Cochrane Library

Cochrane Database of Systematic Reviews

Oral oestrogen and combined oestrogen/progestogen therapy versus placebo for hot flushes (Review)

MacLennan AH, Broadbent JL, Lester S, Moore V

MacLennan AH, Broadbent JL, Lester S, Moore V.
Oral oestrogen and combined oestrogen/progestogen therapy versus placebo for hot flushes.
Cochrane Database of Systematic Reviews 2004, Issue 4. Art. No.: CD002978.
DOI: 10.1002/14651858.CD002978.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Oral oestrogen and combined oestrogen/progestogen therapy versus placebo for hot flushes

Alastair H MacLennan[1], Jessica L Broadbent[2], Sue Lester[3], Vivienne Moore[4]

[1]Obstetrics and Gynaecology, The University of Adelaide, North Adelaide, Australia. [2]Discipline of Obstetrics & Gynaecology, The University of Adelaide, North Adelaide, Australia. [3]Arthritis Research Lab, Hanson Institute, Adelaide, Australia. [4]Department of Public Health, University of Adelaide, Adelaide, Australia

Contact address: Alastair H MacLennan, Obstetrics and Gynaecology, The University of Adelaide, Women's and Children's Hospital, 72 King William Road, North Adelaide, South Australia, 5006, Australia. alastair.maclennan@adelaide.edu.au.

Editorial group: Cochrane Gynaecology and Fertility Group.
Publication status and date: Edited (no change to conclusions), published in Issue 1, 2009.

Citation: MacLennan AH, Broadbent JL, Lester S, Moore V. Oral oestrogen and combined oestrogen/progestogen therapy versus placebo for hot flushes. *Cochrane Database of Systematic Reviews* 2004, Issue 4. Art. No.: CD002978. DOI: 10.1002/14651858.CD002978.pub2.

ABSTRACT

Background

Hot flushes and night sweats are common symptoms experienced by menopausal women. Hormone therapy (HT), containing oestrogens alone or oestrogens together with progestogens in a cyclic or continuous regimen, is often recommended for their alleviation.

Objectives

To examine the effect of oral HT compared to placebo on these vasomotor symptoms and the risk of early onset side-effects.

Search methods

We searched the Cochrane Menstrual Disorders Group and Subfertility Group trials register (searched May 2002). This register is based on regular searches of MEDLINE, EMBASE, CINAHL, the Cochrane Central Register of Controlled Trials (CENTRAL), PsycINFO, the handsearching of 20 relevant journals and conference proceedings, and searches of several key grey literature sources. We also contacted all relevant pharmaceutical companies, The Journal of the International Menopause Society and Climacteric.

Selection criteria

Double-blind, randomised, placebo-controlled trials of oral HT for at least three months duration.

Data collection and analysis

Study quality and outcome data were assessed independently. Random effects models were considered appropriate due to the variety of trial methodologies. The meta-analyses were explored for sensitivity to trial quality and therapy duration. Symptom frequency and severity were assessed separately, together with withdrawals and side-effects. Frequency data were analysed using the Weighted Mean Difference (WMD) between treatment and placebo outcomes. For severity data, odds ratios were estimated from the proportional odds model. From 115 references originally identified, 24 trials meeting the selection criteria were included in the review. Study participants totaled 3,329. Trial duration ranged from three months to three years.

Main results

There was a significant reduction in the weekly hot flush frequency for HT compared to placebo (WMD -17.92, 95% CI -22.86 to -12.99). This was equivalent to a 75% reduction in frequency (95% CI 64.3 to 82.3) for HT relative to placebo. Symptom severity was also significantly reduced compared to placebo (OR 0.13, 95% CI 0.07 to 0.23). Withdrawal for lack of efficacy occurred significantly more often on placebo therapy (OR 10.51, 95% CI 5.00 to 22.09). Withdrawal for adverse events, commonly breast tenderness, oedema, joint pain and

psychological symptoms, was not significantly increased (OR 1.25, 95% CI 0.83 to 1.90), although the occurrence of any adverse events was significantly increased for HT (OR 1.41, 95% CI 1.00 to 1.99). In women who were randomised to placebo treatment, a 57.7% (95% CI 45.1 to 67.7) reduction in hot flushes was observed between baseline and end of study.

Authors' conclusions

Oral HT is highly effective in alleviating hot flushes and night sweats. Therapies purported to reduce such symptoms must be assessed in blinded trials against a placebo or a validated therapy because of the large placebo effect seen in well conducted randomised controlled trials, and also because during menopause symptoms may fluctuate and after menopause symptoms often decline. Withdrawals due to side-effects were only marginally increased in the HT groups despite the inability to tailor HT in these fixed dose trials. Comparisons of hormonal doses, product types or regimens require analysis of trials with these specific "within study" comparisons.

PLAIN LANGUAGE SUMMARY

Oral hormone therapies help reduce the frequency and severity of hot flushes and night sweats caused by menopause.

Hot flushes and night sweats are common symptoms around the menopause (the end of menstrual periods in a woman's life). During menopause there is a major reduction in sex hormones produced by the ovaries that cause these symptoms. The review of scientifically well conducted trials found that taking oral oestrogen or combined oestrogen and progestogen hormone replacement therapy greatly reduces the frequency and severity of these symptoms. This effect was significantly greater than the reduction of symptoms seen with placebo (dummy tablets) over time. No adverse effects were found but as the trials were only short term, more research is needed.

BACKGROUND

The hot flush (or flash) is the most characteristic manifestation of the climacteric. The climacteric can be defined as all the physiological and pathological events that directly follow the onset of reduced ovarian function, both before and after the last menstrual period (the menopause). The aetiology of the hot flush is complex and is still uncertain but is probably caused by lability in the thermoregulatory centre of the hypothalamus induced by falling oestrogen and progesterone levels (Freedman 1995). Instability of the thermoregulatory centre leads to sudden, transient and erratic peripheral vasodilation in the skin blood vessels with a concomitant sensation of flushing and a measurable increase in skin temperature. Hot flushes can occur with differing severity and at different frequency during the day or night (Porter 1996). At night, these changes may be recognised and referred to as night sweats. Together, hot flushes and night sweats are described as vasomotor symptoms. Hot flushes are not unique to the menopause or to hormonal fluctuations (Mohyi 1997) but they are a very common symptom in the peri-menopausal and early post-menopausal phases of life with up to 75% of women experiencing differing degrees of this symptom (Sturdee 1988).

Oestrogens and progestogens, either in combination or separately, have been used to ameliorate or eliminate vasomotor symptoms around the menopause. It is not clear if a combination of an oestrogen and a progestogen has an additive or synergistic effect on the frequency and severity of vasomotor symptoms. Nor is it clear if these hormones are more or less effective in controlling these symptoms during the peri-menopausal years immediately prior to menopause when ovarian hormonal production is erratic and declining compared to their effect after menopause when ovarian oestrogen production has ceased and endogenous fluctuations are unlikely.

Oestrogen and progestogen therapy may potentially be associated with the onset of early side-effects such as breast tenderness, nausea, atypical uterine bleeding, bloating and a perception of weight gain. Longer term potential adverse outcomes such as any cancer, thromboembolism, stroke and gall bladder disease require assessment in long term randomised control trials such as the Women's Health Initiative (WHI 2002; WHI 2004).

OBJECTIVES

To determine the effectiveness of oral oestrogens alone or in combination with a progestogen, in the amelioration of hot flushes. Hormone Therapy (HT) will be defined as oestrogen therapy or oestrogen therapy with combined, cyclic or continuous progestogen therapy. HT was previously known as hormone replacement therapy but as only selected hormones are given at maintenance doses and do not usually achieve premenopausal serum levels, hormone therapy is a more accurate term (Sturdee 2003).

We wish to examine the following.

(1) The effect of oral HT on hot flushes when compared to placebo.

(2) The separate effects on hot flushes of unopposed oestrogen (E) therapies and combined oestrogen and progestogen therapies (E +P) versus placebo.

(3) The separate effect of these therapies (E and E+P) during the peri-menopause and the post-menopause versus placebo.

(4) To assess the risk of early onset side-effects (as delineated in "Types of Outcome Measures" section) associated with these therapies (E and E+P) versus placebo.

(5) To assess the reduction in hot flushes between baseline and end of study in women randomised to placebo therapy.

METHODS

Criteria for considering studies for this review

Types of studies

All double-blind, randomised, placebo-controlled trials were considered for inclusion in this review.

Exclusion criteria:

- Trial unblinded or single-blinded
- HT and placebo packaging not identical
- Participants not randomised
- No placebo

Types of participants

Suitable participants were defined as menopausal women recruited from any health care setting or a population based sample who may have had either spontaneous menopause or bilateral oophorectomy (removal of both ovaries).

Peri-menopausal women were defined as women with spontaneous menopause who had menstruated irregularly within the last 12 months. Post-menopausal women were defined as women with surgical menopause or women with spontaneous menopause and amenorrhoea for more than 12 months.

Exclusion Criteria:

- Non-menopausal women
- Major intercurrent disease
- Previous HT within one month of commencement of the study

Types of interventions

All oral oestrogens with or without concomitant progestogens (administered as sequential or continuous progestogen therapy) for a minimum treatment period of three months.

Exclusion Criteria:

- Non-oral HT
- Treatment period of less than three months
- Oestrogen content not clear (as assessed by AHM)
- Co-interventions that may potentially affect hot flush outcomes (as assessed by AHM)

Types of outcome measures

The primary outcome was hot flushes which includes the symptoms of night sweats and is defined as any otherwise unexplained sensation of flushing/sweating experienced by the woman being studied. Studies were included that measured a hot

flush (with or without night sweats) outcome such as frequency, severity, presence versus absence, or a combination measure of frequency and severity as either primary or secondary outcomes. Vasomotor outcomes were sought at baseline, three months and/or end of study.

Frequency and/or severity outcomes were also sought for early onset side-effects such as:

- atypical bleeding
- nausea
- vomiting
- breast tenderness
- headaches
- weight changes
- dizziness
- thrombosis (superficial and deep)
- rash
- pruritis (itch)
- other
- mortality
- any adverse events

Other outcomes were:

- adherence to therapy (assessed as both withdrawals from therapy due to adverse events and withdrawals from therapy due to lack of effect)
- quality of life score (any)

Exclusion Criteria:

- no hot flush/vasomotor outcomes

Search methods for identification of studies

We searched for all publications which describe (or might describe) randomised placebo-controlled trials for the treatment of menopausal symptoms. Original searches were performed in February 1998 and November 2000. Updated searches were completed in May 2002.

(1) The Menstrual Disorders and Subfertility Group's trials register was searched for any trials (searched 10 May 2002). This register is based on regular searches of MEDLINE, EMBASE, CINAHL, The Cochrane Central Register of Controlled Trials (CENTRAL), PsycINFO, the handsearching of 20 relevant journals and conference proceedings, and searches of several key grey literature sources.

(2) The citation lists of relevant publications, review articles, and included studies were searched.

(3) All pharmaceutical companies manufacturing oral HT products were contacted to request all data on all randomised controlled trials (RCT's) on their files, including unpublished trials.

(4) The editorial services of the Journal of the International Menopause Society, Climacteric and its members' newsletter were used to request details of unpublished randomised controlled trials.

Data collection and analysis

Selection of studies

Each study identified by the search strategy was assessed against the inclusion criteria by one of the reviewers (AHM for the original and JB for the update). This assessment was performed unblinded. Where there was uncertainty regarding eligibility, a second reviewer (SL for both the original and the update) also assessed the study and a decision reached through discussion.

Quality assessment

Quality assessment of studies that met the inclusion criteria was performed by two independent reviewers, one content expert (AHM) and two non-content experts (SH for the original and JB for the update). Any disagreements in quality ratings were resolved by discussion. Although only double-blind trials were considered for this review, standards for blinding may vary, so studies were assessed with regard to these procedures. The quality related attributes of studies were graded as outlined in Table 1.

Data extraction

Further details related to the characteristics of the included studies were collected independently by two reviewers (AHM and SL for the original, JB for the update).

