A meta-analysis of the effect of hormone replacement therapy upon depressed mood
Zweifel J E, O'Brien W H

Authors' objectives
To address fundamental questions about the effectiveness of hormone replacement therapy (HRT) on menopausal depressed mood. In particular, the authors sought to evaluate the following:

the overall effect of HRT upon depressed mood; and

the extent to which treatment effects (type of HRT, dose, length of treatment, and route of administration) and methodological factors accounted for variance in treatment outcomes.

Searching
Three search strategies were employed to locate relevant studies.

1. A computer-aided search was conducted of Psychological Abstracts (from 1974 to November 1995) and MEDLINE (from 1966 to November 1995), using four key terms: 'Menopause and Drug Therapy', 'Menopause and Depression', 'Menopause and Psychotherapy', and 'Menopause and Behavior Therapy'. In addition, a manual search of Dissertation Abstracts International was conducted from 1970 to November 1995, which examined studies under the major categories of Health Sciences and Psychology, and used the key term 'menopause'.

2. Articles obtained through the first strategy were reviewed carefully, examining the reference list of each article retrieved for additional studies. These additional articles were then obtained.

3. Major contributors to this research area were contacted for unpublished data and information about studies that may have been missed using the first two strategies.

Study selection
Study designs of evaluations included in the review
Studies were included if hormone therapy was used, and if outcome data on at least one quantitative measure of depression or depressed mood was reported.

There were no restrictions on the study designs included in the review. Studies utilised placebo control groups, no-treatment control groups, treatment versus placebo crossover design, and no control group.

The length of treatment ranged from 1 month to 2 years.

Specific interventions included in the review
HRT. The treatments included oestrogen alone, progesterone alone and in combination with oestrogen, and androgen alone and in combination with oestrogen.

The most common oestrogen treatment was premarin, taken orally at a dosage of either 0.625 or 1.25 mg/day. The most common type of progesterone used was norethisterone.

Participants included in the review
The mean age of the participants was 49.5 years (standard deviation 3.02).

Menopausal status: 16 studies (61.5%) utilised only postmenopausal participants, 2 studies (7.7%) used only perimenopausal participants, and 6 studies (23.1%) used both postmenopausal and perimenopausal participants. The remaining 2 studies (7.7%) did not report the menopausal status of their participants.
Symptomology: most of the studies (61.5%) utilised women who were experiencing menopausal symptoms, 3 studies (11.5%) employed women who were not experiencing menopausal symptoms, and 7 studies (26.9%) failed to report whether participants were symptomatic.

Participant recruitment: 11 studies (42.3%) recruited participants from menopausal clinics, 7 studies (26.9%) used another form of recruitment (e.g. referral from private practices, recruitment through newspaper advertisements), and 8 studies (30.8%) did not report how the participants were recruited.

Outcomes assessed in the review
Depression was measured. The Hamilton Rating Scale for Depression and the Beck Depression Inventory were the most commonly used measures of depression. Other measures were: Blatt Menopausal Inventory, MMPI, Sabbatsberg, Kupperman Index, Visual Analogue Scale, Moos Menstrual Distress Questionnaire, Kellner and Schefield's, Daily Menopausal Rating Scale, MAACL, POMS, Leeds Scale, and the Psychological General Well-being Scale.

How were decisions on the relevance of primary studies made?
The authors do not state how the papers were selected for the review, or how many of the authors performed the selection.

Assessment of study quality
Each study was coded along the following dimensions:

- participant characteristics, e.g. demographics and menopausal status;
- study design, e.g. use of control groups and the method of group assignment;
- treatment parameters;
- dependent variables; and
- outcome statistics. The papers were assessed independently by two Psychology graduate students. An assessment of inter-rater reliability was conducted on six (23%) of the studies. Inter-rater agreement was 93.2%.

Data extraction
The authors do not state how the data were extracted for the review, or how many of the authors performed the data extraction.

Methods of synthesis
How were the studies combined?
The meta-analysis used Cohen's $d$ to measure effect sizes. An effect size (d-value) was calculated for each comparison between treatment and control groups. The effect sizes were corrected for unreliability in the dependent variable measurement, by dividing individual d-values by the square root of the reliability coefficient. They were then averaged (mean-d) to provide summary information on the relationship between HRT and menopausal depressed mood. Several additional mean-ds were calculated to examine the effectiveness of various hormone combinations (e.g. oestrogen alone, and oestrogen combined with progesterone).

In addition, aggregated p-values were calculated using the Stouffer method (see Other Publications of Related Interest).

The fail-safe number, i.e. an estimate of the number of additional null comparisons needed to disconfirm the conclusion that a relationship exists, was calculated for each mean-d.

How were differences between studies investigated?
The authors did not present a specific test for heterogeneity. However, they investigated the following moderator variables as sources for heterogeneity: type of depression measure used; study design; sample size; menopausal status;
Results of the review

Twenty-six studies were included in the meta-analysis. Fourteen studies assigned participants to treatment or control groups through random assignment. Four additional studies used random assignment to assign patients to alternate HRT groups. The remaining studies either did not use random assignment (n=7) or did not report how participants were assigned to groups (n=1).

