Longitudinal changes in menopausal symptoms comparing women randomized to low-dose oral conjugated estrogens or transdermal estradiol plus micronized progesterone versus placebo: the Kronos Early Estrogen Prevention Study

Nanette Santoro, University of Colorado
Amanda Allshouse, University of Colorado
Genevieve Neal-Perry, Albert Einstein College of Medicine
Lubna Pal, Yale University
Rogerio A. Lobo, Columbia University
Frederick Naftolin, New York University
Dennis M. Black, University of California San Francisco
Eliot A. Brinton, Utah Foundation for Biomedical Research
Matthew J. Budoff, University of California Los Angeles
Marcelle I. Cedars, University of California at San Francisco

Only first 10 authors above; see publication for full author list.

Journal Title: Menopause
Volume: Volume 24, Number 3
Publisher: Lippincott, Williams & Wilkins | 2017-03-01, Pages 238-246
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1097/GME.0000000000000756
Permanent URL: https://pid.emory.edu/ark:/25593/s8bbr

Final published version: http://dx.doi.org/10.1097/GME.0000000000000756

Copyright information:

Accessed December 18, 2018 6:02 AM EST
Longitudinal Changes in Menopausal Symptoms Comparing Women Randomized to Low-Dose Oral Conjugated Estrogens or Transdermal Estradiol Plus Micronized Progesterone Versus Placebo: the Kronos Early Estrogen Prevention Study (KEEPS)

Nanette Santoro, MD, Amanda Allshouse, MS, Genevieve Neal-Perry, MD, PhD, Lubna Pal, MD, Rogerio A. Lobo, MD, Frederick Naftolin, MD, PhD, Dennis M. Black, PhD, Eliot A. Brinton, MD, Matthew J. Budoff, MD, Marcelle I. Cedars, MD, N. Maritza Dowling, PhD, Mary Dunn, RN, Carey E. Gleason, PhD, Howard N. Hodis, MD, Barbara Isaac, RN, Maureen Magnani, RN, JoAnn E. Manson, MD, DrPH, Virginia M. Miller, PhD, MS, Corresponding Author (to whom reprint requests should be addressed): Nanette Santoro, MD, Department of Obstetrics & Gynecology, University of Colorado School of Medicine, 12631 E 17th Avenue, Mail Stop B-198, Aurora, CO 80045, 303-724-2041, 303-724-2061 (FAX), Nanette.Santoro@ucdenver.edu.

Conflict of Interest disclosures:
Nanette Santoro, MD, reports investigator initiated grant support from Bayer, Inc and stock options in Menogenix, outside the submitted work
Dennis Black, PhD, reports grant and personal fees from Novartis, personal fees from Merck, Amgen, and Eli Lilly, outside the submitted work
Eliot Brinton, MD, reports personal fees from Alexion, Amarin, Amgen, Arazel, Janssen, Kowa, Merck, Regeneron, Sanofi Aventis and Takeda
S Mitchell Harman, MD, reports grants from The Aurora Foundation and Pfizer Pharmaceuticals, non-financial support from Abbott Laboratories and non-financial support from Bayer Healthcare during the conduct of the KEEPS study
Lubna Pal, MD, MSc, reports personal fees from Merck
Roger Lobo M.D. reports consultation fees from Pfizer, Amigen, Teva and grant support from TherapeuticsMD.
Hugh Taylor, MD, PhD, reports grant support from Pfizer through Yale University and personal fees from Pfizer
Elin Wolfa MD, PhD, reports being an employee of Celmatix, Inc.
Other coauthors report no disclosures.

KEEPS Investigators and Staff
Albert Einstein College of Medicine: Genevieve Neal-Perry, Ruth Freeman, Hussein Amin, Barbara Isaac, Maureen Magnani, Rachel Wildman
Brigham and Women’s Hospital/Harvard Medical School: JoAnn E. Manson, Maria Bueche, Marie Gerhard-Herman, Kate Kalan, Jan Lieson, Kathryn M. Rexrode, Barbara Richmond, Frank Rybicki, Brian Walsh
Columbia College of Physicians and Surgeons: Rogerio A. Lobo, Luz Sanabria, Maria Soto, Michelle P. Warren, Ralf C. Zimmerman
Kronos Longevity Research Institute: S. Mitchell Harman, Mary Dunn, Panayiotis D. Tsitouras, Viola Zepeda
Mayo Clinic: Virginia M. Miller, Philip A. Araoz, Rebecca Beck, Dalene Bott-Kitslaar, Sharon L. Mulvagh, Lynne T. Shuster, Teresa G. Zais (deceased)
University of California, Los Angeles, CAC Reading Center: Matthew J. Budoff, Chris Dailing, Yanlin Gao, Angel Solano
University of California, San Francisco, Medical Center: Marcelle I. Cedars, Nancy Jancar, Jean Perry, Rebecca S. Wong, Robyn Pearl, Judy Yee, Brett Elicker, Gretchen A. W. Gooding
University of California, San Francisco, Statistical Center: Dennis M. Black, Eric Vittinghoff, Lisa Palermo
University of Southern California, Atherosclerosis Research Unit/Core Imaging and Reading Center: Howard N. Hodis, Yanjie Li, Mingzhu Yan
University of Utah School of Medicine: Eliot A. Brinton, Paul N. Hopkins, M. Nazeem Nanjee, Kirtly Jones, Timothy Beals, Stacey Larrinaga-Shum
Veterans Affairs Puget Sound Health Care System and University of Washington School of Medicine: George R. Merriam, Pamela Asberry, Sue Ann Brickle, Colleen Carney, Molly Carr, Monica Kletke, Lynna C. Smith
Yale University School of Medicine: Hugh S. Taylor, Kathryn Czarkowski, Lubna Pal, Linda McDonald, Mary Jane Minkin, Diane Wall, Erin Wolf*
Other: Frederick Naftolin (New York University), Nanette Santoro (University of Colorado)
*Now at the National Institutes of Health/Eunice Kennedy Shriver National Institute of Child Health and Human Development
Abstract