Outcome data were extracted independently by two reviewers (VM and SL for the original, JB and SL for the update), using a pro forma, and checked for agreement. Any disagreements were resolved by discussion.

Where necessary, additional information was sought from the principal investigators of the study. A number of requests for information are currently outstanding and, if and when received, will be included in an updated version of this review.

The original review included eligible trials that had been located as of January 2000. This current update includes subsequent eligible trials located as of May 2002.

Analysis

Statistical analyses were undertaken following the guidelines developed by the Menstrual Disorders and Subfertility Review Group. For cross-over trials, only results from the end of the first phase (before the treatment cross-over) were used because of the potential carry-over effect of HT from the first treatment phase. Treatment effects were defined as the difference between the HT and placebo group outcomes at the end of study, and measures of the difference between the changes from baseline for the two groups were not used (except for evaluation of the placebo effect). Participants from individual trials were included only once in each comparison and multiple HT arms within each trial were combined as appropriate. For continuous data, multiple HT arms were combined by estimating common means and variances. For dichotomous data, HT arms were combined by simple summation. For studies which reported sequential trials on the same participants, one trial was selected (the first trial reporting vasomotor outcomes) for inclusion.

Odds ratios (OR) were the outcome measure of choice for dichotomous data for hot flush outcomes, as this review includes trials with a differing degree of baseline "risk" of hot flushes (i.e. some studies specifically selected women with hot flushes and some studies did not). Random effects models were considered appropriate as there were many differences between trials (e.g.

participant inclusion criteria, menopausal status, treatment types and dosages) that may potentially influence the size of the treatment effect, although the heterogeneity associated with fixed effects models was always assessed. Fixed effects results were only reported when a random effects analysis could not be performed due to negative "between study" variance for analyses using the inverse variance weighting method. The meta-analyses were explored for sensitivity to individual studies, quality aspects (allocation score and losses), and duration of therapy. Sensitivity to individual studies was assessed by examining their individual contribution to both the overall treatment effect and heterogeneity by the method of Baujat 1999. Concealment of treatment allocation (adequate, not adequate or not specified), losses (<10%, >10%) and duration of therapy beyond three months were each assessed by "between studies" meta-regression with a single covariate, using a moment estimator for between study variance (Thompson 1999).

Data were collected at three months and/or end of study for the following outcomes:

- Hot flush frequency - continuous
- Hot flush weekly weighted score - continuous
- Hot flush severity - continuous, dichotomous (present versus absent, moderate-severe versus mild-absent) or categorical (i.e. the number of women in each severity category)
- Withdrawals from treatment due to lack of effect (dichotomous)
- Withdrawals from treatment due to any adverse events (dichotomous)
- Quality of life scale (e.g. GHQ-11) - continuous

Hot flush frequency data were reported in a variety of formats from daily to weekly to fortnightly frequencies. These data were adjusted to a weekly equivalence by the appropriate multiplicative scaling factor to enable weighted mean differences (WMD) to be used for the comparison. There was evidence of skew ness in these data as assessed by the ratio of mean: standard deviation < 1.64, indicating that natural logarithm (log) transformations may be appropriate. However, transformation of such data requires a prior linear transformation of the continuous data (i.e. x+1), as the log distribution is undefined at x = 0. Log transformations were performed on individual participant data when available. When such data were not available, the mean and standard deviations of the log transformed data were approximated from the original summary statistics by the method-of-moments approach (Whitehead 1999a). The WMD estimate from the log transformed data were then back-transformed to express the treatment effect as the percent change relative to placebo. The formula for back transformation for the % change is *100*{exp(WMD) - 1}*.

The hot flush weekly weighted score (HFWWS) weights the hot flush frequency by the severity score and was treated as a separate outcome. This continuous outcome was also log transformed.

Hot flush severity was predominantly scored on the Kupperman's four point (0 to 3) severity scale, although the Greene's vasomotor sub scale was also included. These data are problematical as they were variously reported as continuous, dichotomous splits and ordinal categorical data. These data were used in three "separate" outcomes in RevMan (continuous severity score, dichotomous presence versus absence of symptoms and dichotomous moderate-severe versus mild-absent symptoms) which resulted in a fragmented analysis. Data from these studies were combined for meta-analysis using the proportional odds model (Whitehead 1999a).

Proportional odds is a property of the logistic distribution (which is very similar to the normal distribution) and by definition, implies that the data can be described by a single odds ratio which is the same for all possible dichotomous splits of the data. Proportional odds ratios were estimated from ordinal data by logistic regression utilising all informative binary splits. Categories where there were no women in either the placebo or HT groups were ignored, and if there were zero cell counts in any other categories, a constant of 0.5 was added to all cells. Standard odds ratios were estimated from dichotomous data and again a constant of 0.5 was added to all cells if there were any zero cell counts. Odds ratios from the summary statistics of continuous data were estimated using an approximation (Whitehead 1999b) which assumed the underlying logistic distribution had variance equal to the pooled variance (sp^2) of the HT and placebo groups. The logistic distribution shape parameter (beta) could then be estimated from the formula, $\beta = (sp/\pi) * 1.732$, and the log odds ratio (LOR) estimated from the formula, $LOR = (\mu 1 - \mu 2)/\beta$(where µ1 and µ2 are the means of the HT and placebo groups respectively). The variance was approximated using the logistic density function to fit the number of observations in each group above and below the common mean and estimated in the usual manner.

Adherence to therapy was assessed by meta-analysis for both the number of withdrawals from therapy due to lack of effect and any adverse events. As these are correlated outcomes, independent treatment effects for each outcome were estimated by odds ratios calculated against the same baseline (the number of women completing the trial).
Side-effect data were poorly reported and generally unsuitable for meta-analysis of individual symptoms. These data were handled by a qualitative summary of the trials.

Any recognised, validated quality of life scale reported in the included studies was included for this outcome.

The weighted mean difference between baseline and end of study results for women randomised to the placebo group was estimated for vasomotor outcomes analysed as continuous variables without adjustment for correlation between paired samples. Ignoring the correlation between paired samples is expected to overestimate the standard errors of the difference and hence result in a conservative estimate. Analysis of the placebo effect was not attempted for categorical hot flush severity data as the implications of ignoring the correlation between paired samples for odds ratios is not as clear.

Minor changes to the protocol

During the course of preparing this review, there were minor changes made to the protocol. The overall aims remained the same, but a question about the magnitude of the reduction in vasomotor symptoms in women randomised to placebo therapy was added as this is of interest in interpreting results from trials evaluating other menopausal therapies that may not have included a placebo or other reference treatment group. We also realised that the protocol is not optimal for definitively comparing oestrogen alone with combined oestrogen and progestogen therapies as the review requirement for a placebo group will exclude some informative trials. However, the comparison of oestrogen only and combined oestrogen plus progestogen trials was retained in the protocol as

it is valid to explore this as a potential source of heterogeneity between the trials included in the review.

RESULTS

Description of studies

One hundred and fifteen references were assessed for inclusion in the review up until September 2002. Sixty six trials (a total of 73 references) were excluded. The predominant reasons for exclusion were: no placebo (23 studies), treatment period less than three months (13 studies), no oral oestrogen administered (11 studies), and no vasomotor/hot flush outcomes (11 studies). Twenty four trials (a total of 37 references) were included in the review. There are a further three trials awaiting assessment pending publication and contact with the authors. The biggest trial to date, namely the Women's Health Initiative (WHI), discouraged women with vasomotor symptoms entering the trial and therefore was not eligible to be included in this review.

Six of the 24 included trials reported results for multiple trials or treatment arms that were not all included in this review. Campbell 1976 reported results for two trials of which only Study II was included as the treatment phase of Study I was only two months. Two arms (an HT arm and a placebo arm) from Hagen 1982 were excluded as both these groups were given thiazide. A progestogen only treatment arm from Dennerstein 1978 was also excluded. Jensen J 1983 reported the results from four sequential studies on the same participants. Study III and Study IV were ineligible for this review as there was no HT administered in Study III and the participants in Study IV had taken HT within one month. Study I and Study II were both eligible, but could not both be included as the Study II participants were the placebo group from Study I. Study II was chosen for inclusion as hot flush outcomes were not reported for Study I. Jensen P 1987 reported four HT groups of which only two were oral HT and Marslew 1992 reported two consecutive studies both with multiple treatment arms. Treatment group A2 from Study A and all of Study B were excluded as they were single-blind (i.e. HT and placebo not identical).

The included studies represent a total of 3329 participants randomised (median 91, range 23 to 875) and 2992 participants analysed (median 83, range 20 to 846), although not all hot flush outcomes were reported in a form suitable for inclusion in the meta-analysis for all participants.

All of the trials were double-blind, placebo-controlled, randomised clinical trials, although for both the Jensen J 1983 and Hagen 1982 trials, double-blinding was implied rather than explicitly stated. Seven of the trials were of cross-over design (Campbell 1976; Chung 1996; Coope 1975; Coope 1981; Davidsen 1974; Dennerstein 1978; Paterson 1982a), and nine were multi centre studies (Archer 1992; Baerug 1998; Baumgardner 1978; Derman 1995; Notelovitz 2000a; PEPI 1998; Symons 2000 Study 1; Symons 2000 Study 2; Viklylaeva 1997). The trials were located in nine different countries, specifically Denmark (Bech 1998; Davidsen 1974; Hagen 1982; Jensen J 1983; Jensen P 1987; Marslew 1992), USA (Archer 1992; Baumgardner 1978; Derman 1995; Martin 1971; Notelovitz 2000a; PEPI 1998; Symons 2000 Study 1; Symons 2000 Study 2), United Kingdom (Campbell 1976; Coope 1975; Coope 1981; Paterson 1982a), Australia (Dennerstein 1978), Chile (Blumel 1994), France (Conard 1995), Hong Kong (Chung 1996), Norway (Baerug 1998) and Russia (Viklylaeva 1997). The source of funding was not stated for six trials (Archer 1992; Campbell 1976; Chung 1996; Davidsen 1974; Jensen P 1987; Martin 1971), and of the remaining trials only one (Blumel 1994) did not list pharmaceutical companies as a source of funds and/or drugs.

The majority of studies recruited healthy menopausal women from a clinical setting (predominantly menopause clinics), three studies (Bech 1998; Dennerstein 1978; Jensen J 1983) recruited women from a population base through advertisements etc and Notelovitz 2000a recruited women from both. Menopausal status was commonly defined through a combination of age, amenorrhea and symptoms attributable to menopause. Confirmation of ovarian failure by measurement of FSH levels was reported in eleven trials (Archer 1999; Bech 1998; Blumel 1994; Chung 1996; Coope 1981; Derman 1995; Notelovitz 2000a; Paterson 1982a; PEPI 1998; Symons 2000 Study 1; Symons 2000 Study 2). Five trials (Archer 1992; Coope 1975; Martin 1971; Paterson 1982a; PEPI 1998) included surgical as well as natural menopausal women, and a further two trials (Chung 1996; Dennerstein 1978) included only surgical menopausal women. It was possible to classify all participants from only two trials (Davidsen 1974; Viklylaeva 1997) as peri-menopausal women and participants from only five trials (Chung 1996; Dennerstein 1978; Jensen J 1983; Symons 2000 Study 1; PEPI 1998) as post-menopausal women according to the review protocol. Participants from the Baerug 1998 trial were subdivided post-randomisation into peri- and post-menopausal. Not every participant in each trial reported vasomotor symptoms at baseline as they were not specified in the inclusion criteria for eleven trials (Bech 1998; Campbell 1976; Chung 1996; Coope 1981; Davidsen 1974; Dennerstein 1978; Hagen 1982; Jensen J 1983; Jensen P 1987; Marslew 1992; PEPI 1998) and two of these trials (Campbell 1976; PEPI 1998) specifically excluded women with severe vasomotor symptoms (although women with lesser vasomotor symptoms were not excluded). For four trials (Campbell 1976; Dennerstein 1978; Hagen 1982; Paterson 1982a), it was not clear if women who had taken HT within the previous month were excluded. Similarly, for three trials (Campbell 1976; Davidsen 1974; Paterson 1982a), it was not explicitly stated that women with major intercurrent diseases were excluded. These studies may not have entirely conformed to the review protocol.

