>22</td>
<td>64.3</td>
<td>Mixed</td>
<td>Beck</td>
<td>Normal</td>
<td>Estradiol</td>
<td>1 mg/day, oral×30 days; 2 mg/day, oral×60 days</td>
<td>None</td>
<td>13</td>
<td>100</td>
</tr>
</tbody>
</table>

Beck = Beck Depression Inventory; Beck Short Form = Beck Depression Inventory Short Form; Epidemiologic Studies = Center for Epidemiologic Studies Depression Scale; General Well Being = General Well Being Index; Hamilton = Hamilton Depression Rating Scale; Montgomery-Asberg = Montgomery-Asberg Depression Rating Scale.

*Postmenopausal = age 60 or older, or minimum of 2 years since surgical menopause or LMP; Perimenopausal = <2 years since LMP, or <60 years of age and at least moderate ongoing vasomotor symptoms; Mixed = included subjects meeting perimenopausal and postmenopausal criteria.

*Mean severity of depressive symptoms at baseline based on a validated scale.

†Cycles of 2 mg/day po×22 days, then 1 mg/day×6 days.

‡Excluded from meta-analysis.

§Group 1 (n=50): 2 mg/day po; Group 2 (n=50): cycles of 2 mg/day po×22 days, then 1 mg/day×6 days.

LMP, last menstrual period; NR = not reported.
12 included studies, and was 1208 subjects in the 10 studies included in the meta-analysis. Eleven studies\textsuperscript{7,49,52-60} reported mean age of the study sample; overall mean age for these studies was 59.4 years. The study samples varied by menopausal status (perimenopausal, postmenopausal, and mixed) and by mean baseline depression scores, which ranged by category from normal to moderate depressive symptoms. The intervention in 10 of the 12 included studies\textsuperscript{47,49,51,52,54-57,59,60} was oral or transdermal estradiol in various doses; the intervention in the other two studies\textsuperscript{53,58} was estrone. In four trials,\textsuperscript{51-53,55} treatment included a daily or cyclic progestogen in combination with estrogen, and in the remaining eight trials,\textsuperscript{47,49,54,56-60} subjects in the treatment arm received unopposed estrogen. None of the studies described the use of compounded hormones in the treatment arm; seven described using a specific commercial product,\textsuperscript{47,49,51-53,56,57} one acknowledged support from a pharmaceutical company and described the dose and formulation used in the study but not the specific product\textsuperscript{58} and four described the dose and formulation but gave no indication of how it was supplied.\textsuperscript{34,55,59,60} Each study used either one or two validated scales to measure depressive symptoms; a total of eight different scales were used in the 12 studies. Six of the studies excluded women taking antidepressants\textsuperscript{49,53,56,57,59,60} the remaining studies did not address whether study subjects were concurrently taking antidepressants or other psychoactive medications.

Risk of bias in included studies

Of the 12 included studies, six\textsuperscript{49-54,58} demonstrated methodological quality concerns according to the Cochrane Risk of Bias Tool. All six of these studies were missing key information about outcomes assessments in the study subjects, specifically uneven or unreported attrition rates in the study arms. One study\textsuperscript{58} did not identify a method of randomization. Two studies\textsuperscript{49,53} had potentially inadequate blinding due to vaginal bleeding in the treatment arm, and also had concerns regarding possible reporting bias. Details regarding risk of bias are described in Table 3. We found no evidence of publication bias based on visual inspection of a funnel plot of the 10 studies included in the meta-analysis.

Effects of intervention

Depressive symptoms. Ten of the included studies reported outcomes data on depressive symptoms adequate for inclusion in the meta-analysis (Fig. 2). Overall, there was no significant difference in change in depressive symptoms with treatment compared to placebo (SMD $-0.02$, 95% CI $-0.41$ to $+0.38$). Based on the pooled baseline standard deviation of the study sample in the largest included trial,\textsuperscript{47} the boundaries of the 95% CI translate to a possible mean change on the 60 point Center for Epidemiologic Studies Depression Scale of $-1.4$ to $+1.3$. The two trials that did not report outcomes data in sufficient detail for inclusion in the meta-analysis reported no significant effect of treatment versus placebo on depressive symptoms, and did not change the overall qualitative impression of no clinical benefit. Results for the primary outcome demonstrated significant heterogeneity ($\chi^2$ $p$-value $<0.00001$, $I^2 = 89\%$). Performance of sensitivity analyses to determine the effect on heterogeneity of removing low quality studies from the analysis revealed no reduction in heterogeneity.