**Objective**—To compare the efficacy of two forms of menopausal hormone therapy in alleviating vasomotor symptoms, insomnia, and irritability in early menopausal women over four years.

**Methods**—727 women, aged 42–58, within three years of their final menstrual period were randomized to receive oral conjugated estrogens (o-CEE) 0.45 mg (n=230) or transdermal estradiol (t-E2) 50mcg (n=225; both with micronized progesterone 200mg for 12 days each month), or placebos (PBO; n=275). Menopausal symptoms were recorded at screening and at 6, 12, 24, 36 and 48 months post-randomization. Differences in proportions of women with symptoms at baseline and at each followup timepoint were compared by treatment arm using exact chi-square tests in an intent-to-treat (ITT) analysis. Differences in treatment effect by race/ethnicity and body mass index (BMI) were tested using generalized linear mixed effects modeling.
Results—Moderate-to-severe hot flashes (from 44% at baseline to 28.3% for PBO, 7.4% for t-E2 and 4.2% for o-CEE) and night sweats (from 35% at baseline to 19% for PBO, 5.3% for t-E2 and 4.7% for o-CEE) were reduced significantly by 6 months in women randomized to either active hormone compared to PBO (P<0.001 for both symptoms), with no significant differences between the active treatment arms. Insomnia and irritability decreased from baseline to 6 months post randomization in all groups. There was an intermittent reduction in insomnia in both active treatment arms vs PBO, with o-CEE more effective than PBO at 36 and 48 months (p=0.002 and 0.05) and t-E2 more effective than PBO at 48 months (p=0.004). Neither hormone treatment significantly affected irritability compared to PBO. Symptom relief for active treatment vs PBO was not significantly modified by BMI or race/ethnicity.

Conclusions—Recently-menopausal women had similar and substantial reductions in hot flashes and night sweats with lower than conventional doses of oral or transdermal estrogen. These reductions were sustained over 4 years. Insomnia was intermittently reduced compared to placebo for both hormone regimens.

Keywords
- hormone therapy; menopause; hot flashes; night sweats; insomnia; irritability

Introduction

Menopausal symptoms are experienced by a majority of women. While women in the perimenopausal age range of 40–55 report many different types of symptoms, not all are necessarily linked to ovarian senescence and the loss of estradiol and progesterone production. It remains challenging to disentangle age-related symptoms from those attributable to, or interacting with, menopause. Vasomotor symptoms (VMS), or hot flashes and sweats, are in many ways the quintessential menopausal symptom. Up to 85% of women report onset or significant increases VMS at menopause. There is little disagreement about the linkage of VMS to ovarian hormone, and particularly, estrogen, production. Menopausal hormone therapy (HT) is highly effective for the treatment of hot flashes. HT treatment is most commonly given for just a few years flanking the final menstrual period (FMP), although some women with prolonged symptoms require longer courses of therapy. In particular, recent studies indicate that hot flashes may persist for 10 years or more after the FMP. While non-hormonal prescription drugs are known to have modest effectiveness against VMS, they remain inferior to estrogen in the magnitude of benefit, and there is only one FDA-approved alternative to estrogen for the treatment of hot flashes (paroxetine mesylate). Fewer studies have examined the role of hormones in the treatment of the less prevalent menopausal symptoms such as depressed mood and disturbed sleep and fewer still have compared HT regimens, especially over extended durations, to determine which are most effective for particular outcomes or symptoms.

