Title
Estrogen replacement therapy for treatment of mild to moderate Alzheimer disease: A randomized controlled trial

Permalink
https://escholarship.org/uc/item/1v73t255

Journal
Journal of the American Medical Association, 283(8)

ISSN
0098-7484

Authors
Mulnard, RA
Cotman, CW
Kawas, C
et al.

Publication Date
2000-02-23

DOI
10.1001/jama.283.8.1007

License
CC BY 4.0

Peer reviewed
Estrogen Replacement Therapy for Treatment of Mild to Moderate Alzheimer Disease
A Randomized Controlled Trial

Ruth A. Mulnard, RN, DNSc
Carl W. Cotman, PhD
Claudia Kawas, MD
Christopher H. van Dyck, MD
Mary Sano, PhD
Rachelle Doody, MD, PhD
Elizabeth Koss, PhD
Eric Pfeiffer, MD
Sheila Jin, MD, MPH
Anthony Gamst, PhD
Michael Grundman, MD, MPH
Ronald Thomas, PhD
Leon J. Thal, MD
for the Alzheimer’s Disease Cooperative Study

ALZHEIMER DISEASE (AD) affects more than 4 million Americans and is one of the most frequent obstacles to healthy aging in this country. Women appear to be at higher risk for developing AD, only in part due to increased longevity. Because women with AD also live longer than men with AD, there are approximately twice as many women as men in the population with this disorder. It has been suggested that the abrupt decline of estrogen production in postmenopausal women may be associated with a vulnerability of women to develop AD. Men, in contrast, have an intrinsic supply of estrogen.

Context Several reports from small clinical trials have suggested that estrogen replacement therapy may be useful for the treatment of Alzheimer disease (AD) in women.

Objective To determine whether estrogen replacement therapy affects global, cognitive, or functional decline in women with mild to moderate AD.

Design The Alzheimer’s Disease Cooperative Study, a randomized, double-blind, placebo-controlled clinical trial conducted between October 1995 and January 1999.

Setting Thirty-two study sites in the United States.

Participants A total of 120 women with mild to moderate AD and a Mini-Mental State Examination score between 12 and 28 who had had a hysterectomy.

Interventions Participants were randomized to estrogen, 0.625 mg/d (n = 42), or 1.25 mg/d (n = 39), or to identically appearing placebo (n = 39). One subject withdrew after randomization but before receiving medication; 97 subjects completed the trial.

Main Outcome Measures The primary outcome measure was change on the Clinical Global Impression of Change (CGIC) 7-point scale, analyzed by intent to treat; secondary outcome measures included other global measures as well as measures of mood, specific cognitive domains (memory, attention, and language), motor function, and activities of daily living; compared by the combined estrogen groups vs the placebo group at 2, 6, 12, and 15 months of follow-up.

Results The CGIC score for estrogen vs placebo was 5.1 vs 5.0 (P = .43); 80% of participants taking estrogen vs 74% of participants taking placebo worsened (P = .48). Secondary outcome measures also showed no significant differences, with the exception of the Clinical Dementia Rating Scale, which suggested worsening among patients taking estrogen (mean posttreatment change in score for estrogen, 0.5 vs 0.2 for placebo; P = .01).

Conclusions Estrogen replacement therapy for 1 year did not slow disease progression nor did it improve global, cognitive, or functional outcomes in women with mild to moderate AD. The study does not support the role of estrogen for the treatment of this disease. The potential role of estrogen in the prevention of AD, however, requires further research.

JAMA. 2000;283:1007-1015 www.jama.com

For editorial comment see p 1055.
gen by having the ability to aromatize testosterone into estrogen in the brain.

Considerable evidence has emerged from neuropathological studies,7,8 animal behavioral studies,9,10 and human investigations11,12 to suggest that estrogen may be beneficial in improving cognition and mood in AD. Several open-label clinical trials9,11 and 1 randomized clinical trial13 reported selective cognitive improvement in women with dementia who received estrogen replacement therapy (ERT). However, these studies have all been relatively brief, generally ranging from 6 to 8 weeks of treatment with estrogen. Similarly, the number of subjects taking the drug in these trials has been small, ranging from 7 to 12, and most have not used standardized diagnostic criteria. Thus, the evidence for using estrogen as treatment for clinically diagnosed AD is modest at best. Since the potential role of estrogen for the treatment of AD is of public health significance, this intervention, if scientifically supported, could become a routine part of the management regimen for women with AD.

We conducted this study to provide a definitive, placebo-controlled, double-blind, randomized clinical trial of adequate duration and size to determine the benefit of ERT for the treatment of women with mild to moderate AD. We chose to study unopposed estrogens because previous investigations have suggested that progesterone may mitigate some of estrogen’s beneficial effects in the central nervous system.