Only one of the studies included in the meta-analysis\textsuperscript{58} evaluated a bioidentical estrogen other than estradiol. We elected to remove this study from the subgroup analyses because doing so allowed us to evaluate treatment with a specific agent (estradiol), and because this study was the largest contributor to heterogeneity and had serious methodologic concerns, particularly lack of reporting regarding randomization. For the remaining nine studies, we analyzed treatment effect in subgroups based on menopausal status, with studies that limited enrollment to perimenopausal women in one group and studies of either postmenopausal or both post- and perimenopausal women in the other group (Fig. 3). There was no significant effect of treatment on depressive symptoms in either subgroup (SMD $-0.10$, 95% CI $-0.49$ to $0.29$ in the postmenopausal and mixed group; SMD $-0.38$, 95% CI $-1.12$ to $0.36$ in the perimenopausal group), and heterogeneity remained significant in both groups ($\chi^2$ $p$-value $<0.0001$). We also analyzed treatment effect in studies of unopposed estradiol and studies using a combination of estradiol and an adjunctive progestogen (Fig. 4). There was no significant effect of treatment in either intervention subgroup (SMD 0.05, 95% CI $-0.23$ to $0.33$ in the adjunctive progestogen group, SMD $-0.35$, 95% CI $-0.81$ to $0.12$ in the unopposed estradiol group). Heterogeneity remained significant in the unopposed estradiol group ($\chi^2$ $p$-value $<0.0001$), but was not significant in the adjunctive progestogen subgroup ($\chi^2$ $p$-value $=0.40$).

Secondary outcomes. Five studies\textsuperscript{51,52,54,56,59} using estradiol reported effect of treatment on vasomotor symptoms using a validated scale. All five studies showed a significant benefit of treatment compared to placebo.

Adverse events. Three studies\textsuperscript{49,57,59} reported more frequent vaginal bleeding in women taking bioidentical estrogens compared to placebo; one study reported no heavy or irregular bleeding in either study arm,\textsuperscript{54} and the remaining studies did not specifically comment on rates of vaginal bleeding. One study reported no serious adverse events in either arm but was not clear regarding the types of events that were considered. One study\textsuperscript{58} specifically cited that no thromboembolic events occurred and there was no difference in weight gain between the study arms. None of the studies reported on other harms.

Discussion

Summary of main results

The available evidence does not support a clinically significant effect of bioidentical estrogen on depressive symptoms in a broadly defined population of perimenopausal and postmenopausal women. Prespecified subgroup analysis indicated a statistically nonsignificant improvement in depressive symptoms favoring estradiol when given without a progestogen. Based on 95% CIs of this subgroup analysis, any true benefit is of marginal clinical relevance. However, more than half of the subjects included in this subanalysis were from two studies of older postmenopausal women that showed no benefit of treatment. This may have obscured a benefit of unopposed estradiol that was seen in three of the four studies of younger peri- and early postmenopausal women, in whom depressive symptoms would be more plausibly related to hormonal changes.
<table>
<thead>
<tr>
<th>Study</th>
<th>Sequence generation</th>
<th>Allocation concealment</th>
<th>Participants</th>
<th>Study personnel</th>
<th>Outcomes assessment</th>
<th>Completeness of outcome data</th>
<th>Absence of reporting bias</th>
<th>Other sources of bias and clarifying comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thomson and Oswald(^{58})</td>
<td>High risk</td>
<td>High risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear</td>
<td>High risk</td>
<td>Low risk</td>
<td>No description of randomization. Distribution of attrition between study arms not reported.</td>
</tr>
<tr>
<td>Wiklund et al.(^{59})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>None</td>
</tr>
<tr>
<td>Derman et al.(^{52})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>More women in placebo group dropped out due to ongoing hot flushes.</td>
</tr>
<tr>
<td>Saletu et al.(^{56})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>None</td>
</tr>
<tr>
<td>Klaiber et al.(^{53})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>Blinding “violated” by high rates of vaginal bleeding in treatment group; unclear if these subjects were included in analysis. Attrition rates not documented. High number of outcomes/associations reported, with unclear rationale.</td>
</tr>
<tr>
<td>Bech et al.(^{51})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High risk</td>
<td>Low risk</td>
<td>Discrepancies in n’s reported in text and in outcome tables. No explanations given for attrition. More drop out and incomplete data for intervention no. 1 than intervention no. 2 or control.</td>
</tr>
<tr>
<td>Soares et al.(^{57})</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>None</td>
</tr>
<tr>
<td>Morrison et al.(^{49})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High risk</td>
<td>High risk</td>
<td>Unclear</td>
<td>High risk</td>
<td>High risk</td>
<td>Four women in intervention arm had vaginal bleeding, one dropped out due to severe breast tenderness. One drop out from control arm to seek depression treatment. Reported subgroup analysis that was not predetermined (subjects with history of major depression).</td>
</tr>
<tr>
<td>Yaffe et al.(^{47})</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>None</td>
</tr>
<tr>
<td>Pefanco et al.(^{55})</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>None</td>
</tr>
<tr>
<td>Marinho et al.(^{54})</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High risk</td>
<td>Low risk</td>
<td>Unclear how many women stopped estrogen within a short time (3 months) before trial.</td>
</tr>
<tr>
<td>Newhouse et al.(^{60})</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>None</td>
</tr>
</tbody>
</table>