The Kronos Early Estrogen Prevention Study (KEEPS) was a multicenter clinical trial designed to compare effects of low-dose oral conjugated equine estrogens (o-CEE) to transdermal estradiol (t-E2) versus placebo (PBO) on cardiovascular end points in recently-menopausal women. Herein, we report a comparison of self-reported symptoms over time in KEEPS participants randomized to o-CEE, t-E2, or PBO in whom the prevalence of
vasomotor, mood, and sleep symptoms were assessed and the presence and severity of symptoms were compared across treatments over four years of trial duration. We hypothesized that menopausal symptoms would be prevalent in the KEEPS cohort at baseline, that both hormone regimens would be more effective in alleviating symptoms compared to PBO, and that symptoms would subside over time in all groups. We also hypothesized that baseline demographics and anthropometric measures would influence prevalence and severity of symptoms at baseline and their response to treatment.

Methods

KEEPS enrolled 727 women ages 42–58 years who were ≥6 months but <36 months from their last menses with an FSH level ≥5 ng/ml and/or estradiol (E$_2$) <40 pg/ml. Age at menopause was determined at phone screening and verified at the baseline visit. Women were randomized to either: o-CEE 0.45mg daily (n=230) or t-E2 50mcg daily (n=225), both with oral micronized progesterone 200mg daily for 12 days each month and both with a placebo (PBO) for the treatment not given versus a control arm receiving triple PBO (patch and pills) group (n=275). Nine recruitment sites from across the USA participated in the KEEPS trial (ClinicalTrials.gov; trial number NCT00154180). Details on the enrolled population have been reported. All women had a benign Pap smear and a normal mammogram within one year prior to randomization. Past or current users of HT were screened only after at least a 90-day washout period. Women with a history of clinical cardiovascular disease (CVD), including myocardial infarction (MI), angina, congestive heart failure, stroke, transient ischemic attacks, or thromboembolic disease were excluded, as were those who reported smoking more than 10 cigarettes per day. All women meeting initial eligibility criteria had a complete blood count and chemistry panel measured at the clinical laboratories at each study center. Lipid profiles and TSH were measured at the Kronos Science Laboratories (Phoenix, AZ). Women were also screened for coronary artery calcium (CAC) and those with scores ≥50 AU were excluded.

A planned secondary analysis of the KEEPS Study included a central evaluation of circulating estrone (E$_1$) and estradiol (E$_2$) in a subset of participants at baseline and on treatment to determine compliance with the treatment regimen and to assess the relationship of circulating hormone levels on symptom relief. For these analyses, E$_1$ and E$_2$ were measured on a randomly selected subsample of women from all three treatment arms by the Reproductive Endocrine Research laboratory of the USC Keck School of Medicine (Los Angeles, CA) using a highly sensitive, well-validated radioimmunometric method following extraction of 0.8ml of serum with ethyl acetate/hexane (3:2) followed by Celite column partition chromatography. Recovery of tritiated E$_2$ (range 73–86%) was used to correct observed values. The limit of detection for E$_2$ was 2 pg/ml, respectively. Interassay and intra-assay CVs for E$_2$ were 6% and 4%, respectively.

Short follow-up visits were scheduled every 90 days (in person or by phone with medications mailed) to assess adverse events and adherence (pill/patch counts). Longer in-person visits were conducted at 12, 18, 36 and 48 months to measure cardiovascular or cognitive end-points.
All participants completed a menopausal symptom checklist prior to randomization and again at 6, 12, 24, 36 and 48 months. Menopausal symptoms included in the present analysis were self-assessed and included only current symptoms: hot flashes, night sweats, insomnia, and irritability. Symptoms were scored on a 4-point ordinal scale: 0 (no symptoms) to 1 (mild), 2 (moderate) and 3 (severe). “Symptomatic” for each menopausal symptom was dichotomously defined as moderate/severe (vs none/mild) for most analyses; for the comparison of symptom severity to circulating estradiol, 4-category symptom severity score was used (see below).

History of tobacco use was captured based on a combination of a screening question, which screened out women who smoked more than 10 cigarettes per day, and a Tobacco Use Form designed to capture lifetime pack-years of smoking, introduced at the study’s mid-point. Diet and physical activity were assessed at the first clinic screening visit. Additional validated instruments were administered to assess mood and cognition; a standardized, comprehensive cognitive test battery was also performed as a substudy to the KEEPS called the KEEPS Cognitive and Affective Study (KEEPS-Cog); these results have been reported.

Physical measures (height, weight, waist and hip circumference, blood pressure) were obtained at all visits using standardized protocols. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters, squared. BMI was categorized using international classification definitions: ≥30 being obese, 25–<30 being overweight and <25 being normal (or possibly underweight). Race/ethnicity was self-reported. Analysis of genetic polymorphisms in the KEEPS sample confirmed women’s self reports in that the majority of participants were of white, Central European ancestry.

Data management

Data were entered at study centers into secure online forms in PERL (Practical Extraction and Report Language) and transferred to the database management system at the KRONOS Coordinating Center. Data were subsequently uploaded for analysis and converted to SAS datasets at the UCSF data coordinating center.

Data analysis

Data were analyzed based on original treatment assignment (intent-to-treat; ITT). All reported results are from the ITT analysis. Available data were used without imputation for missing values in the primary analysis.

