Supplementary Online Content

Gordon JL, Rubinow DR, Eisenlohr-Moul TA, Xia K, Schmidt PJ, Girdler SS. Efficacy of transdermal estradiol and micronized progesterone in the prevention of depressive symptoms in the menopause transition: a randomized clinical trial. *JAMA Psych*. Published online January 10, 2018. doi:10.1001/jamapsychiatry.2017.3998

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

Participants

Women without a uterus but with at least one ovary retained were included in the study if: 1) they were experiencing vasomotor symptoms and their baseline estradiol levels were above postmenopausal concentrations (> 40 pg/ml) (n =15) or, 2) they were not experiencing vasomotor symptoms but had baseline estradiol > 40 pg/ml and baseline FSH levels > 14 pg/ml (n=2), which is two standard deviations above the mean level of a previously-recruited sample of reproductive-aged women, in line with STRAW staging guidelines[1]. However, these women did not take intermittent micronized progesterone and were not included in the reproductive stage moderation analyses. Women with both ovaries removed were included if the bilateral oophorectomy occurred in the past 24 months and they had been menstruating regularly prior to the procedure (n = 1). These women were not included in reproductive stage moderation analyses.

Exclusion criteria included current major depressive disorder or a trauma- or stress-related disorder with severity greater than mild, a history of severe substance use within the past 10 years, a history of suicide attempts, bipolar or other psychotic disorders, current use of psychotropic medication, hormonal preparations, or herbal compounds indicated for menopausal symptoms (e.g. Black Cohosh) or mood (e.g. St. John's Wort), use of statins or antihypertensive agents other than diuretics, blood pressure >140/90 mmHg, fasting LDL ≥ 190 mg/dl (women with borderline elevated LDL cholesterol (190-200 mg/dl) at baseline were allowed into the trial if their physician was not intending to treat them over the course of the subsequent 12 months) or fasting glucose ≥ 7.0 mmol/l (126 mg/dl), use of anti-inflammatory agents (>10 times/month) unless they were on a stable regimen, endometrial hyperplasia, abnormal uterine anatomy, history of thrombophlebitis or thromboembolic disorders, history of estrogen-dependent neoplasias, gall bladder disease, liver dysfunction or other disorders for which estrogen or progesterone use is contraindicated, any history of any cardiovascular disease, arteriosclerosis, heart attack or stroke, atrial fibrillation, frequent premature atrial or ventricular beats, or other rhythm abnormalities, diabetes (Type I or II), body mass index > 35, smoking >10 cigarettes per day, any history of migraine in current cigarette smokers, known sensitivities to any ingredient in the transdermal estradiol system. To be eligible for the study, women also must have had a normal mammogram within one year of study enrollment. All women underwent a pelvic exam with a study physician to screen for any signs or history of endometrial disorder or abnormal uterine or ovarian anatomy. An endometrial biopsy was performed to rule out endometrial cancer in cases of concerning bleeding patterns. To control as much as possible for extraneous factors that vary with depression and impact cardiovascular risk (e.g. diet and exercise), all participants were given information on the American Heart Association DASH diet and exercise guidelines and were encouraged to adhere to them throughout the study.

The following exclusion criteria were applied to minimize the risk of breast or ovarian cancer: a personal history of breast cancer; a personal history of even one breast biopsy with atypical hyperplasia, though women with a history of > 1 breast biopsy were allowed to enter the trial if there was no personal or family history of breast cancer in first or second degree relatives and if the documented biopsy results are consistent with minimal proliferative disease, as

described by Vogel et al., 2008[2]; more than one first degree relative with breast cancer; premenopausal breast cancer in even one first degree relative; more than three first, second, or third-degree relatives with breast cancer regardless of age; two or more first-degree relatives with any cancer with onset before age 60 (except tobacco-related lung cancer since it is associated with a low-penetrance gene); multiple primary cancers in a single relative (except tobacco-related lung cancer since it is associated with a low-penetrance gene); any male breast cancer (as it is almost exclusively related to the BRCA 1/2 mutation); ovarian cancer in even one first-degree relative since that is associated with a 25% chance of having the BRCA 1 mutation (Schwartz et al., 2008); any known BRCA mutation in first, second or third-degree relatives, unless the woman had tested negative for the BRCA mutation; a personal history of irradiation to the breast or chest wall prior to the age of 30, such as for Hodgkin disease; Ashkenazi Jewish descent (those tracing their roots to central and eastern Europe) since two BRCA1 and one BRCA2 mutations are observed with higher frequency in Ashkenazi Jews, unless there was no family history of cancer.