“Low Risk” = low risk of bias; “High Risk” = high risk of bias; “Unclear” = uncertain risk of bias.
Qualitative analysis indicated improvement in vasomotor symptoms with estradiol, an expected result consistent with previous findings. The potential for at most modest benefits of treatment with estradiol for depressive symptoms must be weighed against the finding in three trials of an increased incidence of abnormal vaginal bleeding in treated subjects compared to controls, and uncertainty regarding the effect on other adverse outcomes due to the short duration of many trials and limited reporting on harms.

Overall completeness and applicability of evidence

We found a paucity of studies evaluating forms of bioidentical estrogen other than estradiol, including only one low quality trial of estrone. We applied strict inclusion criteria to reduce heterogeneity and maximize methodologic quality, which may have excluded some studies addressing our primary research question. The outcomes data from included studies were sufficient to allow us to draw conclusions regarding the effects of estradiol, but not bioidentical estrogens generally, on depressive symptoms. Only three studies evaluated outcomes beyond 16 weeks, providing limited evidence regarding longer durations of treatment that are likely to occur in clinical practice. The findings of this review are applicable to treatment of peri- and postmenopausal women who have either no depressive symptoms, or mild to moderate depressive symptoms. Our findings apply to treatment with estradiol delivered orally or transdermally in doses of up to 2 mg per day.

Quality of the evidence

Five of the studies included in the meta-analysis had no identified methodological quality concerns. The other five had issues related to missing or unevenly reported outcomes data. Reported completeness of follow-up ranged widely among the included trials, from 57% to 100%, and was not reported by one study. We found no evidence of publication bias. We did find evidence of reporting bias in two studies.

Heterogeneity among included studies limited the strength of our conclusions with respect to all pooled outcomes data, and to all subgroups analyzed other than the subgroup treated with adjunctive progestogens. Variability in the dose and route of delivery of the treatments used was possibly an important source of heterogeneity. One trial that used a relatively high dose of unopposed transdermal estradiol in perimenopausal women with moderate depressive symptoms showed a much stronger benefit than other studies included in the review. This outlier introduced considerable heterogeneity into the

---

**FIG. 2.** Effect of bioidentical estrogen on depressive symptoms as compared with placebo.

**FIG. 3.** Subgroup analysis of the effect of estradiol on depressive symptoms as compared with placebo, in perimenopausal women versus postmenopausal or mixed menopausal status women.
mix of included studies, but the methodological quality of this study was among the highest of all included trials, and because of the relatively small number of total trials and substantial differences in study samples and intervention between trials, we felt it was important to include it in the primary analysis.