Comparisons of baseline characteristics across the 3 treatment groups are reported as frequency and percent for categorical, or mean (SD) for continuous variables, with p-values from chi square (exact where possible), or analysis of variance (ANOVA), comparing across all 3 groups. In a logistic model of each symptom at baseline, 3-category BMI and 3-category race/ethnicity were included as covariates to compare overall symptom prevalence by race/ethnicity and BMI. Racial and ethnic groups compared included white, black and other (comprised of women reporting Asian, Hispanic, or any other ethnicity that could not be considered either white or black). Comparisons of symptoms by race and BMI are reported as odds ratios (OR) and 95% confidence intervals (CI).
The percentage women who were symptomatic at screening and at each follow-up time-point was assessed. To test if either hormone regimen was more effective than PBO in alleviating symptoms, the proportion of women who were symptomatic (reported moderate/severe symptoms) at each follow-up visit was summarized by treatment, with comparisons tested using logistic regression between each of the treated groups vs. placebo, and also between treatment groups (as an exploratory analysis).

In a secondary analysis, symptoms over time were modeled using two general linear mixed effects models for the logistic distribution with an interaction between (1) treatment arm and (2) 3-category race/ethnicity, and separately 3-category BMI.

The relationship between severity of vasomotor symptoms and circulating E2, concentrations were compared across 4-category vasomotor symptoms at baseline (reflecting the relationship between endogenous E2 and VMS) and 12 months (reflecting the relationship between t-E2 and VMS), with differences tested using ANOVA and pairwise comparisons within a General Linear Model (GLM). Since the E2 distribution was right-skewed, data were log-transformed for analysis and back-transformed for reporting of geometric mean concentrations with 95% CI. The linear relationship of log-scale E2 at month 12 was compared within a GLM by testing differences across treatment group in slope of E2.

As a sensitivity analysis, all women with missing symptom data were assigned the worst symptomatic category, and all analyses were performed again. A second sensitivity analysis was performed to examine whether a different cut point for symptoms (any versus none) changed the results.

A p-value <0.05 was used to determine statistical significance. SAS 9.4 was used to carry out the statistical analysis, and graphics were created using GraphPad Prism 6.

**Results**

Recruitment and flow of participants through the study have been previously reported\(^\text{16}\) and is shown in Supplemental Figure 1. Due to study dropout, there were fewer women available to report symptoms at each successive time point. Participant attrition from screening to 48 months was similar in all groups such that 173 of 230 (75.2%) women randomized to o-CEE, 170 of 222 (76.6%) women randomized to t-E2, and 211 of 275 (76.7%) women randomized to PBO, 211 completed the end-of-study assessment (Figure 1). Baseline characteristics of the KEEPS cohort and baseline symptom reporting of moderate-to-severe symptoms by randomization assignment are shown in Table 1.

**Vasomotor Symptoms: Hot Flashes/Night Sweats**

Unadjusted prevalence of hot flashes and night sweats are shown in Table 2 and Figure 2 for all time points. At screening, 86% of all participants reported at least mild hot flashes, while moderate-severe hot flashes were reported by 44%. By 6 months post-randomization, moderate-severe hot flashes had decreased to 28.3% of women randomized to PBO, 7.4% of women randomized to t-E2 and 4.2% of women randomized to o-CEE (p<0.001 for each
active treatment vs PBO). Night sweats were reported by 68% of women at screening, with 35% being moderate-severe. At 6 months, moderate-severe night sweats declined to 19% with PBO, 5.3% with t-E2 and 4.7% with o-CEE (p<0.0001 for each active treatment vs PBO). This initial magnitude of symptom reduction was maintained throughout the study in all treatment groups.

**Insomnia and Irritability (Table 2, Figure 2)**

At baseline, the proportion of women reporting insomnia did not differ between treatment groups (PBO 34%, o-CEE 29%, and t-E2 35%, p=0.3). Insomnia decreased substantially and comparably by 6 months in all groups and this decrease was maintained throughout the trial. At 36 and 48 months, o-CEE was significantly more effective in reducing insomnia vs PBO (p=0.002 and 0.05), and at 48 months t-E2 was more effective than PBO (p=0.004). Baseline reports of irritability were similar between treatment groups (PBO 15%, o-CEE 17%, and t-E2 19%, p=0.6) and decreased comparably by about half in all groups at 6 months, to 7.5%, 6.9% and 5.8%, respectively, and did not differ between treatment groups at any time point.

**Differences in treatment effect by demographics**

For each symptom, the relationship of race/ethnicity and BMI to treatment effect was calculated. Due to small numbers of women for some of the time points, a fully-interacted model could not be constructed for night sweats or irritability. The effects of o-CEE as well as t-E2 vs PBO on hot flashes and insomnia showed no significant interaction by BMI or race/ethnicity.