The overall age range of participants was 34 to 64 years, with mean age in the majority of trials approximately 50 years, although two trials (Baumgardner 1978; Bech 1998) did not report the age of participants. In terms of mean age, participants from the Chung 1996 study were the youngest (mean = 43.8 years, standard deviation (SD) = 4.9) and the oldest participants were from the PEPI 1998 study (mean = 56.1 years, SD = 4.3 years). The racial background of the participants was generally not stated.

Twelve trials reported on multiple treatment arms (Archer 1992; Baerug 1998; Baumgardner 1978; Bech 1998; Conard 1995; Dennerstein 1978; Jensen J 1983; Martin 1971; Notelovitz 2000a; PEPI 1998; Symons 2000 Study 1; Symons 2000 Study 2) resulting in a total of forty seven HT arms from twenty four studies. The predominant oestrogens were oestradiol (23 HT arms), conjugated equine oestrogens (CEE, 9 HT arms) and ethinyl oestradiol (9 HT arms) although Mestranol (Martin 1971; Paterson 1982a), Quinestrol (Baumgardner 1978) and piperazine oestrone sulphate (Coope 1981) were also used. The dosages of oestradiol and CEE were predominantly in the moderate range (1 to 2mg oestradiol, 0.625 to 1.25 mg CEE) and only two trials (Hagen 1982; Jensen J

1983) included high doses (mg oestradiol). Thirty of the HT arms were oestrogen plus progestogen (E+P) compared to seventeen with oestrogen therapy alone (E only). Sixteen of the E+P arms were cyclic/sequential therapy as opposed to fourteen with continuous combined therapy. The predominant progestogens used were 19-nor testosterone derivatives (norethisterone acetate, norethisterone, levonorgestrel, nomogestrol acetate) but others included medroxyprogesterone acetate (Blumel 1994; PEPI 1998), cyproterone acetate (Jensen P 1987; Marslew 1992) and micronized progesterone (PEPI 1998). For trials with multiple treatment arms, eight compared oestrogen dose (Archer 1992; Baumgardner 1978; Conard 1995; Jensen J 1983; Martin 1971; Notelovitz 2000a; Symons 2000 Study 1; Symons 2000 Study 2), two compared oestrogen type (Archer 1992; Baumgardner 1978), two compared E only versus E +P (Dennerstein 1978; PEPI 1998) and six compared different types, dose or administration of progestogen (Baerug 1998; Bech 1998; Conard 1995; PEPI 1998; Symons 2000 Study 1; Symons 2000 Study 2).

The duration of therapy (taken as the end of the first phase for cross-over studies) ranged from three months (inclusion criteria) to 36 months with the majority of trials (18/24) of six months or less duration. Two trials were of 12 months duration (Bech 1998; Jensen J 1983), three trials were of 24 months duration (Hagen 1982; Jensen P 1987; Marslew 1992) and one trial was of 36 months duration (PEPI 1998).

Not all studies reported all hot flush outcomes, and for six of the 24 studies (Archer 1992; Campbell 1976; Davidsen 1974; Dennerstein 1978; Hagen 1982; Martin 1971), the hot flush data were unsuitable for inclusion in any of the vasomotor outcome meta-analyses. Repeated attempts to contact the authors to obtain further information were unsuccessful with the exception of the Dennerstein 1978 trial whose first author confirmed that the data from the end of the first phase of the cross-over trial are no longer available.

Risk of bias in included studies

All of the twenty four trials in this review were double-blind, randomised, placebo-controlled trials. However, according to the methodological quality criteria (Table 1), considerable variation existed between the trials (*See* Other Data Table: Analysis 1.1). Fifteen trials received an A for allocation concealment, which was unclear for the remaining nine trials. Nineteen trials contained a statement that HT and placebo were identical (treatment blinding = A) and this was not explicitly stated for five trials. All trials scored A for standardised outcome assessment. Baseline equality in terms of age, menopause status and menopause symptoms was reported for eighteen trials (score A), was unreported for four trials (score B) and some baseline inequality in hot flush outcomes were evident in two trials (Bech 1998; PEPI 1998). Losses to follow-up and intention-to-treat analysis were the quality criteria which received the poorest overall scores. Only six trials reported losses to follow-up of less than 10%. Losses were unclear in two trials and more than 10% in the remaining sixteen trials. Of these sixteen trials, seven (Archer 1992; Bech 1998; Dennerstein 1978; Jensen P 1987; Marslew 1992; Martin 1971; Paterson 1982a) reported losses between 20 to 30%. Only three trials (Derman 1995; PEPI 1998; Symons 2000 Study 2) were clearly analysed on an intention-to-treat basis.

Only one trial received an A quality score in all six categories (Symons 2000 Study 2). Seven trials (Baerug 1998; Baumgardner 1978; Blumel 1994; Derman 1995; PEPI 1998; Symons 2000 Study 1; Viklylaeva 1997) scored A in five categories, six trials (Chung 1996; Coope 1975; Coope 1981; Marslew 1992; Notelovitz 2000a; Paterson 1982a) scored A in four categories, four trials (Archer 1992; Conard 1995; Jensen J 1983; Martin 1971) scored A in three categories, four trials (Bech 1998; Dennerstein 1978; Hagen 1982; Jensen P 1987) scored A in two categories, and the remaining two trials (Campbell 1976; Davidsen 1974) scored A in one category only.

Effects of interventions

Comparison 1: Risk of bias score of included studies
These results have been discussed previously in the Methodological risk of bias of included studies section.

Comparison 2: Any HT versus placebo: vasomotor outcomes at end of study
There were data from nine trials (Baerug 1998; Conard 1995; Coope 1975; Coope 1981; Derman 1995; Notelovitz 2000a; Symons 2000 Study 1; Symons 2000 Study 2; Viklylaeva 1997) with a total of 1104 participants, for hot flush frequency (Comparison 2.1). All of the trials demonstrated a reduction in hot flush frequency for HT compared to the placebo which was statistically significant in eight of the trials, and of borderline significance (p = 0.05) in the ninth trial (Coope 1981). There was a significant mean reduction of approximately 18 hot flushes per week for HT compared to placebo (WMD -17.9, 95% CI -22.9 to -13.0, random effects). The fixed effects analysis indicated substantial heterogeneity (Q = 17.38, df = 8, P = 0.03). It is plausible that the absolute reduction in hot flushes may be positively correlated with the untreated hot flush frequency. In support of this, the treatment effect was significantly greater (P = 0.009) in the three studies with the highest mean placebo group hot flush frequency (>30 hot flushes/week) at the end of the study (Baerug 1998; Conard 1995; Symons 2000 Study 2). Sensitivity analysis indicated that differences in trial duration, allocation score and losses to follow- up did not significantly influence the results.

The log-transformed hot flush frequency data (Comparison 2.2) showed a significant mean percent reduction of approximately 75% (% reduction 75.3, 95% CI 64.3 to 82.3 random effects). These results must be interpreted with caution as an approximation from the summary statistics was used for the log transformation for the majority of trials. However, individual data were available for two trials (Conard 1995; Coope 1975) and the estimate from these two trials (% reduction 73.5, 95% CI 59.5 to 82.6, fixed effects) was in close agreement with the overall estimate. Again, the fixed effects analysis was substantially heterogeneous (Q = 50.11, df = 8, P < 0.0001). No formal analysis of heterogeneity was performed as the approximation used for the log transformations may have contributed to this result.

Two studies (Baerug 1998; Notelovitz 2000a) reported the effect of HT on the HFWWS (Comparisons 2.3, 2.4). Although heterogeneous, the results indicated a significant reduction for HT compared to placebo therapy (% reduction 78.8, 95% CI 17.3 to 94.6, random effects).

Data were available from seven trials for analysis of hot flush severity score as a continuous variable (Comparison 2.5). One of these trials (Derman 1995) employed a different severity scoring scale (0 to 12) instead of the four point scale (0 to 3) employed by the other studies. The standardised mean difference (SMD) is therefore the appropriate statistic for comparisons where this trial

is included, otherwise, the weighted mean difference (WMD) is appropriate.

Data were available from eight trials for the present versus absent dichotomous split (Comparison 2.6.1) and for four trials for the moderate-severe versus mild-absent dichotomous split (Comparison 2.6.2) of the severity data. The moderate-severe versus mild-absent dichotomous split was imprecisely estimated as there were relatively low numbers of women with moderate-severe symptoms in three of the four studies. There was a significant reduction in hot flush severity for HT compared to placebo for each of these measures.

The use of the proportional odds model, as outlined in the Methods, enabled the combination of these results for an overall meta-analysis with 13 trials and a total of 1718 participants (Comparison 2.7). There was a significant reduction in vasomotor severity for HT compared to placebo as assessed by the proportional odds ratios (OR 0.13, 95% CI 0.07 to 0.23, random effects), and the fixed effects analysis revealed substantial heterogeneity (Q = 68.48, df = 12, P < 0.0001). This result for 13 trials is virtually identical to the result obtained for the dichotomous present versus absent data (Outcome 07, Sub-category 01) for eight trials (OR 0.13, 95% CI 0.06 to 0.27, random effects).

Analysis of the sensitivity to individual trials by the method of Baujat 1999 indicated that the PEPI 1998 trial, which demonstrated the smallest (but highly significant) treatment effect, contributed substantially to both the odds ratio and heterogeneity in the fixed effects analysis. However, exclusion of this trial did not substantially change the estimate (OR 0.12, 95% CI 0.08 to 0.18, random effects) and there was still evidence of some heterogeneity in the fixed effects analysis (Q = 19.41, df = 11, P = 0.05).

There are several distinguishing features of the PEPI 1998 trial. It was by far the largest trial with a total of 846 participants, thereby contributing approximately half the participants to the overall analysis. It was one of only three trials analysed on an intention-to-treat basis, and used a dosage of 0.625 mg/day of conjugated equine oestrogens (with and without progestogen) which may be considered in the lower dosage range. It was also by far the longest study with data for 36 months duration of HT. Two trials in the analysis (Jensen P 1987; Marslew 1992) were of 24 months duration of HT and the remaining trials were of 12 months or less duration of HT. In this context, the authors of the PEPI 1998 trial reported an apparent decrease in the size of the treatment effect (as assessed by covariate adjusted odds ratios) between 12 months and 36 months, the only outcome in this trial to do so. The participants in the PEPI 1998 trial were all post-menopausal women (as were participants from the Chung 1996 and Jensen J 1983 trials; the remaining trials included a mixture of peri- and post-menopausal women) and were, on average, the oldest participants across all of the trials, a difference which would be compounded by the relatively long duration of the trial. Furthermore, women with severe vasomotor symptoms were specifically excluded from the trial.

Sensitivity analysis by meta-regression using therapy duration, allocation score and losses in turn as covariates indicated that these covariates were apparently significant in a fixed effects analysis, but not significant in a random effects analysis. This disparity was due to the PEPI 1998 trial as there was no evidence of any effect of these covariates in either a fixed or random effects re-analysis after exclusion of this trial.

Overall, a highly significant benefit of HT compared to placebo was demonstrated for hot flush frequency and severity outcomes, although there was evidence of significant heterogeneity in the size of the treatment effect between trials. The results from trials included in this review, but with hot flush data unsuitable for the meta-analysis (Comparison 2.8), are consistent with this analysis.

Comparison 3: Any HT versus placebo: vasomotor outcomes at 3 months

In addition to the six trials with HT duration ending at three months (Conard 1995; Coope 1975; Baerug 1998; Notelovitz 2000a; Paterson 1982a; Symons 2000 Study 2), data were available at three months from five additional trials of longer duration (Baumgardner 1978; Blumel 1994; Jensen J 1983; Jensen P 1987; Symons 2000 Study 1) resulting in a total of 1,289 participants. Not all trials reported all hot flush outcomes. The meta-analysis of three month data included hot flush frequency (Comparisons 3.1, 3.2), HFWWS (Comparisons 3.3, 3.4) hot flush severity (WMD, Comparison 3.5), odds ratios for the presence versus absence of hot flushes (Comparison 3.6.1) and moderate-severe versus present-absent (Comparison 3.6.2), and proportional odds ratios (Comparison 3.7, Table 02). These results at three months duration of HT are comparable to the overall results for all trials at end of study. This indicates that the benefit of HT therapy on hot flushes is evident at three months, and as noted earlier, the duration of therapy, with the possible exception of the 36 month PEPI 1998 trial, does not appear to have contributed significantly to heterogeneity in the treatment effect as assessed over all the trials included in this review.