One potential source of bias in this review is the definition of “bioidentical” that we used. Consensus on the definition of bioidentical is elusive. We chose a conservative definition that assured that the included estrogens would be broadly accepted as bioidentical. However, this decision may have introduced bias by leading us to exclude trials of estrogens such as estradiol valerate that may be functionally bioidentical. Another source of bias may have been the inclusion of studies that used nonbioidentical forms of adjunctive progestogens. We elected to include trials that utilized any form of progestogen in combination with bioidentical estrogen, and to perform a subgroup analysis comparing trials with and without adjunctive progestogens. Clinicians who advocate treatment with bioidentical estrogens are also likely to favor the use of bioidentical progesterone over a nonbioidentical progestogen; results for women treated with combination hormone therapy may have less relevance for these clinicians and their patients. Only one trial in which adjunctive bioidentical progesterone was used met inclusion criteria for this review. Finally, the failure of half of the included studies to exclude or document the use of antidepressants in enrolled subjects was a significant barrier to addressing a possible important confounder.

Agreements and disagreements with other studies or reviews

This is the first systematic review of bioidentical hormone replacement therapy. Our results conflict with an older systematic review that included studies of both bioidentical and nonbioidentical estrogens and found a benefit for treatment of depressive symptoms. However, our results are consistent with reports from the WHI, which showed no improvement in women’s depressive symptoms for CEE combined with medroxyprogesterone.

Conclusions

Implications for practice

Bioidentical estradiol appears to offer no clinically meaningful benefit for improvement of depressive symptoms in postmenopausal women. For treatment of mild to moderate depressive symptoms in perimenopausal women, the available data do not allow us to draw conclusions, but the possibility of a clinically meaningful effect of estradiol cannot be excluded. The effectiveness of bioidentical estrogens other than estradiol for depressive symptoms in peri- and postmenopausal women cannot be determined from currently available literature. For women complaining of vasomotor symptoms, the evidence suggests that bioidentical estradiol is an effective treatment option.

We found no suggestion of risk for bioidentical estrogen over what might be expected with nonbioidentical estrogen therapy, but most included studies were of short duration and serious adverse events were unlikely to be captured. In the absence of more definitive evidence, it is most reasonable to counsel women that all of the long-term health risks associated with CEE may also be associated with bioidentical estrogens. Women considering treatment with bioidentical estrogens should also be informed that treatment appears to be associated with increased incidence of irregular vaginal bleeding. It is unknown whether use of bioidentical progesterone as an adjunct to bioidentical estradiol is safer or more effective than adjunctive use of a nonbioidentical progestogen.

Implications for research

From a practical standpoint, perimenopausal and early postmenopausal women with active vasomotor symptoms are the most likely group for whom estradiol would be considered as a favorable alternative for treating depressive symptoms, and three studies included in our review suggest a potential benefit in this clinical scenario. Additional research is needed to evaluate a possible benefit of bioidentical estrogen in perimenopausal women with mild to moderate depressive symptoms at baseline. Based on currently
available evidence, the most promising treatment for further study is unopposed estradiol. The most appropriate treatment dose is uncertain, but higher doses, in the range of 50–100 μg transdermal estradiol or 1–2 mg oral estradiol daily may be more effective. A better understanding of the benefits and harms of bioidentical estrogens would allow women to more accurately weigh the risks and benefits of bioidentical hormone therapy as an alternative to no treatment or to standard pharmacologic and nonpharmacologic therapies for mild to moderate depressive symptoms.

Additional research is also needed to compare bioidentical estrogen with nonbioidentical estrogen for treatment of depressive symptoms. This review was not designed to determine whether bioidentical estrogens are superior to nonbioidentical estrogens for treatment of depressive symptoms or other symptoms experienced by menopausal women, nor does it address the safety or efficacy of individually compounded forms of bioidentical estrogen. Randomized trials comparing different types of estrogen therapy for short-term treatment in perimenopause could provide valuable information for women and practitioners making treatment decisions regarding hormone therapy.

**Author Disclosure Statement**

No competing financial interests exist.

**References**


Address correspondence to:
James M. Whedon, DC, MS
University Health System
Southern California University of Health Sciences
16200 Amber Valley Drive
Whittier, CA 90604

E-mail: jameswhedon@scuhs.edu