METHODS

Objectives

The specific aims of the study were to (1) determine if women with AD treated with unopposed estrogens would experience improvement or stability in cognition and other parameters during or after 12 months of therapy; (2) determine which components of the psychometric assessment were improved or stabilized by treatment; (3) determine whether there was a differential response to 2 dosages of estrogen; and (4) establish the safety and tolerability of estrogen in elderly women with AD.

Participants

Participants for this study were recruited from participating sites of the Alzheimer’s Disease Cooperative Study (ADCS; a consortium supported by the National Institute on Aging), with enrollment of 120 women with hysterecomies between October 1995 and January 1999. Selection of women with hysterecomies allowed for the subjects’ long-term exposure (1 year) to unopposed estrogen therapy while eliminating the risk and safety concern of endometrial hyperplasia, which occurs in prolonged unopposed estrogen administration in women with an intact uterus. Thirty-two sites (designated as AD centers by the National Institute on Aging and/or selected sites of the ADCS) participated in the recruitment of these subjects. General inclusion criteria included a diagnosis of probable AD according to National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer’s Disease and Related Disorders Association criteria13 in the mild to moderate stage (the study protocol specified a Mini-Mental State Examination [MMSE]14 score of 14-28; several exceptions were made by the project director to allow for participants with MMSE scores as low as 12); female sex; previous hysterectomy (oophorectomy not required); age older than 60 years; absence of major clinical depressive disorder (as measured by scores of <17 on the Hamilton Depression Rating Scale [Ham-D]15); and normal gynecological, breast, and mammography examination results.

Exclusion criteria were myocardial infarction within 1 year, history of thromboembolic disease or hypercoagulable state, hyperlipidemia, or use of excluded medications (ie, estrogens within 3 months; current use of antipsychotics, anticonvulsants, anti-coagulants, β-blockers, narcotics, methyldopa, clonidine, or prescription cognitive-enhancing or antiparkinson medications, including experimental medications within 60 days prior to baseline. Stable dosages of neuroleptics, antidepressants, anxiolytics, sedatives, and hypnotics were allowed). At the initiation of the protocol, individuals treated with donepezil or tacrine were excluded, but a protocol amendment after 20 months of enrollment allowed the stable use (minimum of 4 weeks) of these medications before screening for the study.

Intervention

This study used a 12-month, randomized, double-blind, placebo-controlled, parallel-group design, in which 120 women were enrolled and randomized to receive a single daily dose of placebo, Premarin (conjugated equine estrogens [CEEs], Wyeth-Ayerst Pharmaceuticals, St Davids, Pa), 0.625 mg, or Premarin, 1.25 mg, followed by a 3-month, single-blind placebo washout phase for all women. Conjugated equine estrogens were chosen for this trial because this formulation is the most commonly prescribed form of ERT. Additionally, CEEs include multiple components, some of which have been shown to have neurotrophic properties, which could be beneficial to brain function.

Participants were randomly allocated to 1 of the 3 treatment arms (Figure 1). Treatments were assigned in randomized permuted blocks of 6 and shipped to each site at the start of the study. The ADCS biostatistical division generated and archived the randomization list. Boxes containing 7-day blister cards of study medication were packaged and shipped by the manufacturer according to instructions provided by the ADCS. Patients were instructed to take 2 identically appearing tablets of study medication each morning. Subjects assigned to estrogen, 1.25 mg/d, received 2 estrogen 0.625-mg tablets; those assigned to estrogen, 0.625 mg/d, received 1 placebo tablet and 1 estrogen 0.625-mg tablet; those assigned to placebo received 2 placebo tablets. Hereafter, the 0.625-mg/d and 1.25-mg/d estrogen dosages are referred to as low and high dosages, respectively. All subjects in this study were required to have a caregiver who administered the investigational agent.
during the trial. Compliance monitoring was done through plasma estradiol level evaluation at each visit and pill count.

Cognitive, global, and other outcome measures were evaluated at screening, baseline, and 2, 6, 12, and 15 months. A telephone check was performed at the 4-month interval to verify ongoing administration of the experimental medication and the status of concurrent medications and to address any issues or concerns, and a brief safety visit was conducted at the 9-month interval. The study was reviewed and approved by the institutional review board at each site. Written informed consent was obtained from all participants.

**Outcome Measures**

The ADCS version of the Clinical Global Impression (CGIC) scale, developed as a semistructured interview from the traditional CGIC scale,\(^\text{16}\) was the primary outcome measure used to assess change from baseline. On this scale, scores of 1, 2, and 3 represent marked, moderate, and mild improvement, respectively; 4 represents no change; and 5, 6, and 7 represent mild, moderate, and marked worsening, respectively.