Comparison 4: Unopposed oestrogen (E only) versus placebo and oestrogen + progestogen (E + P) versus placebo for vasomotor outcomes

The most appropriate method to compare E-only and E+P therapies is a meta-analysis of direct "within studies" comparisons. This cannot be performed definitively in this review as we may have excluded appropriate studies through the placebo requirement. Between studies comparisons are difficult to interpret because other characteristics that differ between the studies may influence the results. Nevertheless, between studies comparisons were performed for E versus placebo and E+P versus placebo trials (Comparisons 4.1 to 4.7). For the PEPI 1998 trial, the relevant treatment arms from the trial were included in both the E and E+P analysis.

For the hot flush frequency outcomes, there were data for three E-only trials and six E+P trials. The treatment effect for the reduction in weekly hot flush frequency was slightly greater for the E+P versus placebo analysis compared to the E only versus placebo analysis (Comparisons 4.1, 4.2), but this difference was not statistically significant (P = 0.46). There was no evidence of heterogeneity between the three E only trials, but there was evidence of heterogeneity between the six E+P trials.

There were insufficient data for between studies (E versus placebo and E+P versus placebo) analyses for HFWWS (Comparisons 4.3, 4.4).

The analysis of hot flush severity included mean severity score (Comparison 4.5, one E only , six E+P), odds ratios for presence versus absence of hot flushes (Comparison 4.6.1, three E only, six E+P) and moderate-severe versus mild-absent hot flushes (Comparison 4.6.2, two E only, two E+P) and did not show clear differences. However, when all trials with severity data (four E

only, ten E+P) were combined using proportional odds ratios (Comparison 4.7, Table 02), there was an apparent different in the size of the treatment effects between the E versus placebo and E+P versus placebo analyses because the confidence intervals for the two separate comparisons did not overlap. The estimate for the E+P versus placebo analysis was OR 0.10 (95% CI 0.06 to 0.19, random effects) compared to OR 0.35 (95% CI 0.22 to 0.56, random effects) for the E versus placebo analysis. While there was no evidence of heterogeneity between the four E only versus placebo trials (Q = 4.53, df = 3, P = 0.21), there was substantial heterogeneity between the ten E+P versus placebo trials (Q = 31.40, df = 9, P = 0.0003). Again the PEPI 1998 trial was identified as contributing substantially to this heterogeneity. Removal of this trial (Table 02) did not appreciably alter the results but reduced the remaining heterogeneity. The remaining heterogeneity was so low that random effects models could not be estimated due to negative "between study" variance.

The apparent difference in the size of the treatment effect between E only versus placebo and E+P versus placebo trials must be interpreted with caution. Other "between trial" covariates that may have contributed to this apparent effect are the type of oestrogen (the E only versus placebo trials used predominantly conjugated equine oestrogen compared to the predominant use of oestradiol in the E+P versus placebo trials) and differences in participants (the four E only trials included surgical menopause women compared to only two of the ten E+P trials). Relative differences can only be truly assessed within trials.

Comparison 5: Perimenopausal and post-menopausal women: any HT versus placebo (vasomotor outcomes)

It was not possible to compare the meta-analyses of HT versus placebo treatment effects for hot flush outcomes in trials with peri-menopausal women and trials with post-menopausal women as the majority of the trials included a mixture of peri- and post-menopausal women. However, one trial (Baerug 1998), in an apparently pre-planned analysis, grouped participants into peri- and post-menopausal for analysis of the hot flush weekly weighted score (Comparisons 5.3, 5.4). There was a greater absolute reduction in the HFWWS in the post-menopausal women (post-menopausal: WMD -70.1, 95% CI -76.2 to -64.0; peri-menopausal: WMD -35.5, 95%CI -41.8 to -29.2). However, back transformation of the log transformed data indicated comparable treatment effects when expressed as the percent reduction compared to placebo (post-menopausal: % reduction 88.6, 95% CI 86.9 to 90.0; peri-menopausal: % reduction 91.5, 95%CI 89.6 to 93.2).

Comparison 6: Other outcomes

There was an increased risk of withdrawals explicitly due to lack of effect for women randomised to placebo therapy compared to women randomised to placebo HT (Comparison 6.1: OR 10.51, 95% CI 5.00 to 22.09, random effects). Martin 1971 analysed prior data from participants who withdrew due to lack of effect and these women had more severe vasomotor symptoms than women who remained in the trial (P < 0.001). In the PEPI 1998 trial, approximately 11% of women randomised to placebo therapy had begun taking privately prescribed HT by the end of the trial. This bias in withdrawals due to lack of effect has potential to bias against an HT treatment effect and is a methodological problem with these trials which is difficult to address. Analysis of withdrawals due to any adverse events (Comparison 6.2) indicates a small, non-significantly increased risk for women randomised to HT compared to women randomised to placebo therapy (OR 1.25, 95% CI 0.83 to 1.90, random effects). However, reasons for withdrawals from therapy were poorly reported and not fully ascertained. Consequently, these analyses may be subject to reporting and ascertainment bias.

While withdrawal from therapy due to adverse events was not significantly greater with HT therapy, HT was associated with a significantly increased risk in the occurrence of any adverse events (Comparison 6.3: OR 1.41, 95% CI 1.00 to 1.99, P = 0.05, random effects). Adverse events and side-effect data were, in general, anecdotally reported, and not suitable for meta-analysis. Summaries of these data for each trial are listed in Comparison 6.4. Recurrent reasons for withdrawals due to adverse events included breast tenderness, oedema, joint pain and nervous/psychiatric problems. The incidence of serious adverse events was apparently low. Only one study (Coope 1981) reported any participant mortality. Two deaths occurred in this study, one from recurrent gastric carcinoma and one from epileptic seizure and whether these deaths occurred in the HT or placebo phase of this cross-over trial was not reported. Thrombosis was apparently rare but was reported in two studies (Coope 1981 - one case, not clear if associated with HT; PEPI 1998 - two cases, both on HT). Breast cancer was reported in three trials (Bech 1998 - two cases, both on HT; Hagen 1982 - three cases, two in placebo group, one on HT; PEPI 1998 - six cases, all on HT). In the PEPI 1998 trial, four of the six cases of breast cancer were associated with the micronised progestogen combined HT treatment group.

Without meta-analysis it is difficult to draw any meaningful conclusions regarding side-effects. However, as expected, many trials reported withdrawal bleeding associated with combined HT, which was possibly associated with the menopausal status of the participants (Baerug 1998; Coope 1975). Breast tenderness was also common and four studies reported a significant increase with HT compared to placebo (Conard 1995; Bech 1998; Hagen 1982; PEPI 1998) which was possibly associated with oestrogen dose (Archer 1992; Martin 1971).

Only one trial reported quality of life data (Bech 1998, GHQ-11, Comparison 6.5). These results showed a small, non-significantly improved quality of life on HT therapy compared to placebo therapy. Quality of life was a secondary outcome in the PEPI 1998 trial. These data have been requested and will be added to the review if and when it becomes available.

Comparison 7: Comparison of end of study and baseline vasomotor outcomes for women randomised to placebo therapy

In Comparisons 7.1 to 7.5, vasomotor outcomes at baseline and end of study for women randomised to placebo therapy were analysed. These comparisons have ignored the fact that the comparisons are actually paired data (i.e. measured on the same participants), therefore the standard error for each within study comparison is likely to be overestimated, which will result in a conservative overall estimate.

There was a substantial, significant reduction in hot flush frequency and severity outcomes at end of study compared to baseline in women randomised to placebo therapy. For example, back transformation of the log transformed hot flush frequency data (Comparison 7.2) suggests that an approximate 58% reduction in the hot flush frequency for placebo therapy at end of study

compared to baseline (% reduction 57.7, 95%CI 45.1 to 67.7, random effects) is characteristic of these trials.

Comparison 8: Investigation of assumptions

The proportional odds assumption and the approximation used for estimation the log odds ratio (LOR) from summary statistics requires investigation. Full categorical (0 to 3 scale) severity data were available for three studies (Conard 1995; Chung 1996; Blumel 1994) at both baseline and end of study with additional three month data (as well as end of study data) available for Blumel 1994. There were a total of three baseline data sets and four treatment data sets for evaluation of both the proportional odds assumption and the odds ratio approximation used for estimating the odds ratio from the summary statistics (Comparison 8.1). There was no indication of lack of fit due to the proportional odds assumption, as estimated by the residual deviance from the logistic regression model (Whitehead 1999a), in each of these data sets. The approximation to the log odds ratio (LOR) from the summary statistics was generally in close agreement with that derived by logistic regression from the categorical data. Therefore, the proportional odds assumption and its method of approximation from summary statistics, appear valid for these data.

DISCUSSION

This was the first systematic review of the effect of HT on vasomotor symptoms. This update which adds three more trials reinforces but does not change the conclusion of the original review that HT is highly effective in the control of vasomotor symptoms. There were a large number of studies in the international literature on the topic. Due to the variety of modes of HT and the variable quality of the studies, it was decided to limit the review to the effect of oral HT versus placebo on hot flushes and night sweats in well conducted, double-blind, randomised trials of at least three months duration prior to any cross-over of therapies. There also had to be no confounding from concurrent therapies in the experimental protocol. These criteria eliminated many otherwise well conducted studies that lacked, for instance, a placebo group or where therapy continued for less than three months. The quality of the included studies was generally high in regard to concealment of treatment allocation, outcome assessment and baseline equality. However, several included studies had a greater then 10% loss to follow- up, and the majority of studies did not analyse the data on an intention-to-treat basis. Failure to conduct an intention-to-treat analysis has the potential to underestimate the treatment effect if there are more withdrawals from the placebo group due to a perceived failure to alleviate vasomotor symptoms. Similarly, this methodological weakness may also underestimate the number of side-effects if these were the reasons for withdrawal in the participants not followed up.

Overall the withdrawal rate was relatively low. When withdrawals were documented as due to lack of therapeutic effect, they were more common in the placebo group (Comparison 6.1). This was the case in both the meta-analysis of all the trials with such data and in the three trials with intention-to-treat . Withdrawals due to adverse events were more common in the HT group and reached significance in the PEPI trial where intention-to-treat analysis was performed.

An approximate 58% reduction in hot flush frequency between baseline and end of study was observed in women randomised to placebo therapy in the trials included in this review. This reduction may be due to a variety of reasons including fluctuating endogenous oestrogen levels and symptoms during the perimenopause, a natural decline in symptoms over time postmenopausally, relief of anxiety over symptoms due to counseling received during the course of the trial, or systematic differences in the self-recording of symptoms over time. These data illustrate the importance of blinded trials which include a placebo or reference treatment group for evaluation of therapy for vasomotor symptoms. Thus, it is very important for the effect of HT or any other therapeutic modality prescribed for vasomotor symptoms to be assessed in a blinded manner against a placebo or reference therapy lest a false impression of efficacy be claimed. Claims of around a 50% effectiveness in the reduction of hot flushes from baseline may be seen in uncontrolled studies for therapies such as phytoestrogens (Nachtigall 1999). Such degrees of effect are not better than the placebo effect commonly seen in rigorous double-blind randomised placebo-controlled trials assessing these outcomes.