**Sensitivity Analyses**

Since severe symptoms might have caused differential loss of data due to selective visit avoidance or study dropout, which could result in artefactual symptom relief, we performed a sensitivity analysis, re-categorizing all missing responses as being due to the presence of moderate to severe symptoms. This analysis did not, however significantly alter any of the above treatment effects or lack thereof, compared to the above results. Changing the point at which the symptom variable was dichotomized to ‘any’ versus ‘none’ resulted in identical patterns of symptom relief among groups (Supplemental Figure 2).

**Relationship of Circulating E2 to Hot Flashes**

Serum E2 levels were available in a subset of 426 women at baseline, (135 o-CEE, 131 t-E2 and 158 PBO), and 424 women at 12 months (194 o-CEE, 181 t-E2 and 227 PBO). E2 concentrations at baseline did not significantly differ by symptom severity score categories for either hot flashes or night sweats. At 12 months, however, among women randomized to t-E2, circulating E2 was significantly higher among non-symptomatic women compared to women reporting moderate or severe symptoms. At 12 months, among women randomized to t-E2, those reporting no hot flashes had a geometric mean E2 of 44.26 pg/ml (95% CI 38.97, 50.27), significantly higher than women reporting moderate (9.12 pg/ml [95% CI 5.85,14.20], p=<0.001) or severe (11.04 pg/ml [95% CI 5.35, 22.77], p=0.01) hot flashes. Similar findings were observed for night sweats at month 12 (data not shown).
Screening Symptoms By Race/Ethnicity

Before randomization, reporting of symptoms differed significantly by race/ethnicity for every symptom, \((p=0.03\) for hot flashes, otherwise \(p<0.001\)) in a model including 3-category race and 3-category BMI. Pairwise differences by race/ethnicity revealed that the differences of greatest magnitude were between black and white women for all symptoms (hot flashes OR 1.74 [1.14, 2.64]; night sweats OR 3.59 [2.3 5.52]; insomnia OR 3.38 [2.20, 5.19]; irritability OR 19.23 [11.72, 31.57]). Comparisons between other vs white women and other vs black women revealed progressively smaller between group differences. In terms of absolute value, unadjusted differences at baseline were highest among black women relative to white women and women of other races for moderate to severe hot flashes (52%, 44%, 41%), night sweats (46%, 34, 35%), and especially moderate-to-severe irritability (33%, 15%, 21%), respectively. Baseline insomnia was reported by similar proportions of black women (35%), white women (32%), and women of other races (36%).

Discussion

This study is the first to compare menopausal symptoms longitudinally by treatment regimen and route of administration in women taking different types of low-dose estrogen therapy in combination with oral micronized progesterone. Overall, HT with either o-CEE or t-E2 was highly effective in relieving the more traditional menopausal symptoms of hot flashes and night sweats, with little difference in effectiveness between either of the two active treatment groups. There was pronounced reduction of moderate-severe symptoms, which typically drive women to seek treatment. Relief from symptoms with HT relative to PBO was maintained for the 48 month follow up period for each treatment arm, despite an overall decrease in hot flashes and night sweats in the PBO group over the course of the study. The clear-cut effect observed for both active treatments is consistent with the current and most accepted indication for HT, i.e., treatment of menopausal symptoms.

In contrast to favorable changes in mood with HT reported in the KEEPS-Cog study, which included detailed psychological and cognitive testing in the women in the KEEPS cohort\(^2^3\), we observed little effect of either of the two active treatments on irritability. Notably, however, a statistically significant treatment-related reduction in insomnia emerged towards the end of the follow-up period, which was largely due to an increase in insomnia after 12 months in the women randomized to PBO. HT-related decreases in hot flashes and night sweats persisted despite decreases in these symptoms in the PBO group over time.

There are several possible explanations for the overall longitudinal decrease in vasomotor symptoms with PBO. The first and most likely is that the timing of the KEEPS intervention coincides with the natural history of these menopausal symptoms, because menopausal symptoms tend to be worst within the year surrounding the final menses and then often spontaneously subside\(^2^6\). Vasomotor symptoms are worst within the year surrounding the final menses\(^2^6,2^7\). Since KEEPS participants were all at least 6 months but no more than 3 years past their final menses (22 months on average, similar in all 3 treatment arms), it is not surprising that vasomotor symptoms would tend to improve over time with PBO. The very large initial decrease in both hot flashes and night sweats from screening to 6 months,
however, may be beyond that from natural history alone and instead may reflect a substantial ‘placebo effect’ with initiation of study treatments.

These data on the effectiveness of hormone therapy for hot flashes and night sweats are in agreement with prior clinical trials in which other doses and formulations of HT were used. Both the WHI and HERS reported improvements in self-reported symptoms with hormone therapy at an o-CEE dose higher than what was administered in KEEPS. The present report underscores the concept that both oral and transdermal HT are very effective treatments for the common menopausal symptoms of hot flashes and night sweats, with some possible effectiveness for sleep complaints as well. It is also notable that o-CEE and t-E2 were remarkably similar in their ability to relieve symptoms. No other trials have compared these low-dose HT treatments to each other, especially over 4 years of follow up.