The MMSE (range, 0-30) and the Clinical Dementia Rating Scale\(^\text{17}\) (CDR; range, 0-5) also were used as global staging instruments. Through the judicious choice of other secondary outcome measures, this study also aimed to address several unanswered questions: (1) independent of mood enhancement, does estrogen therapy improve cognition in AD; (2) by what mechanism is memory improved; and (3) what other clinically relevant benefit does estrogen have in AD?

The first goal was to determine to what degree observed cognitive benefit was associated with improved mood. To this end, it was proposed that the evaluation of estrogen include a traditional measure of depression (Ham-D) and an index of mood state (Multiple Affect Adjective Checklist–Revised [MAACL-R]\(^\text{18}\)) conducted concurrently with other cognitive assessments. To elucidate possible mechanisms for estrogen’s effect on memory that are separate from specific mood alteration, the assessment of memory included measurement of (1) explicit verbal learning (Alzheimer’s Disease Assessment Scale–Cognitive [ADAS-Cog]\(^\text{19}\)), (2) mood-congruent memory (Emotional Face Recognition Test; unpublished data, Elizabeth Koss, PhD, 1995), and (3) visual delayed non-matched to sample recognition (New Dot Test\(^\text{20}\)). We investigated other cognitive benefits that could be associated with estrogen by testing the subjects on measures of attention (Letter Cancellation,\(^\text{21}\) Trail-Making Test A,\(^\text{22}\) and Digit Symbol\(^\text{23}\)), language (Category Fluency\(^\text{24}\) and Letter Fluency\(^\text{25}\)), and motor behavior (Grooved Pegboard Test\(^\text{26}\) and Finger Tapping Test\(^\text{27}\)). Another goal in the choice of secondary measures was to assess the effect of estrogen on activities of daily living abilities in AD, as measured by the Blessed Dementia Rating Scale\(^\text{28}\) and the Dependency Scale.\(^\text{29}\)

**Safety Monitoring**

To reduce risk associated with estrogen administration, screening measures before enrollment included a baseline gynecological and breast examination within the 3 preceding months, a mammogram within the 6 preceding months, and a Papanicolaou test within the previous 3 years. As an additional safety measure, the mammogram was repeated at the end of the 12-month double-blind phase of the study to monitor for any breast complications. At each visit, blood pressure, body weight, and fluid retention (ankle swelling) were monitored. An additional lipid profile was performed at the 2-month visit to detect the complication of hyperlipidemia as a rare reaction to estrogen compounds. Adverse event reports were reviewed quarterly by the independent ADCS safety monitoring committee, who found no necessity to break the blind or interrupt the trial at any time.

**Patient Populations**

For purposes of analysis, 3 patient populations were defined. These included traditional intent-to-treat (based on a last observation carried forward imputation scheme) and completers.
populations (12-month visit completed) as well as a compliers population (defined as all individuals who completed the study and ingested at least 80% of the randomized agent by pill count). Of the 120 randomized subjects, 119 were exposed to the investigational agent (1 subject dropped out because of medical problems before starting the medication). Figure 1 shows the subject flow and disposition through the course of the trial. Of the 120 randomized subjects, 97 completed the trial. There was no attempt to balance the use of donepezil across treatment groups. Table 1 shows that more patients in the ERT groups took donepezil during the course of the trial compared with the placebo group.

**Statistical Analyses**

Power calculations were performed using data from a clinical trial with a similar design that included only women aged at least 60 years with baseline MMSE scores of 16 to 28. Based on the data from this similar trial, with 40 subjects receiving placebo and 80 subjects receiving estrogen, the design power was 81% to detect a 29% difference in the proportion of subjects who worsen in the 2 groups (60% worse in the placebo group vs 31% worse in the estrogen group) using a 2-tailed alpha = .05. Since this was a dosage-finding study, a large effect size was sought to ensure clinical meaningfulness and to estimate the signal size for a possible follow-up trial, if the findings were positive.

In all analyses, a set of predefined covariates was assessed as potential confounders (subject age, apolipoprotein e4 allele frequency, subject education). Any variables unbalanced at baseline (P < .15) and significantly associated with response (P < .10) were included in the statistical model. In the 2-group analysis (combined estrogen groups vs placebo group), there was no significant imbalance of the covariates at baseline, negating the inclusion of these prestated potential confounders to the statistical models. In the 3-group analysis (differential dosage response vs placebo), subject age at baseline was marginally unbalanced (P = .07) in the low-dosage estrogen group, and was significantly associated with 3 outcome variables. Therefore, age was included as a covariate in the statistical analysis models for the ADAS-Cog (P = .07), Dependency Scale (P = .07), and Grooved Pegboard (P = .06).