This review shows that oral HT is highly effective in alleviating hot flushes and night sweats compared to placebo. The results were analysed for "end of study" data (ranging from three months to three years) and for three months HT only and these results were comparable. All studies were consistent in showing benefit of HT compared to placebo for vasomotor symptoms, but there was evidence of heterogeneity indicating differences in the size of the effect between studies. For example, for the hot flush severity analysis for the thirteen trials with severity data, there was substantial heterogeneity. ThePEPI 1998 trial, which contributed 37.5% to the total weight of the analysis for the end of study data (Comparison 2.8), contributed substantially to this heterogeneity. The PEPI 1998 trial showed a significant reduction in vasomotor symptoms (OR = 0.42), but this was of a smaller magnitude compared to the effect seen in the other combined trials (OR = 0.12). Reasons why the treatment effect was apparently less in the PEPI 1998 trial may be possibly related to the older participants and/or the longer duration of the trial. Women with severe vasomotor symptoms were specifically excluded from the PEPI 1998 trial, and the participants (with a mean age at baseline of 56 years) were older than the mean age of participants of other trials at baseline. The PEPI study was also by far the longest duration trial (three years) which, in an end of study analysis, would compound this age difference. This age difference may be important as vasomotor symptoms are primarily a feature of the early post-menopausal years and generally disappear over time.

There is potential for publication bias due to selective reporting of vasomotor symptoms in this meta-analysis as nine trials reporting no vasomotor outcomes were excluded from this review. However, the extent of such publication bias is impossible to assess. The clinical effect of HT is broad, not only in relation to the alleviation of symptoms at menopause (of which vasomotor symptoms are only a part), but in relation to other outcomes such as the effect on the endometrium, bowel and breast, bone metabolism, cardiovascular disease, stroke, venous thromboembolism, cognitive function and dementia. The methodologies of all trials were scrutinised carefully for any indication of whether vasomotor symptoms were recorded. Only the large, well funded trials can attempt to assess a broad range of outcomes and it is expected that there are a considerable number of trials which have not addressed vasomotor symptoms. Of the two large trials considered for inclusion in this review, vasomotor data were obtained from the PEPI 1998 trial, and

the authors of the CHART 1996 study confirmed that vasomotor outcomes were not recorded in that trial. The largest trial of HT, WHI, excluded women with prevalent vasomotor symptoms at entry and although menopausal symptoms were recorded in a small minority, this was not a suitable trial to assess the effect of HT on vasomotor symptoms.

Comparing the size of effect in trials that only used combined oestrogen plus progestogen HT versus placebo with trials that only used oestrogen versus placebo has obvious limitations due to possible inter-trial differences in regard to populations, protocols etc. However, an apparently significant trend was seen for combined HT to have a greater effect than oestrogen alone for hot flush severity. When all trials, including PEPI, were combined for severity data (Comparison 4.7), the estimate for oestrogen only therapies versus placebo was OR 0.35 (95% CI 0.22, 0.56) and for combined HT versus placebo, the estimate was OR 0.10 (95% CI 0.06 to 0.19). There was a similar trend in the hot flush frequency analysis, which contained fewer trials, which was not statistically significant (Comparisons 4.1,4.2). These data are suggestive of an additive ameliorative effect on hot flush severity when progestogen is added to oestrogen, but is not conclusive. The most appropriate analysis for comparison of oestrogen alone to combined oestrogen and progestogen therapy is direct, within trial comparisons. Such an analysis could not be performed for this review as we may have excluded informative studies through the placebo requirement. Two trials (Dennerstein 1978; PEPI 1998) in this review did include all three arms in the one trial. The Dennerstein 1978 trial reported no statistically significant difference between the oestrogen alone and combined oestrogen and progestogen arms for hot flush frequency. The PEPI 1998 trial data suggested some increase in effectiveness of combined HT versus placebo (OR 0.38, 95% CI 0.25 to 0.58) over oestrogen alone versus placebo (OR 0.53, 95% CI 0.32 to 0.89), for hot flush severity, but a direct comparison of combined HT versus oestrogen only HT did not reach statistical significance (P = 0.085). The frequency of hot flushes was not assessed in this trial. A separate review is required to examine the relative effectiveness of combined oestrogen and progestogen compared to oestrogen alone.

The oestrogens used in the trials were varied. Twenty trials used oestradiol and nine used conjugated equine oestrogens and ethinyl oestradiol, Mestranol, Quinestrol and piperazine oestrone sulphate were all used at least once in separate trials. Assessment of different oestrogen types and dosages is most appropriately obtained from a meta-analysis of appropriate "within studies" comparisons. The protocol of this review included only placebo-controlled studies and there are separate studies in the literature which compare different oestrogen dosages but do not have a placebo control. A separate review could examine the minimum, clinically effective dose of each oestrogen. However, in clinical practice, each woman absorbs oral oestrogen differently (Helton 1977) and has an individual response to oestrogen such that it is important to use a flexible regimen to find the individual's ideal dose that keeps her between menopausal symptoms and oestrogenic side-effects such as sore breasts or nausea. In clinical practice, younger women in general need higher doses for symptomatic relief than older women and an empirical starting dose can be selected on this basis and varied with the clinical response. Clinical trials of fixed regimens are usually not ideal regimens in clinical practice as flexibility and individualisation of therapy is required to achieve the best clinical response and the greatest chances of adherence to therapy. The optimal dose found in a trial population may not be the optimal dose for an individual and tailoring of hormone replacement regimens and doses is likely to give a better balance between symptoms and side-effects.

The progestogens used in the trials in this review were norethisterone, norethisterone acetate, levonorgestrel, nomogestrol acetate, medroxyprogesterone acetate, cyproterone acetate and micronized progesterone. As with comparisons of oestrogens and oestrogen doses, evaluation of the effect of different types, doses and modes of progestogen therapy on vasomotor symptoms would be best estimated in a separate review with trials selected specifically to address these questions.

Only one trial (Baerug 1998) grouped participants as peri-menopausal or post-menopausal. There were more symptoms in the recently post-menopausal women as assessed by the hot flush weekly weighted score. However, there was no apparent difference in the HT treatment effect when analysed as the percentage reduction relative to the placebo group between peri- and post-menopausal women. Many trials included both peri- and post-menopausal women but did not analyse for this variable. There were insufficient data overall to ascertain whether HT had any greater or lesser effect on vasomotor symptoms before or after menopause.

There were data on quality of life assessment in only one study (Bech 1998). These data showed a small, non-significantly improved quality of life on HT compared to placebo therapy.

The description of possible side-effects and adverse events in the trials assessed was inconsistently reported and often not expressed numerically to allow meta-analysis of these events. The data and descriptions available have been summarised in Comparison 6.4. The incidence of any adverse events was significantly higher (Comparison 6.3), and rates of withdrawal from therapy due to adverse events were marginally, but not significantly, higher for women randomised to HT compared to women randomised to placebo therapy (Comparison 6.2). Uterine bleeding in women on continuous combined HT, nausea and breast tenderness were the commonest reported side-effects. This has been reported further in a recent Cochrane Review (Lethaby 2004). With the exception of uterine bleeding in women on continuous combined HT regimens, most studies reported a low prevalence of side-effects. Clinical practice and other trials (MacLennan 1993) with continuous combined HT show that initial bleeding or spotting is very common but mostly disappears after several months. Headaches were not clearly influenced by HT with some studies reporting a reduction in baseline headaches with HT but a small increase in women on HT who had no headaches at baseline. Headaches as a potential menopausal symptom and the effect of HT on headaches and migraine deserves a formal systematic review of its own where trials have defined this symptom at entry and as an outcome. Similarly, weight gain or loss on HT has been the topic of a separate Cochrane review (Norman 1999) and no change in weight or body mass index compared to placebo therapy was seen in trials where this was a measured outcome. Other reported side-effects and serious adverse events in placebo and treatment groups were reported rarely without any apparent preponderance in either group. Nausea and sore breasts are often described as a start up symptom especially in older women taking HT for the first time. In clinical practice, where these symptoms persist, HT

dosages are usually reduced to try to eliminate these side-effects. An artificial aspect of most clinical trials is that during the trial, the dose of HT is fixed. Thus there may be more withdrawals from therapy due to side-effects than might happen if tailoring of the dose to the individual was allowed. The incidence of rarer, more serious side effects such as breast cancer and stroke have recently been reported for the combined HT and oestrogen only arms of WHI (WHI 2002; WHI 2004). It is important to differentiate between the global index of major adverse outcomes under age 60, when most HT is used and the higher incidence of these morbidities seen when HT was initiated in the older age groups of WHI.

AUTHORS' CONCLUSIONS

Implications for practice

HT is a highly effective therapy for the treatment of hot flushes and night sweats and its effect was sustained in trials of three months to three years duration. Apart from bleeding on combined continuous HT, side-effects were not commonly reported. There were no reports of serious adverse events during these short term trials. The effectiveness and short term safety of oral HT for the alleviation of hot flushes and night sweats appears to be well established in trials performed to date in symptomatic women around menopause. Longer term safety has been addressed in other trials (WHI 2002; WHI 2004), and this may be an issue for women wanting to take HT for longer periods to control symptoms. It is beyond the scope of this review to differentiate between oral HT products, combinations, doses or regimens.

Implications for research

More research is not necessary to confirm the efficacy of oestrogen or combined oestrogen and progestogen in ameliorating hot flushes and night sweats. The effect is very strong. However, further research is merited on the lowest effective doses of HT and whether combinations of low dose oestrogen and progestogens may achieve the equivalent effect of a higher dose of oestrogen when used alone. Although it may be difficult to define equipotency between different oestrogens and progestogens, trials of product versus product and regimen versus regimen would also be helpful to detect any clinical differences, advantages and disadvantages. It would be helpful if such trials used a validated symptom scoring system and methods of describing frequency and severity were standardised and combined into one numerical score. Such trials should carefully monitor reasons for withdrawal for the study and intention-to-treat analysis should be conducted. In time, similar systematic reviews of the effect of HT on other putative menopausal symptoms should be performed e.g. psychological, locomotor and urogenital symptoms, such as a recent Cochrane review which has examined the use of local oestrogen for vaginal atrophy in postmenopausal women (Suckling 2004). The effect of other routes of HT e.g. transdermal and other menopausal therapies e.g. tibolone, also require such reviews.

ACKNOWLEDGEMENTS

The authors wish to thank Sue Hill (SH) for her role as an independent assessor of trial quality. We also wish to thank the authors of publications who responded to requests for further information and the pharmaceutical companies, Pharmacia & Upjohn and Wyeth Ayerst, who provided access to their literature databases for searching for relevant trials.

REFERENCES

References to studies included in this review

Archer 1992 *{published data only}*

* Archer DR, Fischer LA, Rich D, Schade GH, Schwartz S, Wittoff H, et al. Estrace vs Premarin for treatment of menopausal symptoms: dosage comparison study. *Advances in Therapy* 1992;**9**(1):21-31.

Baerug 1998 *{published and unpublished data}*

* Baerug U, Winge T, Nordland G, Faber-Swensson E, Heldaas K, Norling B, et al. Do combinations of 1 mg estradiol and low doses of NETA effectively control menopausal symptoms?. *Climacteric* 1998;**1**:219-28.

Baumgardner 1978 *{published data only}*

* Baumgardner SB, Condrea H, Daane TA, Dorsey JH, Jurow HN, Shively JP, et al. Replacement estrogen therapy for menopausal vasomotor flushes. Comparison of Quinestrol and conjugated estrogens. *Obstetrics and Gynecology* 1978;**51**(4):445-52. [MEDLINE: 78200565]

Bech 1998 *{published data only}*

* Bech P, Munk-Jensen N, Obel EB, Ulrich LG, Eiken P, Pors Nielsen S. Combined versus sequential hormonal replacement therapy: a double-blind, placebo-controlled study on quality of life-related outcome measures. *Psychotherapy and Psychosomatics* 1998;**67**:259-65. [MEDLINE: 98360054]

Obel EB, Munk-Jensen N, Svenstrup B, Bennett P, Micic S, Henrik-Nielsen R, et al. A two-year double-blind controlled study of the clinical effect of combined and sequential postmenopausal replacement therapy and steroid metabolism during treatment. *Maturitas* 1993;**16**:13-21. [MEDLINE: 93156625]

Blumel 1994 *{published and unpublished data}*

* Blumel JE, Roncagliolo ME, Gramegna G, Vasquez R, Estartus AT. Double blind method of the effect of menopause symptoms, lipid profile, and endometrial thickness of continuous therapy with estradiol valerate and medroxyprogesterone acetate [Estudio doble ciego del efecto sobre la sintomatologia menopausica, et perfil lipidico y el grosor endometrial de una terapia continua de valerato de estradiol mas acetato de medroxiprogesterona]. *Revista Chilena de Obstetricia y Ginecologia* 1994;**59**(5):354-60.