The prevalence of insomnia reported by KEEPS participants and its increase over time are similar to the 30.8% of women in a comparable population of midlife women from the Study of Women’s Health Across the Nation (SWAN), which observed an increased in this complaint over the menopausal transition to >40% during the late transition and postmenopause. The increase of moderate to severe symptoms of insomnia in the PBO group began to appear at 12 months, and led to a statistically significant improvement in insomnia for both active treatment arms at years 3 and 4. This finding may indicate a late effect of menopause on sleep efficiency that has not been previously appreciated, but the finding should be interpreted with caution since it was towards the end of the study when fewer participants provided data for the analysis. Poor sleep of various types, including insomnia, has been attributed to be a cardinal symptom of the menopause transition by some, but not other investigators. It is possible that the timing of this symptom is later after menopause than previously believed, and its relationship to a woman’s hormone status is more complex than a simple model of estrogen and progesterone withdrawal. It is also possible that the use of a non-validated survey to self-report insomnia lacked the sensitivity to detect between group differences earlier in the study.

Irritability declined in all groups, regardless of whether the women were treated with HT or placebo. In a prior analysis from the KEEPS-Cog study, o-CEE reduced scores for depression and anxiety, whereas t-E2 did not. Only a single aspect of mood, irritability, was measured and this may have been insufficiently sensitive to detect a treatment effect.

The favorable impact of HT is further underscored by the inverse correlation between circulating E2 levels and the prevalence of hot flashes in the t-E2 group, as well as an inverse relationship between self-reported adherence to treatments and menopausal symptoms in both active treatment arms. Lower symptom prevalence was associated with higher circulating E2 among women randomized to t-E2. Women demonstrated a wide range of E2 levels, from as low as 11 pg/ml in women reporting severe hot flashes to as high as 44 pg/ml in women reporting no hot flashes. It is possible that this variation is in past due to a lack of strictly controlled timing of blood draws in relation to patch application and pill ingestion. The only other study, to our knowledge, that has examined hot flash relief in relation to circulating E2 in women taking t-E2 was performed more than 3 decades ago, used a reservoir (not a matrix) transdermal E2 delivery system, and reported a therapeutic
range for serum E2 of 61 and 122 pg/ml for a 50% and 100% (theoretical) reduction in hot flashes\(^4\). This study used a radioimmunoassay without a chromatographic separation step for E2 determination, which lacks sensitivity at the low levels found in menopausal women\(^34\). Due to chemical differences between the treatments, E2 levels are much lower with o-CEE than with t-E2 but instead the majority of the estrogenic effect with o-CEE is due to circulating E1 and other B-ring equine estrogens (which were not measured). E1 can be converted to E2 to some degree so effective concentrations of E2 at the tissue level in the oCEE group may be of clinical importance.

Despite a rigorous trial design and large study sample, limitations of this study merit consideration. Although dropout was comparable with other trials of hormone therapy, it was still substantial. This trial was not powered to assess superiority of one route of estrogen therapy over another. In addition, the frequency of menopausal symptoms was not queried, but instead only their severity. This lack of data may have obscured differences between active treatment and placebo. Also, questionnaires only asked about current symptoms, and did not require participants to recall symptoms over the prior weeks or months. Although reporting of current symptoms tends to be more valid than recall of prior symptoms, it will underestimate prevalence of symptoms that fluctuate greatly over time. KEEPS did not use lengthier, well-validated, multidimensional measures of the symptoms under study because it was necessary to balance participant burden against the information to be gained. Study of symptom relief was a planned secondary end point of the KEEPS. Nonetheless, the use of a subjective severity measure for hot flashes and other symptoms is validated by other clinical trials\(^7,33,35,36\). Given that the KEEPS population is predominantly non-Hispanic and white, and were generally healthier and better educated than the average US woman, this may influence the generalizability of the findings and limit the ability to reliably assess ethnic and racial differences in symptom relief from hormone therapy among non-white women. Finally, the use of multiple comparisons in this study could have led to positive findings due to chance.

In addition, using a sensitivity analysis that imputed moderate/severe symptoms in all of the participants who dropped out, the relationships observed between hormone therapy and symptom relief remained remarkably similar, making differential dropout of the most symptomatic women (who would then have presumably gone onto HT in a setting outside of the study) an unlikely explanation for these findings.