The primary end point used to evaluate the differential effect of estrogen on progression of AD was the ADCS-CGIC. For this analysis, the 7-point scale was collapsed to 5 points because of lack of subjects in the marked

### Table 1. Baseline Demographic and Clinical Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Placebo (n = 39)</th>
<th>Low-Dosage Estrogen (n = 42)†</th>
<th>High-Dosage Estrogen (n = 39)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD) [range], y</td>
<td>74.1 (16.6) [62-87]</td>
<td>76.8 (6.5) [60-91]</td>
<td>74.2 (7.4) [56-89]</td>
</tr>
<tr>
<td>Race, No. (%)</td>
<td>White</td>
<td>32 (82)</td>
<td>39 (93)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>7 (18)</td>
<td>3 (7)</td>
</tr>
<tr>
<td>Weight, mean (SD) [range], kg</td>
<td>64.8 (16.8) [40-104]</td>
<td>60.3 (10.0) [44-86]</td>
<td>66.0 (12.9) [41-109]</td>
</tr>
<tr>
<td>Education, mean (SD), y</td>
<td>12.1 (3.0)</td>
<td>12.6 (3.3)</td>
<td>12 (3.6)</td>
</tr>
<tr>
<td>Oophorectomy status, No. (%)</td>
<td>Yes</td>
<td>20 (53)</td>
<td>21 (50)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>10 (26)</td>
<td>6 (14)</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>8 (21)</td>
<td>15 (36)</td>
</tr>
<tr>
<td>Prior estrogen exposure</td>
<td>Yes/no/unknown, No.</td>
<td>12/25/1</td>
<td>17/20/5</td>
</tr>
<tr>
<td>Mean duration (SD) [range], y</td>
<td>5.4 (7.6) [1-24]</td>
<td>16.8 (12.9) [1-38]</td>
<td>6.4 (7.3) [1-25]</td>
</tr>
<tr>
<td>Serum estradiol level, mean, pg/mL§</td>
<td>22.7</td>
<td>48.0</td>
<td>58.4</td>
</tr>
<tr>
<td>Hamilton Depression Scale score, mean (SD)</td>
<td>3.8 (4.0)</td>
<td>3.4 (4.0)</td>
<td>3.2 (3.0)</td>
</tr>
<tr>
<td>Clinical Dementia Rating Scale score, mean (SD)</td>
<td>1.0 (0.5)</td>
<td>1.2 (0.5)</td>
<td>1.1 (0.5)</td>
</tr>
<tr>
<td>Mini-Mental State Examination score, mean (SD)</td>
<td>21.1 (3.3)</td>
<td>20.2 (4.7)</td>
<td>20.8 (4.2)</td>
</tr>
<tr>
<td>Alzheimer disease severity, No. (%)</td>
<td>Mild</td>
<td>29 (74)</td>
<td>23 (55)</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>10 (26)</td>
<td>19 (45)</td>
</tr>
<tr>
<td>Donepezil use, No. (%)</td>
<td>None</td>
<td>5 (13)</td>
<td>10 (24)</td>
</tr>
<tr>
<td>Apolipoprotein e4 allele frequency, No.</td>
<td>None</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

*No significant group differences were observed on any characteristic at baseline (P > .05 for all).
†Low-dosage estrogen indicates conjugated equine estrogens, 0.625 mg/d.
‡High-dosage estrogen indicates conjugated equine estrogens, 1.25 mg/d.
§To convert estradiol to picomoles per liter, multiply by 3.67.
or moderately improved categories and was analyzed using ordinal logistic regression. Other primary and secondary end-point treatment differences were analyzed as follows: for continuous measures, linear regression adjusting for baseline values (analysis of covariance) was applied. For categorical measures, logistic regression adjusting for baseline values was used. Confirmatory analyses were conducted using linear regression on changes in scores with no baseline adjustment and $\chi^2$ tests (with Mantel-Haenszel adjustment if necessary) for categorical measures. No interim analyses were performed.

RESULTS

The demographic and clinical characteristics of each group at baseline are illustrated in Table 1. No significant group differences were seen on any characteristic at baseline (all $P$ values > .05).

### Two-Group Comparison

To assess the overall efficacy of estrogen, the low- and high-dosage estrogen groups were combined into a single group of 81 women with AD and hysterectomies who took estrogen. These 81 women were compared with the 39 placebo subjects regarding performance on the ADCS-CGIC, MMSE, ADAS-Cog, and the CDR at 12 months to assess change. The primary intent-to-treat analysis comparing the combined estrogen groups with the placebo group showed no difference between groups for the percentage of patients who worsened on the ADCS-CGIC ($P = .48$), the ADCS-CGIC score ($P = .43$), the MMSE score ($P = .51$), or the ADAS-Cog score ($P = .13$). However, a significant difference was seen on the CDR ($P = .01$) favoring the placebo group (Table 2 and Figure 2). Repeated completers and compliers analyses on the primary outcome measure likewise did not reveal any differential treatment effects on the ADAS-CGIC (data available from the authors on request).