Campbell 1976 *{published data only}*

* Campbell S. Double blind psychometric studies on the effects of natural estrogens on post-menopausal women. In: Campbell S editor(s). The Management of the Menopause and Post Menopausal Years. Lancaster: MTP, 1976.

Campbell S, Whitehead M. Oestrogen therapy and the menopausal syndrome. *Clinics in Obstetrics and Gynecology* 1977;**4**(1):31-46. [MEDLINE: 77160842]

Chung 1996 *{published and unpublished data}*

* Chung TKH, Yip SK, Lam P, Chang AMZ, Haines CJ. A randomized, double-blind, placebo-controlled, crossover study on the effect of oral oestradiol on acute menopausal symptoms. *Maturitas* 1996;**25**:115-23. [MEDLINE: 97061572]

Conard 1995 *{published and unpublished data}*

* Conard J, Basdevant A, Thomas J-L, Ochsenbein E, Denis C, Guyene TT, et al. Cardiovascular risk factors and combined estrogen-progestin replacement therapy: a placebo-controlled study with nomegestrol acetate and estradiol. *Fertility and Sterility* 1995;**64**(5):957-62. [MEDLINE: 96049049]

Coope 1975 *{published and unpublished data}*

Coope J. Double blind cross-over study of estrogen replacement therapy. In: Campbell S editor(s). The Management of the Menopause and Post-Menopausal Years. Lancaster: MTP, 1976.

Coope J, Thomson JM, Poller L. Effects of "natural oestrogen" replacement therapy on menopausal symptoms and blood clotting. *British Medical Journal* 1975;**4**(5989):139-43. [MEDLINE: 76063274]

Coope 1981 *{published and unpublished data}*

Coope J. Is oestrogen therapy effective in the treatment of menopausal depression?. *Journal of the Royal College of General Practitioners* 1981;**31**(224):134-40. [MEDLINE: 82010374]

Poller L, Coope J. A double-blind cross-over study of piperazine oestrone sulphate and placebo with coagulation studies. *British Journal of Obstetrics and Gynaecology* 1980;**87**(8):718-25. [MEDLINE: 81039858]

Davidsen 1974 *{published data only}*

* Davidsen PCB, Secher HJ, Pedersen GT, Christennnsen N. Cyclical therapy with oestradiol valerale/norgestrel at the menopause. Report of a double-blind investigation [Sekvensterapi med ostradiolvalerat/norgestrel i klimakteriet]. *Ugeskrift for Laeger* 1974;**136**:1805-8.

Dennerstein 1978 *{published data only}*

Dennerstein L, Burrows GD, Hyman GJ, Sharpe K. Hormone therapy and affect. *Maturitas* 1979;**1**(4):247-59. [MEDLINE: 81011448]

Dennerstein L, Fotherby K, Burrows GD, Senior J, Wood C. Hormone replacement therapy at the menopause. Double blind controlled study of women's preferences. *Australian and New Zealand Journal of Obstetrics and Gynaecology* 1978;**18**:139-43.

* Dennerstein LL, Burrows GD, Hyman G. Menopausal hot flushes: a double blind comparison of placebo, ethinyl oestradiol and norgestrel. *British Journal of Obstetrics and Gynaecology* 1978;**85**(11):852-6. [MEDLINE: 79062144]

Derman 1995 *{published data only}*

* Derman RJ, Dawood MY, Stone S. Quality of life during sequential hormone replacement therapy - a placebo controlled study. *Internationa Journal of Fertility and Menopausal Studies* 1995;**40**(2):73-8. [MEDLINE: 95323009]

Hagen 1982 *{published data only}*

Christiansen C, Christensen MS, McNair P, Hagen C, Stocklund K, Transbol IB. Prevention of early postmenopausal bone loss: controlled 2-year study in 315 normal females. *European Journal of Clinical Investigation* 1980;**10**(4):273-9. [MEDLINE: 81044145]

Hagen C, Christensen M, Christiansen C, Stocklund K-E, Transbol I. Effects of two years' estrogen-gestagen replacement on climacteric symptoms and gonadotropins in the early postmenopausal period. *Acta Obstetrica et Gyanecologica Scandinavica* 1982;**61**(3):237-41. [MEDLINE: 83018970]

Jensen J 1983 *{published data only}*

Christensen M, Hagen C, Christiansen C, Transbol I. Dose-response evaluation of cyclic estrogen/gestagen in postmenopausal women: Placebo-controlled trial of its gynecologic and metabolic actions. *American Journal of Obstetrics and Gynecology* 1982;**144**(8):873-9. [MEDLINE: 83071947]

* Jensen J, Christiansen C. Dose-response and withdrawal effects on climacteric symptoms after hormonal replacement therapy. A placebo-controlled therapeutic trial. *Maturitas* 1983;**5**:125-33. [MEDLINE: 84038950]

Jensen P 1987 *{published data only}*

Christiansen C, Nilas L, Riis BJ, Rodbro P, Deftos L. Uncoupling of bone formation and resorption by combined osetrogen and progestagen therapy in postmenopausal osteoporosis. *Lancet* 1985;**2**:800-1. [MEDLINE: 86013161]

Jensen J, Riis BJ, Christiansen C. Cyproterone acetate, an alternative progestogen in postmenopausl hormone replacement therapy? Effects on serum lipids and lipoproteins. *British Journal of Obstetrics and Gynaecology* 1987;**94**:136-41.

* Jensen PB, Jensen J, Riis BJ, Rodbro P, Strom V, Christiansen C. Climacteric symptoms after oral and percutaneous hormone replacement therapy. *Maturitas* 1987;**9**(3):207-15. [MEDLINE: 88121740]

Marslew 1992 *{published data only}*

* Marlsew U, Overgaard K, Riis BJ, Christiansen C. Two new combinations of estrogen and progestogen for prevention of postmenopausal bone loss: long-term effects on bone, calcium and lipid metabolism, climacteric symptoms and bleeding. *Obsterics and Gynecology* 1992;**79**(2):202-9. [MEDLINE: 92115236]

Marslew U, Riis B, Christiansen C. Progestogens: therapeutic and adverse effects in early post-menopausal women. *Maturitas* 1991;**13**(1):7-16. [MEDLINE: 91319014]

Martin 1971 *{published data only}*

* Martin PL, Burnier AM, Segre EJ, Huix FJ. Graded sequential therapy in the menopause: a double-blind study. *American Journal of Obstetrics and Gynecology* 1971;**111**(2):178-6.

Notelovitz 2000a *{published and unpublished data}*

Notelovitz M, Arce JC, Nanavate N, Huang WC. One milligram 17 beta-estradiol is the optimal dose for treatment of moderate to severe vasomotor symptoms: a dose-response study (Abstract P-3, ninth annual NAMS meeting). Menopause. 1998; Vol. 5:251.

* Notelovitz M, Lenihan JP, McDermott M, Kerber IJ, Nanavati N, Arce J-C. Initial 17b-estradiol dose for treating vasomotor symptoms. *Obstetrics and Gynecology* 2000;**95**(5):726-31.

Paterson 1982a *{published and unpublished data}*

* Paterson ME. A randomised, double-blind, cross-over study into the effect of sequential mestranol and norethisterone on climacteric symptoms and biochemical parameters. *Maturitas* 1982;**4**(2):83-94. [MEDLINE: 83012227]

PEPI 1998 *{published and unpublished data}*

Espeland MA, Bush TL, Mebane-Sims I, Stefanick ML, Johnson S, Sherwin R et al. PEPI Trial Investigators. Rationale, design and conduct of the PEPI trial. Postmenopausal Estrogen/Progestin Interventions. *Controlled Clinical Trials* 1995;**16**(4 Suppl):3S-19S. [MEDLINE: 96000488]

* Greendale GA, Reboussin BA, Hogan P, Barnabel VM, Shumaker S, Johnson S, et al. Symptom relief and side effects of postmenopausal hormones: results from the Postmenopausal Estrogen/Progestin Inverventions Trial. *Obsterics and Gynecology* 1998;**92**(6):982-8. [MEDLINE: 99054298]

Miller VT, Byington RL, Espeland MA, Langer R, Marcus R, et al. Baseline characteristics of the PEPI participants. Postmenopausal Estrogen/Progestin Interventions. *Controlled Clinical Trials* 1995;**16**(4 Suppl):54S-65S. [MEDLINE: 96000491]

The Writing Group for the PEPI Trial. Effects of estrogenor estrogen/progestin regimes on heart disease risk factors in postmenopausal women. *JAMA: Journal of the American Medical Association* 1995;**273**:199-208. [MEDLINE: 95106422]

Symons 2000 Study 1 *{published and unpublished data}*

* Symons J, Kempfert N, Speroff L for the FemHRT Study investigators. Vaginal bleeding in postmenopausal women taking low-dose norethindrone acetate and ethinyl estradiol combinations. *Obstetrics and Gynecology* 2000;**96**(3):366-72.

Symons 2000 Study 2 *{published and unpublished data}*

* Symons J, Kempfert N, Speroff L for the FemHRT Study investigators. Vaginal bleeding in postmenopausal women taking low-dose norethindrone acetate and ethinyl estradiol combinations. *Obstetrics and Gynecology* 2000;**96**(3):366-72.

Viklylaeva 1997 *{published and unpublished data}*

* Vikhlyaeva EM, Zaidiieva YaZ, Lobova TA, Shishkina AV, Larsen S. Trisequens in perimenopausal women with climacteric syndrome: a randomised double-blind trial. Acta Obstetrica et Gynecologica Scandanavica. 1997; Vol. 76, issue 167S:58 (abstract P86.19).

References to studies excluded from this review

Agnusdei 1995 *{published data only}*

* Agnusdei A, Gennari C, Bufalino L. Prevention of early postmenopausal bone loss using low doses of conjugated estrogens and the non-hormonal, bone-active drug Iprilflavone.

Osteoporosis International 1995;**5**(6):462-6. [MEDLINE: 96241283]

Al-Azzawi 1999 *{published data only}*

* Al-Azzawi F, Wahab M, Habiba M, Akkad A, Mason T. Continuous combined hormone replacement therapy compared with tibolone. *Obsterics and Gynecology* 1999;**93**(2):258-64. [MEDLINE: 99129346]

Archer 1999 *{published data only}*

* Archer DF, Furst K, Tipping D, Dain MP, Vandepol C. A randomized comparison of continuous combined transdermal delivery of estradiol-norethindrone acetate and estradiol alone for menopause. CombiPatch Study Group. *Obsterics and Gynecology* 1999;**94**(4):498-503. [MEDLINE: 99439366]

Aslaksen 1982 *{published data only}*

* Aslaksen K, Frankendal B. Effect of oral medroxyprogesterone acetate on menopausal symptoms in patients with endometrial carcinoma. *Acta Obstetrica et Gyanecologica Scandinavica* 1982;**61**(5):423-8. [MEDLINE: 83123032]

Aylward 1974 *{published data only}*

* Aylward M, Holly F, Parker RJ. An evaluation of clinical response to piperazine oestrone sulphate ('Harmogen') in menopausal patients. *Current Medical Research and Opinion* 1974;**2**(7):417-23. [MEDLINE: 75111779]

Bakke 1965 *{published data only}*

* Bakke JL. A double-blind study of a progestin-estrogen combination in the management of the menopause. *Pacific Medicine & Surgery* 1965;**May-June**:200-5.

Barret-Connor 1999 *{published data only}*

* Barrett-Connor E, Young R, Notelovitz M, Sullivan J, Wiita B, Yan HM, et al. A two-year, double-blind comparison of estrogen-androgen and conjugated estrogens in surgically menopausal women. Effects on bone mineral density, symptoms and lipid profile. *Journal of Reproductive Medicine* 1999;**44**(12):1012-20. [MEDLINE: 20114071]

Blahey 1953 *{published data only}*

* Blahey PR. Double-blind and prolonged constant quinestrol therapy in the menopausal patient. *Current Therapeutic Research, Clinical & Experimental* 1969;**11**(12):755-62. [MEDLINE: 70107715]

Blatt 1953 *{published data only}*

* Blatt MH, Wiesbader H, Kupperman HS. Vitamin E and climacteric syndrome. *A.M.A. Archives of Internal Medicine* 1953:792-9.