Modifications to eligibility criteria initiated following trial commencement: Due to challenges with participant recruitment, several changes were made to the originally-proposed eligibility criteria to facilitate recruitment, resulting in the above-described criteria. These included the following changes: early postmenopausal (in STRAW +1b) were initially excluded but later allowed to enter the trial; the acceptable age range was extended from 45-55 to 45-60; an initial requirement that all women have an FSH level >2 SD above mean premenopausal levels was removed; women without a uterus or having undergone an endometrial ablation were initially excluded but later allowed to participate if they met the above-mentioned requirements related to estradiol and/or FSH levels; having had a bilateral oophorectomy was initially exclusionary but was later allowed if it had occurred within 24 months and the participant had been menstruating regularly prior to the procedure; an initial requirement of having a baseline CES-D score ≤8 in the women without a history of depression was removed; women with borderline elevated LDL cholesterol (190-200 mg/dl) were initially ineligible. More details about amendments made to participant eligibility criteria can be found in the Trial Protocol.

Details about the Intervention

The placebo patches generated by 3M Pharmaceuticals were not identical to the Climara® TE+IMP system in that the active Climara® patches was labeled with the brand and dosage of the patch, which was missing from the placebo patch. In an effort to maintain the double-blind despite this difference in appearance, 2/3 of active patches were the generic version of Climara® and study personnel were instructed that multiple forms of both the placebo and active patches existed. All placebo patches were 3M placebo patches. The patch that any particular woman received remained identical throughout her involvement. Women randomized to receive active TE+IMP wore a patch to deliver 0.025 mg of estradiol over a 24-hour period for the first two weeks, 0.05 mg for the next four weeks and 0.10 mg of estradiol for the remainder of the study. Though 0.10 mg was the target dose, five women were maintained on 0.05 mg and three were maintained at 0.025 mg to mitigate side effects. To monitor compliance, participants' used patches were collected and counted at each study visit. Plasma estradiol levels were also measured at months 6 and 12.

The study statistician created the randomization scheme using a 1:1 ratio of permuted blocks of size 4 by two levels of depression history, with the maximum imbalance of size 2 and

the UNC Hospitals Investigational Drug Services managed the randomization and dispensing of all study medication in blinded form. The safety of the study was monitored by a Data and Safety Monitoring Board. All study personnel with any direct or indirect contact with participants, including physicians and research assistants, remained blinded throughout the trial to the best of our abilities. Because one of the PERT study's primary aims was to determine whether a positive history of depression would predict the mood benefits of TE+IMP, randomization was stratified based on histories of depression (yes, no) and, in the past depression group, stratified based on one versus more than one episode of depression.

Sample Size

Based on the work of Freeman et al.,[3] we anticipated the rates of clinically significant depressive symptom occurrence, operationalized as CES-D scores ≥ 16[4,5], in untreated women would be 37.5% and that TE+IMP would reduce this rate by 50% to a rate of 18.75%. Using these rates and anticipated effect size, sample size calculations indicated that 183 trial completers would be needed to achieve 80% power to detect a significant effect of treatment on the rate of clinically significant depressive symptoms (CES-D score ≥ 16), the current study's primary outcome, using logistic regression and a two-tailed significance level of 0.05. This target was not reached because substantially more perimenopausal women did not meet criteria for being medically healthy (e.g. obesity, diabetes, hypertension) than we had anticipated and thus were not eligible for TE+IMP.

Stressful Life Events Assessed in the Life Experiences Survey[6].

1. Got married or engaged
2. Got separated, divorced or had a break-up of long-term committed relationship
3. Major change in closeness to a family member (e.g. estrangement)
4. Death of close family member
5. Death of a close friend
6. Serious illness of close family member
7. Serious illness of a very close friend
8. Trouble with your employer (e.g. in danger of losing your job)
9. Loss of your job
10. Trouble finding employment
11. Worked 60+ hours/week for at least 1 month
12. Major worsening of your financial status or major chronic financial problems
13. Major illness, chronic health problem or injury
14. You were in a motor vehicle accident
15. Hospitalized overnight for illness or injury
16. Physically attacked or assaulted or had your life threatened
17. Problems feeling safe in your neighbourhood
18. Arrested for a serious crime
19. Convicted of a crime and went to jail

20. Mate or close relative was arrested for a serious crime or went to jail
21. You were robbed or your home was burglarized
22. Your partner became pregnant, had or adopted a baby
23. Ending of your formal schooling
24. You left home for the first time
25. Son or daughter moved away from home for the first time
26. Moved residence more than once in the past 6 months