Analysis of secondary outcomes with sensitivity to changes in mood (Hamilton Depression Rating Scale and Dependency Scale) and activities of daily living (Blessed Dementia Rating Scale and Dependency Scale) showed no significant differences between treatment and placebo groups (Table 2). Among the language measures, Category Fluency favored the placebo group ($P = .05$), yet there were no group dif-

### Table 2. Intent-to-Treatment Analysis of Combined Estrogen Groups vs Placebo Group: Mean Change in Scores at 12 Months*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Placebo Group, Mean (SD)</th>
<th>Estrogen Groups, Mean (SD)</th>
<th>$P$ Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients who worsened on ADCS-CGIC, No. (%)</td>
<td>28 (74)</td>
<td>64 (80)</td>
<td>.48</td>
</tr>
<tr>
<td>ADCS-CGIC score‡</td>
<td>5.0 (1.1)</td>
<td>5.1 (0.9)</td>
<td>.43</td>
</tr>
<tr>
<td><strong>Secondary Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mini-Mental State Examination§</td>
<td>−3.1 (4.1)</td>
<td>−2.7 (3.7)</td>
<td>.51</td>
</tr>
<tr>
<td>ADAS-Cog‡</td>
<td>3.6 (4.7)</td>
<td>5.6 (7.3)</td>
<td>.13</td>
</tr>
<tr>
<td>Clinical Dementia Rating Scale‡</td>
<td>0.2 (0.4)</td>
<td>0.5 (0.6)</td>
<td>.01</td>
</tr>
<tr>
<td>Mood scores, Hamilton Depression Rating Scale‡</td>
<td>0.0 (3.9)</td>
<td>0.2 (4.0)</td>
<td>.99</td>
</tr>
<tr>
<td>Memory scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional Face Recognition Test§</td>
<td>−5.7 (22.4)</td>
<td>−9.7 (14.2)</td>
<td>.15</td>
</tr>
<tr>
<td>New Dot Test§</td>
<td>−0.9 (3.1)</td>
<td>−1.5 (3.1)</td>
<td>.32</td>
</tr>
<tr>
<td>Attention scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter Cancellation§</td>
<td>−1.3 (5.5)</td>
<td>−1.4 (7.5)</td>
<td>.80</td>
</tr>
<tr>
<td>Trail-Making Test A‡</td>
<td>18.6 (43.4)</td>
<td>18.9 (48.6)</td>
<td>.94</td>
</tr>
<tr>
<td>Digit Symbol§</td>
<td>−3.9 (6.8)</td>
<td>−3.4 (7.7)</td>
<td>.68</td>
</tr>
<tr>
<td>Language scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category Fluency§§</td>
<td>−2.9 (6.6)</td>
<td>−5.7 (7.6)</td>
<td>.05§</td>
</tr>
<tr>
<td>Letter Fluency§§</td>
<td>−1.7 (6.8)</td>
<td>−2.7 (6.4)</td>
<td>.46</td>
</tr>
<tr>
<td>Motor scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grooved Pegboard Test§</td>
<td>−5.2 (42.4)</td>
<td>3.7 (66.4)</td>
<td>.95</td>
</tr>
<tr>
<td>Finger Tapping Test§</td>
<td>4.0 (9.6)</td>
<td>0.1 (8.8)</td>
<td>.05§</td>
</tr>
<tr>
<td>Activities of daily living scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blessed Dementia Rating Scale part 1‡</td>
<td>1.2 (1.5)</td>
<td>1.0 (1.2)</td>
<td>.60</td>
</tr>
<tr>
<td>Blessed Dementia Rating Scale part 2‡</td>
<td>0.8 (1.6)</td>
<td>1.0 (1.4)</td>
<td>.60</td>
</tr>
<tr>
<td>Dependency Scale‡</td>
<td>0.4 (1.1)</td>
<td>0.4 (0.9)</td>
<td>.30</td>
</tr>
</tbody>
</table>

*ADCS-CGIC indicates Alzheimer’s Disease Cooperative Study version of the Clinical Global Impression of Change scale; ADAS-Cog, Alzheimer’s Disease Assessment Scale–Cognitive. The ADCS-CGIC values reflect mean scores on the 7-point scale.†$P$ values for percentages reflect 2-tailed Fisher exact test; $P$ values for the ADCS-CGIC and Clinical Dementia Rating Scale reflect ordinal logistic regression analyses. All other $P$ values reflect analysis of covariance.

| Results significantly favor placebo group. |

©2000 American Medical Association. All rights reserved.
ferences on Letter Fluency. Of the 2 motor measures, the Grooved Pegboard Test did not detect group differences, but the Finger Tapping Test favored the placebo group \( (P = .05) \). Similar analyses on completer and complier populations showed consistent results (data available on request).