**Conclusions**

In summary, there was significant relief of the menopausal symptoms of hot flashes, night sweats, and self-reported insomnia for 4 years and it was comparable between oral conjugated estrogens (o-CEE) 0.45mg daily and transdermal E2 (t-E2) 50mcg/day, combined with cyclic, oral micronized progesterone. Other symptoms, such as irritability and insomnia, were less influenced by HT.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.
Acknowledgments

The authors thank the investigators and staff at the KEEPS clinical centers, the KEEPS Data Coordinating Center at the Kronos Longevity Research Institute, and the National Institutes of Health institutions supporting ancillary studies, listed below. They also thank the participants for their dedication and commitment to the KEEPS research program. The authors dedicate this publication to the memory of Dr. George R. Merriam, principal KEEPS investigator at the Veterans Affairs Puget Sound Health Care System and University of Washington study site, who consistently and cheerfully volunteered to take on many responsibilities essential to the planning, execution, and completion of KEEPS. Dr. Merriam was an outstanding researcher, consummate clinician, and dear friend and colleague and is sorely missed by all.

KEEPS was funded by grants from the Aurora Foundation to the Kronos Longevity Research Institute; the National Institutes of Health (grant HL90639 to Dr. Miller); Mayo Clinic Clinical and Translational Science Award UL1 RR024150; the Mayo Foundation; Brigham and Women’s Hospital/Harvard Medical School Clinical and Translational Science Award UL1 RR024139; and the University of California, San Francisco, Clinical and Translational Science Award UL1 RR024131 from the National Center for Advancing Translational Sciences, a component of the National Institutes of Health and the National Institutes of Health Roadmap for Medical Research. Study medications were supplied in part by Bayer HealthCare and AbbVie Pharmaceuticals.

References


24. Wharton W, Gleason CE, Miller VM, Asthana S. Rationale and design of the Kronos Early Estrogen Prevention Study (KEEPS) and the KEEPS cognitive and affective sub study (KEEPS Cog). Brain research. 2013


Figure 1.
Retention rate at each time point by treatment allocation.
Figure 2.
Unadjusted cross-sectional prevalence of symptoms over time. The proportion of women reporting moderate-severe symptoms is shown at each assessment. The X axis indicates the month of study. Significant differences from PBO indicated by additional symbol: X when p < 0.01, and + where 0.01 < p < 0.05.
### Table 1
Demographic summary of the KEEPS screening sample at baseline, according to randomization assignment

<table>
<thead>
<tr>
<th>Variable</th>
<th>Placebo N=275</th>
<th>o-CEE (Oral) N=230</th>
<th>t-E2 (Patch) N=222</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>52.5(2.5)</td>
<td>52.8(2.6)</td>
<td>52.7(2.6)</td>
<td>0.374</td>
</tr>
<tr>
<td>Time since FMP (months)</td>
<td>21.3(9.6)</td>
<td>21.8(10.2)</td>
<td>22.2(8.7)</td>
<td>0.555</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.4(4.3)</td>
<td>26.0(4.3)</td>
<td>26.0(4.4)</td>
<td>0.503</td>
</tr>
<tr>
<td>BMI (kg/m²)²</td>
<td></td>
<td></td>
<td></td>
<td>0.941</td>
</tr>
<tr>
<td>Normal and underweight (BMI&lt;25kg/m²)</td>
<td>116(42.2)</td>
<td>105(45.7)</td>
<td>98(44.1)</td>
<td></td>
</tr>
<tr>
<td>Overweight (BMI 25 &lt;30 kg/m²)</td>
<td>97(35.3)</td>
<td>79(34.3)</td>
<td>76(34.2)</td>
<td></td>
</tr>
<tr>
<td>Obese (BMI ≥30 kg/m²)</td>
<td>62(22.5)</td>
<td>46(20.0)</td>
<td>48(21.6)</td>
<td></td>
</tr>
<tr>
<td>Ethnic Group²</td>
<td></td>
<td></td>
<td></td>
<td>0.927</td>
</tr>
<tr>
<td>White</td>
<td>211(76.7)</td>
<td>177(77.0)</td>
<td>169(76.1)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>23(8.4)</td>
<td>17(7.4)</td>
<td>14(6.3)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>20(7.3)</td>
<td>17(7.4)</td>
<td>16(7.2)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>7(2.5)</td>
<td>8(3.5)</td>
<td>6(2.7)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>14(5.1)</td>
<td>11(4.8)</td>
<td>17(7.7)</td>
<td></td>
</tr>
<tr>
<td>Education²</td>
<td></td>
<td></td>
<td></td>
<td>0.34</td>
</tr>
<tr>
<td>High School, GED, or less</td>
<td>28(10.2)</td>
<td>16(7.0)</td>
<td>14(6.3)</td>
<td></td>
</tr>
<tr>
<td>Some College/Vocational</td>
<td>46(16.7)</td>
<td>47(20.4)</td>
<td>39(17.6)</td>
<td></td>
</tr>
<tr>
<td>College Degree or Higher</td>
<td>195(70.9)</td>
<td>166(72.2)</td>
<td>166(74.8)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>6(2.2)</td>
<td>1(0.4)</td>
<td>3(1.4)</td>
<td></td>
</tr>
<tr>
<td>Prior Hormone Use ³</td>
<td>52(18.91)</td>
<td>59(25.65)</td>
<td>41(18.47)</td>
<td>0.102</td>
</tr>
<tr>
<td>Currently Use tobacco ³</td>
<td>19(6.91)</td>
<td>14(6.09)</td>
<td>17(7.66)</td>
<td>0.809</td>
</tr>
<tr>
<td>Symptomatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hot Flash ³</td>
<td>126(45.82)</td>
<td>100(43.48)</td>
<td>92(41.44)</td>
<td>0.62</td>
</tr>
<tr>
<td>Night Sweats ³</td>
<td>99(36.00)</td>
<td>83(36.09)</td>
<td>72(32.43)</td>
<td>0.653</td>
</tr>
<tr>
<td>Insomnia ³</td>
<td>93(33.82)</td>
<td>66(28.70)</td>
<td>78(35.14)</td>
<td>0.298</td>
</tr>
<tr>
<td>Irritability ³</td>
<td>42(15.27)</td>
<td>39(16.96)</td>
<td>42(18.92)</td>
<td>0.571</td>
</tr>
<tr>
<td>Mood Swings ³</td>
<td>43(15.64)</td>
<td>34(14.78)</td>
<td>38(17.12)</td>
<td>0.801</td>
</tr>
</tbody>
</table>