Cano 1999 *{published data only}*

* Cano A, Tarin JJ, Duenas JL. Two-year prospective, randomized trial comparing an innovative twice-a week progestin regimen with a continuous combined regimen as postmenopausal therapy. *Fertility and Sterility* 1999;**71**(1):129-36. [MEDLINE: 99131828]

CHART 1996 *{published data only}*

* Speroff L, Rowan J, Symons J, Genant H, Wilborn W. The comparative effect on bone density, endometrium, and lipids of continuous hormones as replacement therapy (CHART study). A randomized controlled trial. *JAMA: Journal of the American Medical Association* 1996;**276**(17):1397-404. [MEDLINE: 97047873]

Cheng 1993 *{published data only}*

* Cheng GJ, Liu JL, Zhang Q, Fan W, Ye HF, Wang ZQ, et al. Nylestriol replacement therapy in postmenopausal women. A three-year prospective study. *Chinese Medical Journal* 1993;**106**(12):911-6.

Cooper 1974 *{published data only}*

* Cooper W, Marnie M, Daw E. A clinical trial of piperazine oestrone sulphate (Harmogen) in the treatment of menopausal symptoms. *Current Medical Research and Opinion* 1974;**2**(5):260-3. [MEDLINE: 75006139]

Crona 1988 *{published data only}*

* Crona N, Samsioe G, Lindberg UB, Silfverstolpe G. Treatment of climacteric complaints with Org OD 14: a comparative study with oestradiol valerate and placebo. *Maturitas* 1988;**9**(4):303-8. [MEDLINE: 88246245]

Dennerstein 1980 *{published data only}*

* Dennerstein L, Fotherby K, Burrows GD, Laby B, Wood C. Plasma levels of ethinyl oestradiol and norgestrel during hormone replacement therapy. *Maturitas* 1980;**2**(2):147-54. [MEDLINE: 81148087]

Ditkoff 1991 *{published data only}*

* Ditkoff EC, Crary WG, Cristo M, Lobo RA. Estrogen improves psychological function in asymptomatic postmenopausal women. *Obsterics and Gynecology* 1991;**78**(6):991-5. [MEDLINE: 92050590]

Fedor-Freybergh 1977 *{published data only}*

* Fedor-Freybergh P. The influence of oestrogens on the wellbeing and mental performance in climacteric and postmenopausal women. *Acta Obstetrica et Gyanecologica Scandinavica* 1977;**64**(Suppl):1-91. [MEDLINE: 77263024]

Fritsch 1997 *{published data only}*

* Fritsch H, Fritsch W, Julius U, Rehak E, Wieczorek A, Hanefeld M. Postmenopausal hormone substitution and its effects on the climacteric syndrome, body weight and blood pressure. *Zentralblatt fur Gynakologie* 1997;**119**(3):128-32. [MEDLINE: 97269350]

Furuhjelm 1976 *{published data only}*

* Furuhjelm M, Fedor-Freybergh P. The influence of estrogens on the psyche in climacteric and post-menopausal women. 1st International Congress on the Menopause. La Grande Motte, France. 1976:84-93.

Furuhjelm 1984 *{published data only}*

* Furuhjelm M, Karlgren E, Carlstrom K. The effect of estrogen therapy on somatic and psychical symptoms in

postmenopausal women. *Acta Obstetrica et Gyanecologica Scandinavica* 1984;**63**:655-61.

Geola 1980 *{published data only}*

* Geola FL, Frumar AM, Tataryn IV, Lu KH, Hershman JM, Eggena P, Sambhi MP, Judd HL. Biological effects of various doses of conjugated equine estrogens in postmenopausal women. *Journal of Clinical Endocrinology and Metabolism* 1980;**51**(3):620-5. [MEDLINE: 81007214]

Greenblatt 1959 *{published data only}*

* Greenblatt RB, Barfield WE, Garner JF, Calk GL, Harrod JP. Evaluation of an estrogen, androgen, estrogen-androgen combination, and a placebo in the treatment of the menopause. *Journal of Clinical Endocrinology* 1950;**10**:1547-58.

Hailes 1981 *{published data only}*

* Hailes JD, Nelson JB, Schneider M, Rennie GC, Burger HG. Conjugated equine oestrogen versus placebo in the management of menopausal symptoms. *Medical Journal of Australia* 1981;**2**(7):340-2. [MEDLINE: 82057645]

Hammar 1998 *{published data only}*

* Hammer M, Christau S, Nathorst-Boos J, Rud T, Garre K. A double-blind, randomised trial comparing the effects of tibolone and continuous combined hormone replacement therapy in postmenopausal women with menopausal symptoms. *British Journal of Obstetrics and Gynaecology* 1998;**105**(8):904-1. [MEDLINE: 98417467]

Hammar 1999 *{published data only}*

* Hammar M, Ekblad S, Lonnberg B, Berg G, Lindgren R, Wyon Y. Postmenopausal women without previous or current vasomotor symptoms do not flush after abruptly abandoning estrogen replacement therapy. *Maturitas* 1999;**31**(2):117-22. [MEDLINE: 99243591]

Hovic 1989 *{published data only}*

* Hovik P, Sundsbak HP, Gaasemyr M, Sandvik L. Comparison of continuous and sequential oestrogen-progestogen treatment in women with climacteric symptoms. *Maturitas* 1989;**11**(1):75-82. [MEDLINE: 89261212]

Jarvinen 1971 *{published data only}*

* Jarvinen PA, Kokkonen J, Ryhanen P. Oestriol succinate in the treatment of climacteric women (a double-blind trial). *Acta Obstetricia et Gynecologica Scandinavica* 1971;**Supplement 50**:26.

Jones 1977 *{published data only}*

* Jones MM, Marshall DH, Nordin BEC. Quantitation of menopausal symptomatology and its response to ethinyl oestradiol and piperazine oestrone sulphate. *Current Medical Research and Opinion* 1977;**4**(Suppl 3):12-20.

Khoo 1998 *{published data only}*

* Khoo SK, Coglan MJ, Wright GR, DeVoss KN, Battistutta D. Hormone therapy in women in the menopause transition. Randomised, double-blind, placebo-controlled trial of effects on body weight, blood pressure, lipoprotein levels, antithrombin III activity, and the endometrium. *Medical Journal of Australia* 1998;**268**(5):216-20. [MEDLINE: 98200719]

Kirkham 1991 *{published data only}*

* Kirkham C, Hahn PM, Van Vugt DA, Carmichael JA, Reid RL. A randomized, double-blind, placebo- controlled, cross-over trial to assess the side effects of medroxyprogesterone acetate in hormone replacement therapy.. *Obsterics and Gynecology* 1991;**78**(1):93-7. [MEDLINE: 91261347]

Klaiber 1996 *{published data only}*

* Klaiber EL, Broverman DM, Vogel W, Peterson L, Snyder MB. Individual differences in changes in mood and platelet monoamine oxidase (MAO) activity during hormonal replacement therapy in menopausal women. *Psychoneuroendocrinology* 1996;**21**(7):575-92. [MEDLINE: 97197381]

Kupperman 1953 *{published data only}*

* Kupperman HS, Blatt MHG, Wiesbader H, Filler W. Comparative clinical evaluation of estrogenic preparations by the menopausal amd amenorrheal indices. *Journal of Clinical Endocrinology* 1953;**13**:688-703.

Lagrelius 1986 *{published data only}*

* Lagrelius A, Fredricsson B, Hirt M, Weintraub L. Clinical experience with a low-dose combination of estradiol valerate and levonorgestrel. Double-blind comparative study between SH D 386 F and Cyclabil. Effects on symptoms, lipids and endometrial condition. *Acta Obstetricia et Gynecologica Scandinavica - Supplement* 1986;**134**:97-101. [MEDLINE: 87123387]

Leonetti 1999 *{published data only}*

* Leonetti HB, Longo S, Anasti JN. Transdermal progesterone cream for vasomotor symptoms and postmenopausal bone loss. *Obsterics and Gynecology* 1999;**94**(2):225-8. [MEDLINE: 99359079]

Lind 1979 *{published data only}*

* Lind T, Cameron EC, Hunter WM, Leon C, Moran PF, Oxley A, et al. A prospective, controlled trial of six forms of hormone replacement therapy given to postmenopausal women. *British Journal of Obstetrics and Gynaecology* 1979;**86**(Suppl 3):1-29. [MEDLINE: 80109672]

Martin 1972 *{published data only}*

* Martin PL, Burnier AM, Greaney MO. Oral menopausal therapy using 17- micronized estradiol. A preliminary study of effectiveness, tolerance and patient preference. *Obsterics and Gynecology* 1972;**39**(5):771-4. [MEDLINE: 72176681]

Mattsson 1999 Study1 *{published data only}*

* Mattsson LA. Clinical experience with continuous combined transdermal hormone replacement therapy. *Menopause* 1999;**6**(1):25-9.

Mattsson 1999 Study2 *{published data only}*

* Mattsson LA. Clinical experience with continuous combined transdermal hormone replacement therapy. *Menopause* 1999;**6**(1):25-9.

Morrison 1980 *{published data only}*

* Morrison JC, Martin DC, Blair RA, Anderson GD, Kincheloe BW, Bates GW, et al. The use of medroxyprogesterone acetate for relief of climacteric symptoms. *American Journal of Obstetrics and Gynecology* 1980;**138**(1):99-104. [MEDLINE: 81019192]

Myers 1990 *{published data only}*

* Myers LS, Dixen J, Morrissette D, Carmichael M, Davidson JM. Effects of estrogen, androgen, and progestin on sexual psychophysiology and behavior in postmenopausal women. *Journal of Clinical Endocrinology and Metabolism* 1990;**70**(4):1124-31. [MEDLINE: 90203143]

Nand 1998 *{published data only}*

* Nand SL, Webster MA, Baber R, Heller GZ. Ogen/Provera Study Group. Menopausal symptom control and side-effects on continuous estrone sulfate and three doses of medroxyprogesterone acetate. *Climacteric* 1998;**1**:211-8.

Nand SL, Wren BG. Continuous oestrone sulphate and medroxyprogesterone acetate hormone replacement therapy - a study of bone density over 2 years. Proceedings of the 8th International Congress on the Menopause. Sydney, Australia, November 1996:Abstract S17.

Natchigall 1979 *{published data only}*

Nachtigall LE, Nachtigall RH, Nachtigall RD, Beckman EM. Estrogen replacement therapy II: a prospective study in the relationship to carcinoma and cardiovascular and metabolic problems. *Obstetrics and Gynecology* 1979;**54**(1):74-9.