**Three-Group Comparison**

A separate analysis addressed treatment effect differences between each of the 2 dosages of estrogen compared with placebo. Again, the 12-month intent-to-treat analysis showed no difference between groups for the proportion who worsened on the ADCS-CGIC \( (P = .73) \) for both low and high dosages, the ADCS-CGIC score (low dosage, \( P = .66 \); high dosage, \( P = .36 \)), the MMSE score (low dosage, \( P = .48 \); high dosage, \( P = .64 \)), or the ADAS-Cog score (low dosage, \( P = .09 \); high dosage, \( P = .32 \)). However, a significant difference was seen on the CDR at 12 months (low dosage, \( P = .03 \); high dosage, \( P = .01 \)), again favoring the placebo group (Table 3 and Figure 3). In addition, we found a benefit of low-dosage estrogen on the MMSE change in score after 2 months of exposure \( (low \text{ dosage} = -0.36; \text{ placebo} = -1.64; P = .05) \), but the benefit did not persist with continued treatment. There was no evidence of improvement in global functioning at any point in the trial. Repeated completers and compliers analyses on the primary outcome measure likewise did not reveal any differential treatment effects on the ADCS-CGIC (data available from the authors on request).

Analysis of the remaining secondary outcome variables showed either nonsignificant differences between groups or results that favored the placebo group (Table 3). The Ham-D, a measure of mood, did not differ between groups. The subscale factors of the MAACL-R also did not differ between groups. The composite outcome measure likewise did not reveal any differential treatment effects on the CDR, MMSE and ADAS-Cog (data available from the authors). There were no detectable group differences in memory, attention, or language measures. Among the motor measures, the Grooved Pegboard Test did not detect group differences, but the Finger Tapping Test again favored placebo, but only in the low-dosage group \( (P = .04) \). Measures of activities of daily living were not significantly different between groups. Repeated analyses on completer and complier populations showed the same results (data available from the authors).

Two additional analyses were performed to compare results on the primary outcome measures among women with prior estrogen exposure and then among women with prior donepezil treatment. Within each treatment group (placebo, low-dosage and high-dosage estrogen), the participants with a history of estrogen use were compared with the participants who had no history of estrogen use prior to the present trial on four 12-month outcome variables, the CGIC, CDR, MMSE and ADAS-Cog. As reflected in Table 1, years of prior estrogen exposure were comparable between the placebo and high-dosage estrogen groups (3.4 vs 6.4 mean years). For the placebo group, mean differences on the outcome measures were not significant for prior estrogen use vs nonuse, respectively, \( (CGIC, 5.0 \text{ vs } 5.1, P = .32; \text{ CDR, } 1.3 \text{ vs } 1.2, P = .42; \text{ MMSE, } 18.3 \text{ vs } 17.5, P = .90; \text{ ADAS-Cog, } 27.9 \text{ vs } 25.8, P = .41) \). For the high-dosage group, mean differences were likewise not significant \( (CGIC, 5.4 \text{ vs } 5.0, P = .07; \text{ CDR, } 1.7 \text{ vs } 1.5, P = .24; \text{ MMSE, } 16.0 \text{ vs } 19.4, P = .21; \text{ ADAS-Cog, } 31.4 \text{ vs } 26.2, P = .97) \). However, participants in the low-dosage estrogen group with a longer period of prior estrogen exposure (16.8 mean years) had significantly better mean CGIC scores at 12 months \( (4.8 \text{ vs } 5.5, P = .04) \), but no significant impact on the CDR \( (mean, 1.4 \text{ vs } 1.8, P = .71) \), the MMSE \( (mean, 18.8 \text{ vs } 16.4, P = .61) \), and the ADAS-Cog \( (mean, 26.6 \text{ vs } 34.1, P = .86) \). A similar analysis of donepezil users vs nonusers showed no significant differences in performance on the CGIC, CDR, MMSE, or the ADAS-Cog outcome variables.

**Figure 2. Two-Group Intent-to-Treat Analysis**

Data are the mean changes from baseline for the Alzheimer’s Disease Cooperative Study version of the Clinical Global Impression of Change scale (ADCS-CGIC), Alzheimer’s Disease Assessment Scale–Cognitive (ADAS-Cog), Mini-Mental State Examination (MMSE), and Clinical Dementia Rating Scale (CDR). Plotted values for the CDR scores are computed to 2 decimal places. Error bars represent 95% confidence intervals around the mean values; gray lines indicate baseline. See footnote to Table 1 for explanation of treatment groups.
Estrogen replacement therapy for Alzheimer disease

COMMENT

Estrogen failed to improve cognitive or functional outcomes in this 1-year study of women with mild to moderate AD and hysterectomies. Similar to previous reports, we found a benefit of low-dosage estrogen on the MMSE after brief exposure (2 months; \( P = .05 \)), but the benefit did not persist with continued treatment. In fact, patients receiving estrogen appeared to decline more than those receiving placebo on 1 global clinical measure, the CDR, despite the greater use of donepezil in the estrogen-treated patients. Overall, the results of this study do not support the role of estrogen in the treatment of AD.