¹ Mean(SD) and ANOVA,
² ** frequency (percent) and chi-square,
³ Frequency (percent) and exact chi-square

*Menopause.* Author manuscript; available in PMC 2018 March 01.
Table 2

Prevalence of each symptom at each time point by treatment allocation. P value refers to overall comparison of all three treatments.

<table>
<thead>
<tr>
<th>Time</th>
<th>Hot Flashes</th>
<th></th>
<th>Night Sweats</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PBO</td>
<td>OCEE</td>
<td>t-E2</td>
<td>p</td>
<td>PBO</td>
<td>OCEE</td>
<td>t-E2</td>
</tr>
<tr>
<td>Baseline</td>
<td>45.82</td>
<td>43.48</td>
<td>41.44</td>
<td>0.62</td>
<td>36.00</td>
<td>36.09</td>
<td>32.43</td>
</tr>
<tr>
<td>Month 6</td>
<td>28.32</td>
<td>4.19</td>
<td>7.41</td>
<td>&lt;.001</td>
<td>19.03</td>
<td>4.71</td>
<td>5.29</td>
</tr>
<tr>
<td>Month 12</td>
<td>20.89</td>
<td>4.04</td>
<td>9.63</td>
<td>&lt;.001</td>
<td>14.67</td>
<td>4.55</td>
<td>5.88</td>
</tr>
<tr>
<td>Month 24</td>
<td>19.07</td>
<td>6.70</td>
<td>10.12</td>
<td>&lt;.001</td>
<td>15.81</td>
<td>6.70</td>
<td>8.33</td>
</tr>
<tr>
<td>Month 36</td>
<td>20.00</td>
<td>5.71</td>
<td>10.00</td>
<td>&lt;.001</td>
<td>14.76</td>
<td>4.57</td>
<td>8.24</td>
</tr>
<tr>
<td>Month 48</td>
<td>16.59</td>
<td>6.94</td>
<td>12.94</td>
<td>0.017</td>
<td>15.64</td>
<td>5.78</td>
<td>7.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Insomnia</th>
<th></th>
<th>Irritability</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PBO</td>
<td>OCEE</td>
<td>t-E2</td>
<td>p</td>
<td>PBO</td>
<td>OCEE</td>
<td>t-E2</td>
</tr>
<tr>
<td>Baseline</td>
<td>33.82</td>
<td>28.70</td>
<td>35.14</td>
<td>0.298</td>
<td>15.27</td>
<td>16.96</td>
<td>18.92</td>
</tr>
<tr>
<td>Month 6</td>
<td>17.26</td>
<td>10.99</td>
<td>12.17</td>
<td>0.144</td>
<td>7.52</td>
<td>5.76</td>
<td>6.88</td>
</tr>
<tr>
<td>Month 12</td>
<td>14.22</td>
<td>10.61</td>
<td>10.70</td>
<td>0.429</td>
<td>5.33</td>
<td>6.06</td>
<td>5.35</td>
</tr>
<tr>
<td>Month 24</td>
<td>19.07</td>
<td>12.85</td>
<td>14.29</td>
<td>0.201</td>
<td>4.65</td>
<td>5.03</td>
<td>4.17</td>
</tr>
<tr>
<td>Month 36</td>
<td>22.38</td>
<td>10.29</td>
<td>14.71</td>
<td>0.005</td>
<td>8.10</td>
<td>3.43</td>
<td>5.88</td>
</tr>
<tr>
<td>Month 48</td>
<td>20.38</td>
<td>12.72</td>
<td>9.41</td>
<td>0.007</td>
<td>3.79</td>
<td>4.62</td>
<td>7.65</td>
</tr>
</tbody>
</table>