The total number of participants included in each study ranged from 10 to 110. The average number of students per study was 47 (standard deviation 29).

In most of the analyses, the variance in d-values within different sets of studies was accounted for by sampling error. In four analyses (HRT versus any control, any HRT versus placebo, oestrogen versus any control, and oestrogen versus placebo), a significant amount of residual variance remained after sampling error was corrected for.

Oestrogen significantly reduced depressed mood (effect size 0.69). The first four comparisons (any HRT versus any control, any HRT versus placebo, oestrogen versus any control, and oestrogen versus placebo) yielded mean-ds of approximately 0.70. Similarly, pre-treatment versus post-treatment comparisons for the any HRT and the oestrogen-only treatments yielded mean-ds of approximately 0.78. The authors stated that, assuming a normal distribution of depression scores, an effect size of 0.7 indicated that the average treatment participant had lower levels of depressed mood than 76% of the control participants. Using the criteria of Hunter and Schmidt (see Other Publications of Related Interest no.2) and Cohen (see Other Publications of Related Interest no.3), this effect size would be considered to be 'moderate to large'.

Progesterone alone, and in combination with oestrogen was associated with smaller reductions in depressed mood; the effect sizes were 0.39 and 0.45, respectively. The pre-treatment versus post-treatment comparisons for oestrogen plus progesterone provided a mean-d of 0.50. These values fell in the small to moderate size range (see Other Publications of Related Interest nos.2-3).

The analysis of studies comparing oestrogen alone with oestrogen plus progesterone treatment yielded a mean-d of 0.62. This indicated that the addition of progesterone reduced the effect of HRT on depressed mood.

Androgen alone, and in combination with oestrogen, was associated with smaller reductions in depressed mood; the effect sizes were 1.37 and 0.90, respectively. The pre-treatment versus post-treatment comparisons for androgen alone and androgen plus oestrogen treatments yielded mean-ds of 1.29 and 0.89, respectively. These values can be classified as large effect sizes (see Other Publications of Related Interest nos.2-3).

Studies using the Hamilton and the MAACL yielded larger effect sizes than those using the Beck Depression Inventory. The mean-d associated with control group studies was larger than the that for studies using a crossover design. Studies with small sample sizes yielded a smaller mean-d relative to studies with larger sample sizes. Studies using a treatment length of 1 month or 4 to 6 months yielded smaller mean-ds relative to studies that used other treatment lengths.

The fail-safe number was larger for analyses with a greater number of comparisons, e.g. any HRT versus any control had a fail-safe number of 105.

Authors’ conclusions

The results of this meta-analysis suggested that HRT is effective in reducing menopausal depressed mood. In addition, variation in methodological approaches (i.e. depression measure used, study design, length of treatment) is associated with the variation in effect size.

CRD commentary

The search strategy used to identify relevant studies was very thorough. The research question and the inclusion criteria were clear and comprehensive. The authors provided a validity assessment of the studies to be included in the meta-
Available details of the primary studies were tabulated, with the exception that the number of participants in each study was not reported. The authors found that the majority of the individual studies often did not report important demographic information such as race and socioeconomic status. The inadequate reporting of such information makes it difficult to evaluate the generalisability of the results of HRT treatment.

As discussed by the authors, few studies included participants with significant levels of depression. Many other studies did not report on the participants’ baseline depression scores. The depression scores in the individual papers were reviewed; these indicated that most studies included participants who were either not depressed or who were experiencing only mild levels of depression. There was limited information on the effects of HRT on more significant levels of depression. Thus, the results of this meta-analysis might not be generalisable to clinically depressed populations.

The authors also noted that a large number of the studies utilised women with surgically-induced menopause. This may limit the extent to which these results can be generalised to naturally menopausal women.

The authors noted a further limitation of this meta-analysis, i.e. 35% of the of the studies included did not use a control group. Without the inclusion of a control group, it is impossible to disentangle the treatment and placebo effects, thus the interpretation of these studies must be guarded.

The data synthesis was clear but the authors did not present confidence intervals for the mean-ds.

**Implications of the review for practice and research**

Future research should attempt to answer the following questions.

1. What is the optimal combination of hormones?

2. What are the dose-response relationships?

3. How effective is HRT across demographic variables such as race and socioeconomic status?


Improvements in the reporting of demographic information and information on the effective dose of the hormones used could enhance the researching of these questions.

Future research should attempt to determine why factors such as study design, measure of depression used and length of treatment, influence effect size.

**Bibliographic details**


**PubMedID**

9203229

**Other publications of related interest**

Indexing Status
Subject indexing assigned by NLM

MeSH
Climacteric /drug effects /psychology; Depression /drug therapy /psychology; Double-Blind Method; Estrogen Replacement Therapy; Female; Humans; Middle Aged; Randomized Controlled Trials as Topic

AccessionNumber
11997003496

Date bibliographic record published
31/08/1999

Date abstract record published
31/08/1999

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.