To date, this study is the largest and the longest study to examine estrogen as a treatment for women with AD. Given that patients receiving estrogen did no better or worse than patients receiving placebo, the use of a larger sample size would not have changed this result.

There are several plausible explanations for the difference between the results of this study and previous studies. Animal studies indicate that in the neural tissue, estrogen modulates cholinergic,3 serotonergic,30 and catecholaminergic31 neurotransmitter systems; regulates synaptogenesis during the estrous cycle; regulates neurogenesis32; and is neuroprotective,33 reducing the brain damage associated with ischemic insult.34,35 It is possible that short-term improvements that have been seen in some clinical trials were due to gene-dependent regulation of neurotransmitter systems such as the up-regulation of cholinergic activity.36 A second mechanism of estrogen action could involve surface receptor–associated signaling via ion channels,

Table 3. Differential Dose-Response Intent-to-Treat Analysis: Mean Change in Scores at 12 Months by Treatment Arm*"
Figure 3. Three-Group Intent-to-Treat Analysis

Data are the mean changes from baseline for the Alzheimer’s Disease Cooperative Study version of the cognitive tests (see legend to Figure 2 for explanation of abbreviations). Plotted values for the CDR scores are computed to 2 decimal places. Error bars represent 95% confidence intervals around the mean values; gray lines indicate baseline. See footnote to Table 1 for explanation of treatment groups.

modulating electrical properties of neurons and transmitter release processes.37,38 Such mechanisms are palliative, however, and insufficient to prevent decline over the long term.

In basic studies on the mechanisms underlying neurodegeneration, it has been suggested that there are at least 2 phases in the process, an initiation phase and a propagation phase.39 It is hypothesized that estrogen can delay the initiation phase but is insufficient to slow the propagation phase. Thus, for example, cell culture studies on primary hippocampal neurons show that estrogen is only partially protective against a variety of insults.33 Estrogen appears to operate in part through a gene-dependent up-regulation of antiapoptotic proteins in the bcl-2 family in vitro and vivo.33,34 In the AD brain, these genes are already up-regulated and, thus, further benefit may not be gained.50 Estrogen also has antioxidant properties, though they are relatively weak compared with vitamin E. The present data suggest that the antioxidant capacity of estrogen is evidently insufficient to slow progression. In addition, the anatomical organization of estrogen receptors may favor a role in early stages. Estrogen receptors are most concentrated in brain regions involved with the initial stages of the disease (eg, the limbic system). As degeneration spreads to other regions, estrogen might be unable to regulate gene-dependent defense mechanisms. Other mechanisms also show selectivity for the initiation and propagation phases. For example, apolipoprotein e4 appears to accelerate disease onset, but most studies agree that it does not slow the rate of progression.39,41,42 Although there have been exceptions,33 the mechanisms underlying the present results are as yet unknown, the data suggest that some therapeutic interventions may only act during selective phases of the disease process.

Thus, in the intact healthy brain, estrogen could play a key neuroprotective role by delaying the initiation phase of neurodegenerative disease onset, thereby supporting the finding of reduced risk of dementia from several published epidemiological studies.34-37 Two multicenter prevention trials are currently under way to answer this question prospectively.

Of public health concern is the tendency for experimental treatments to become standard of care before the rigorous scientific evidence is thoroughly gathered. Such is the concern with estrogen administration for women with AD. Numerous publications with broad-based distributions are now supporting the addition of ERT to the armamentarium of treatments for women with AD as a means to enhance cognitive function and delay progression of the disease. Such clinical practice, begun in advance of rigorous clinical trials, could prove to be detrimental to patient outcome. While other ongoing investigations (Women’s Health Initiative–Memory Study, Women’s International Study of Long Duration Oestrogen for Menopause, and Preventing Postmenopausal Memory Loss and Alzheimer’s with Replacement Estrogens study) will provide needed data on hormone replacement therapy in the primary prevention of AD, this study does not support the role of estrogen for treatment of established AD. However, there remains a possibility that estrogen could have an important role as an adjuvant treatment, or as a means of delaying onset of disease. Further investigation of ERT in these areas is still warranted.

©2000 American Medical Association. All rights reserved.
ESTROGEN REPLACEMENT THERAPY FOR ALZHEIMER DISEASE

Ford, J. Eileen Tekrony, University of Kentucky, Lexington; Angela L. Berry, Washington University, St Louis, Mo; Joan Mackell, New York University, New York, NY; Kimberly Schafer, Rebecca E. Ryan-Jones, University of Minnesota, Minneapolis; Santiago City; David Bennett, Neelum Aggarwal, Concetta DeCarli, Michael Korsakoff’s syndrome. J Neurosci. 1996;16:1797–1811.