References

1. Harlow SD, Gass M, Hall JE, et al. Executive summary of the Stages of Reproductive Aging Workshop+ 10: addressing the unfinished agenda of staging reproductive aging. *Climacteric.* 2012;15(2):105-114.
2. Vogel VG. Epidemiology, genetics, and risk evaluation of postmenopausal women at risk of breast cancer. *Menopause.* 2008;15(4):782-789.
3. Freeman EW, Sammel MD, Lin H. Temporal associations of hot flashes and depression in the transition to menopause. *Menopause (New York, NY).* 2009;16(4):728.
4. Boyd JH, Weissman MM, Thompson WD, Myers JK. Screening for depression in community sample. *Arch Gen Psychiatry.* 1982;39:1195-1200.
5. Thomas JL, Jones GN, Scarinci IC, Mehan DJ, Brantley PJ. The utility of the CES-D as a depression screening measure among low-income women attending primary care clinics. *International journal of psychiatry in medicine.* 2001;31(1):25-40.
6. Sarason IG, Johnson JH, Siegel JM. Assessing the impact of life changes: development of the Life Experiences Survey. *Journal of consulting and clinical psychology.* 1978;46(5):932.

eResults

Baseline Stressful Life Events

60% of our sample reported at least one stressful life event within six months of study entry. The most common stressful life events reported included: death or serious illness of a close family member or close friend (30%), major financial difficulties (16%), working at least 60 hours/week for at least 1 month (12%), job loss or unemployment (12%), estrangement or serious arguments with a family member (9%), son or daughter moving away from home for the first time (6%), divorce or separation (5%) and major illness (4%).

Treatment Adherence

Information regarding visit attendance can be found in Figure 1. Treatment adherence was good, with only 12 (7%) of participants being flagged as potentially poor compliers by the research staff based on the number and appearance of the used returned patches. With regards to treatment blinding, 63% and 39% of participants in the TE+IMP and placebo conditions were accurate at every study visit in guessing their treatment assignment, respectively. The remaining participants were either consistently inaccurate (12 vs. 35%), were unsure (3 vs. 6%) or were inconsistent in their beliefs (22% vs. 20%). The accuracy rates of the two treatment groups (63 vs. 39%) were found to be both significantly different from each other ($p < .01$) and significantly different than chance levels in chi-square analyses (ps < .05). For this reason, we performed additional sensitivity analyses to ensure that participant unblinding did not account for our results.

Effect of Treatment on Estradiol Levels

Paired t-tests comparing estradiol levels pre- and post-randomization (averaged across visits 6 and 12) revealed that the TE+IMP group ($t(60) = 4.9$, $p<.0001$) but not the placebo group ($t(75) = -1.1$, $p = .27$) experienced a significant increase in estradiol levels such that the TE+IMP group had significantly higher estradiol levels in the treatment phase compared to the placebo group (*M (SD)* = 148.6 (77.0) versus 100.0 (57.3) pg/ml; $p<.0001$). Women assigned to TE+IMP experienced a significantly greater decrease in vasomotor symptom bother over the course of the study compared to women assigned to placebo (*M (SD)* = -1.7(1.8) vs. -0.8(1.5), $p < .01$).

Effect of Baseline Characteristics on Depressive Symptoms across Both Treatment Groups

Although Table 2 indicates that none of the treatment moderators were predictive of depressive symptoms in statistical models that included treatment assignment and a treatment-by-moderator interaction term, several of the moderators were significant predictors of mean CES-D score when only the moderator variable and baseline CES-D score were included as independent variables. Baseline stressful life events ($\beta(SE) = 1.0(0.2)$, $p < .0001$), a history of depression ($\beta(SE) = 1.54(0.5)$, $p < .01$) and baseline vasomotor symptom bother ($\beta(SE) = 0.34(0.1)$, $p = .02$) predicted a higher mean CES-D score. Reproductive history was also a significant predictor of mean CES-D score ($p = .01$) such that women in the early postmenopausal period (*M(SE)* = 3.9(0.6)) had significantly lower scores compared to early perimenopausal (*M(SE)* = 6.1(0.6); $p <.01$) or late perimenopausal (*M(SE)* = 5.5(0.4), $p = .01$) women. Neither baseline estradiol ($p = .81$) nor abuse history ($p = .13$) significantly predicted depressive symptoms.

eFigure. Model-Based Estimates of Treatment Effect on CES-D Score Over Time Across Reproductive 181 Stages.

A significant effect of treatment is evident among the early perimenopausal women ($p < .001$) 182 but not the late perimenopausal women ($p = .23$) or early postmenopausal women ($p = .92$). TE+IMP: 183 transdermal estradiol plus intermittent micronized progesterone