28. Dubal DB, Shughrue PJ, Wilson ME, Merch-

27. Hidalgo A, Barani K, Herzen K, Goldman SA. Es-

26. Pike CJ. Estrogen modules neuronal bcl-xL ex-


23. Task Force on Alzheimer’s disease. Arch Neu-

22. Dubois B, Jacqmain L, Dartigues JF, et al. Estro-

21. Richman JD, Kreuzberg JD, Liu B, Dluzen DE.


19. Zuckerman M, Rubin B, Robins S. Validation of

18. Wechsler Adult Intelligence Scale—Revised.

17. Zickel GB, Corralina NG, Tausche AK, et al. Estro-

16. Hamilton M, Shulman LR. A rating scale for depression. J Neu-


14. McCann HP, McHugh PR. Mini-

13. Morris JC. The Clinical Dementia Rating Scale: cur-

12. Epstein, Mayo Clinic Jacksonville, Jacksonville, Fla; Cora


9. Mios MB, Albert MS, Butters N, Payne M. Dif-

8. Zuckerman M, Lubin B, Robins S. Validation of

7. Zuckerman M, Rubin B, Robins S. Validation of

6. Zuckerman M, Lubin B, Robins S. Validation of

5. Zuckerman M, Lubin B, Robins S. Validation of

4. Zuckerman M, Lubin B, Robins S. Validation of

3. Zuckerman M, Lubin B, Robins S. Validation of

2. Zuckerman M, Lubin B, Robins S. Validation of

1. Zuckerman M, Lubin B, Robins S. Validation of

REFERENCES

1. Jorm AF, Korten AE, Henderson AS. The preva-

38. Razandi M, Pedram A, Greene GL, Levin ER. Cell

37. Santarosa D, West BM, Cameron JM, et al. Cell


35. Santarosa D, West BM, Cameron JM, et al. Cell

34. Santarosa D, West BM, Cameron JM, et al. Cell

33. Santarosa D, West BM, Cameron JM, et al. Cell

32. Santarosa D, West BM, Cameron JM, et al. Cell


30. Santarosa D, West BM, Cameron JM, et al. Cell

29. Santarosa D, West BM, Cameron JM, et al. Cell


27. Santarosa D, West BM, Cameron JM, et al. Cell


17. Santarosa D, West BM, Cameron JM, et al. Cell


10. Santarosa D, West BM, Cameron JM, et al. Cell


7. Santarosa D, West BM, Cameron JM, et al. Cell


5. Santarosa D, West BM, Cameron JM, et al. Cell


2. Santarosa D, West BM, Cameron JM, et al. Cell


©2000 American Medical Association. All rights reserved.
the cap, whether in quarterly or annually capped plans, we identified the first month of the year in which the capped limit was exceeded. Unlike Rector, we were not able to identify and exclude members who disenrolled voluntarily. Like Rector, we used an extended Cox model with the internally defined time-dependent variable of reaching the cap to analyze the relationship between reaching the cap and disenrollment from the health plan. Models were estimated for each plan and each year controlling for participant age, sex, and chronic disease score.3

Results. The percentages of members reaching their annual prescription cap for plans A, B, and C, respectively, were 22.6%, 0.7%, and 1.6% in 1997 and 12%, 4.1%, and 3.9% in 1998. Disenrollment rates among those enrolled in the first 3 months of each year for plans A, B, and C, respectively, were 19.3%, 28.9%, and 6.8% in 1997 and 10.4%, 22.9%, and 14.0% in 1998. Among those disenrolling in 1997, 21%, 7%, and 7%, respectively, reenrolled in 1998.

The risk of disenrollment across all plans and both years was significantly associated with older age, greater disease burden (ie, higher chronic disease score), and reaching the cap. In 1997, the relative risks (RRs) of disenrollment in any given month for those reaching the cap for the 3 plans were 2.62 (95% confidence interval [CI], 2.15-3.19), 2.21 (95% CI, 1.70-2.88), and 2.24 (95% CI, 1.43-3.50); in 1998, the RRs of disenrollment were 3.04 (95% CI, 2.40-3.86), 1.79 (95% CI, 1.12-2.86), and 2.30 (95% CI, 1.86-2.86) in plans A, B, and C, respectively.

Comment. Exhaustion of prescription coverage, whether administered on a quarterly or annual basis, was associated with a 2- to 3-fold increase in the RR of disenrollment. These findings expand on those of Rector and suggest that this relationship holds under various scenarios including variation in underlying use, cap amounts, and cap administration.

Emily R. Cox, PhD
Brenda R. Motheral, PhD
Kathi Fairman, MA
Express Scripts Inc
Maryland Heights, Mo

Funding/Support: This research was funded by The Robert Wood Johnson Foundation’s Changes in Health Care Financing and Organization (HCFO) Initiative.