Nordin 1980 *{published data only}*

* Nordin BEC, Jones MM, Crilly RG, Marshall DH, Brooke R. A placebo-controlled trial of ethinyl oestradiol and norethisterone in climacteric women. *Maturitas* 1980;**2**(3):247-51. [MEDLINE: 81074365]

Notelovitz 2000b *{published data only}*

* Notelovitz M, Cassel D, Hille D, Furst KW, Dain MP, VandePol C, et al. Efficacy of continuous sequential transdermal estradiol and norethindrone acetate in relieving vasomotor symptoms associated with menopause. *American Journal of Obstetrics and Gynecology* 2000;**182**(1 Pt 1):7-12. [MEDLINE: 20114933]

Paterson 1982b *{published data only}*

* Paterson ME. A randomized double-blind cross-over trial into the effect of norethisterone on climacteric symptoms and biochemical profiles. *British Journal of Obstetrics and Gynaecology* 1982;**89**(6):464-72. [MEDLINE: 82206565]

Pattison 1989 *{published data only}*

* Pattison NS, Uptin T, Knox B, France J. Transdermal oestrogen for postmenopausal women: a double blind crossover comparative study with ethinyl oestradiol. *Australian and New Zealand Journal of Obstetrics and Gynaecology* 1989;**29**(1):62-5. [MEDLINE: 89322087]

Place 1985 *{published data only}*

* Place VA, Powers M, Darley PE, Schenkel L, Good WR. A double-blind comparative study of Estraderm and Premarin in the amelioration of postmenopausal symptoms. *American Journal of Obstetrics and Gynecology* 1985;**152**(8):1092-9. [MEDLINE: 85276529]

Polo-Kantola 1998 *{published data only}*

* Polo-Kantola P, Erkkola R, Helenius H, Irjala K, Polo O. When does estrogen replacement therapy improve sleep quality?. *American Journal of Obstetrics and Gynecology* 1998;**178**(5):1002-9. [MEDLINE: 98270738]

Polo-Kantola P, Portin R, Polo O, Helenius H, Irjala K, Erkkola R. The effect of short-term estrogen replacement therapy on cognition: a randomized, double-blind, cross-over trial in postmenopausal women. *Obstetrics and Gynecology* 1998;**91**(3):459-66. [MEDLINE: 98150842]

Pratt 1937 *{published data only}*

* Pratt JP, Thomas WL. The endocrine treatment of menopausal phenomena. *JAMA: Journal of the American Medical Association* 1937;**109**:1875-80.

Rauramo 1975 *{published data only}*

* Rauramo L, Lagerspetz K, Engblom P, Punnonen R. The effect of castration and peroral estrogen therapy on some psychological functions. *Frontiers of Hormone Research* 1975;**3**:94-104.

Rohr 1999 *{published data only}*

* Rohr UD, Nauert C, Stehle B. 17beta-estradiol delivered by three different matrix patches 50 microg/day: a three way cross-over study in 21 postmenopausal women. *Maturitas* 1999;**33**(1):45-58. [MEDLINE: 20049868]

Saarikoski 1981 *{published data only}*

* Saarikoski S, Niemela A, Jokela H, Pystynen P. Effect of oestriol succinate on serum lipids. *Maturitas* 1981;**3**(3-4):235-9. [MEDLINE: 82147818]

Schiff 1979 *{published data only}*

Regestein QR, Schiff I, Tulchinsky D, Ryan K. Relationships among estrogen-induced psychophysiological changes in hypogonadal women. *Psychosomatic Medicine* 1981;**43**(2):147-55. [MEDLINE: 81273831]

* Schiff I, Regestein Q, Tulchinsky D, Ryan KJ. Effects of estrogens on sleep and psychological state of hypogonadal women. *JAMA: Journal of the American Medical Association* 1979;**242**(22):2405-7. [MEDLINE: 80030020]

Schubert 1977 *{published data only}*

* Schubert L, Sansa G, Nencioni T, Polvani F, Pruneri C, Sostero M, et al. The use of cyclofenil in menopausal women. *Acta Europaea Fertilitatis* 1977;**8**(1):83-110. [MEDLINE: 77198417]

Shargil 1985 *{published data only}*

* Shargil AA. Hormone replacement therapy in perimenopausal women with a triphasic contraceptive compound: a three-year prospective study. *International Journal of Fertility* 1985;**30**(1):15-28. [MEDLINE: 85260280]

Sheffrey 1969 *{published data only}*

* Sheffery JB, Wilson TA, Walsh JC. Double-blind, cross-over study comparing chlordiazepoxide, conjugated estrogens, combined chlordiazepoxide and conjugated estrogens, and placebo in the treatment of the menopause. *Medical Annals of the District of Columbia* 1969;**38**(8):433-6.

Sherwin 1989 *{published data only}*

* Sherwin BB, Gelfand MM. A prospective one-year study of estrogen and progestin in postmenopausal women: effects on clinical symptoms and lipoprotein lipids. *Obsterics and Gynecology* 1989;**73**(5 Pt 1):759-66. [MEDLINE: 89201700]

Spencer 1999 *{published data only}*

* Spencer C, Crook D, Ross D, Cooper A, Whitehead M, Stevenson J. A randomised comparison of the effects of oral versus transdermal 17beta-oestradiol, each combined with sequential oral norethisterone acetate, on serum lipoprotein levels. *British Journal of Obstetrics and Gynaecology* 1999;**106**(9):948-53. [MEDLINE: 99419856]

Strickler 1977 *{published data only}*

* Strickler RC, Borth R, Cecutti A, Cookson BA, Harper JA, Potvin R, et al. The role of oestrogen replacement in the climacteric syndrome. *Psychological Medicine* 1977;**7**(4):631-9. [MEDLINE: 78073665]

Strickler RC, Borth R, Woodlever CA. The climacteric syndrome: an estrogen replacement dilemma. *Canadian Medical Association Journal* 1977;**116**(6):586-7. [MEDLINE: 78125254]

Strickler RC, Woolever CA. Estrogen replacement therapy in peri- and postmenopausal women. Gynecologic and Obstetric Investigation. 1976; Vol. 7:105-6.

Studd 1995 *{published data only}*

Studd JW, MacCarthy K, Zamblera D, Dain MP. Efficacy and safety of Menorest (50 mikrog/day) compared to Premarin 0.625 mg in the treatment of menopausal symptoms and the prevention of bone loss, in menopausal women. A single-center, comparative, randomized, double-blind, double-dummy study. *Scandinavian Journal of Rheumatology Supplement* 1996;**103**:89-90. [MEDLINE: 96272924]

* Studd JW, McCarthy K, Zamblera D, Burger HG, Silberberg S, Wren B, et al. Efficacy and tolerance of Menorest compared to Premarin in the treatment of postmenopausal women. A randomised, multicentre, double-blind, double-dummy study. *Maturitas* 1995;**22**(2):105-14. [MEDLINE: 96104065]

Sulak 1999 *{published data only}*

Corson SL, Richart RM, Caubel P, Lim P. Effect of a unique constant-estrogen, pulsed-progestin hormone replacement therapy containing 17beta-estradiol and norgestimate on endometrial histology. *International Journal of Fertility & Womens Medicine* 1999;**44**(6):279-85. [MEDLINE: 20083539]

* Sulak PJ, Caubel P, Lane R. Efficacy and safety of a constant-estrogen, pulsed-progestin regimen in hormone replacement therapy. *International Journal of Fertility & Womens Medicine* 1999;**44**(6):286-96. [MEDLINE: 20083540]

Taga 1999 *{published data only}*

* Taga M, Tanaka K, Honjo H, Miyake A, Aono T, Ohoshima K, et al. The efficacy of several doses of oral 17b-estradiol for the treatment of vasomotor symptoms in Japanese women. International Menopause Society, Conference Poster, Yokohama Japan. 1999.

Thomson 1976 *{published data only}*

* Thomson J. Double blind study on the effect of oestrogen on sleep, anxiety and depression in perimenopausal women: Preliminary results. *Proceedings of the Royal Society of Medicine* 1976;**68**:829-30.

Utian 1971 *{published data only}*

Utian WH. The mental tonic effect of oestrogens administered to oophorectomized females. *South African Medical Journal* 1972;**46**(3):1079-82. [MEDLINE: 72268148]

* Utian WH. The true clinical features of postmenopause and oophorectomy and their response to oestrogen therapy. *South African Medical Journal* 1971;**46**:732-7.

Volpe 1986 *{published data only}*

* Volpe A, Facchinetti F, Grasso A, Petraglia F, Campanini D, Genazzani AR. Benefits and risks of different hormonal replacement therapies in post-menopausal women. *Maturitas* 1986;**84**(4):327-34. [MEDLINE: 87201099]

References to studies awaiting assessment

Ferguson 1981 *{published data only}*

* Ferguson MM, Carter J, Boyle P, Hart D M, Lindsay R. Oral complaints related to climacteric symptoms in oophorectomized women. *Journal of the Royal Society of Medicine* 1981;**74**:492-8.

Hackman 1977 *{published and unpublished data}*

* Hackman BW, Galbraith D. Six month pilot study of oestrogen replacement therapy with piperazine oestrone sulphate and its effect on memory. *Current Medical Research and Opinion* 1977;**4**(3):21-28.

Notelovitz 1990 *{published data only}*

Notelovitz M, Katz-Karp S, Jennings D, Lancaster J, Green EM, Stroll BS, et al. Effect of cyclic estrone sulfate treatment on lipid profiles of postmenopausal women with elevated cholesterol levels. *Obsterics and Gynecology* 1990;**76**(1):65-70. [MEDLINE: 90295182]

Notelovitz 1998b *{published data only}*

Notelovitz M, Arce J-C, Nanavati M, Huang WC. Norethindrone acetate at 0.5mg dose adds to the efficacy of 1mg 17 beta estradiol on vasomotor symptom relief (Abstract P-2, ninth annual meeting of NAMS). Menopause. 1998; Vol. 5:250.

Rebar 2000 *{published data only}*

Rebar RW, Trabal J, Mortola J. Low-dose esterified estrogens (0.3 mg/day): long term and short-term effects on menopausal symptoms and quality of life in postmenopausal women. *Climacteric* 2000;**3**:1-7.

Simon 1999 *{published data only}*

Simon J, Klaiber E, Wiita B, Bowen A, Yang HM. Differential effects of estrogen-androgen and estrogen-only therapy on vasomotor symptoms, gonadotropin secretion, and endogenous androgen bioavailability in postmenopausal women. *Menopause* 1999;**6**(2):138-46. [MEDLINE: 99302451]

Studd 1999 *{published data only}*

* Studd J, Pornel B, Marton I, Bringer J, Varin C, Tsouderos Y, et al. Efficacy and acceptability of intranasal 17b-oestradiol for menopausal symptoms: randomised dose-response study. *Lancet* 1999;**353**:1574-8.

Additional references

Baujat 1999

Baujat B, Mahe C, Guerin S, Pignon J-P. Detection of outlying and influential trials: a graphical method to explore heterogeneity in meta-analysis. VII Cochrane Colloquium. The best evidence for health care: the role of the Cochrane Collaboration. Rome, October 5-9, 1999.

Freedman 1995

Freedman RR, Norton D, Woodward S, Cornelissen G. Core body temperature and circadian rhythm of hot flashes in menopausal women. *Journal of Clinical Endocrinology and Metabolism* 1995;**80**(8):2354-8. [MEDLINE: 95355539]

Helton 1977

Helton ED, Goldzieher JW. Pharmacokinetics of ethynyl estrogens: a review. *Contraception* 1981;**15**(3):255-84. [MEDLINE: 77224477]

Lethaby 2004

Lethaby A, Suckling J, Barlow D, Farquhar CM, Jepson RG, Roberts H. Hormone replacement therapy in postmenopausal women: endometrial hyperplasia and irregular bleeding. *Cochrane Database of Systematic Reviews* 2004, Issue 3.

MacLennan 1993

MacLennan AH, MacLennan A, Wenzel S, Chambers H, Eckert K. Continuous low-dose oestrogen and progestogen hormone replacement therapy: a randomised trial. *Medical Journal of Australia* 1993;**159**(2):102-6. [MEDLINE: 93330087 93330087]

Mohyi 1997

Mohyi D, Tabassi K, Simon J. Differential diagnosis of hot flashes. *Maturitas* 1997;**27**(3):203-14. [MEDLINE: 97434824]

Nachtigall 1999

Nachtigall LE, Nachtigall LB. The effects of isoflavones derived from red clover on vasomotor symptoms and endometrial thickness. Proceedings of the Ninth International Menopause Society World Congress on the Menopause. Climacteric. 1999; Vol. 2 Suppl 1:128.

Norman 1999

Norman RJ, Flight IHK, Rees MCP. Oestrogen and progestogen hormone replacement therapy for peri-menopausal and post-menopausal women: weight and body fat distribution. *Cochrane Database of Systematic Reviews* 1999, Issue 1.

Porter 1996

Porter M, Penney GC, Russell D, Russell E, Templeton A. A population based survey of women's experience of the menopause. *British Journal of Obstetrics and Gynaecology* 1996;**103**(10):1025-8. [MEDLINE: 97017090]

Sturdee 1988

Sturdee D, Brincat M. The hot flush. In: Studd JWW, Whitehead MI editor(s). The Menopause. London: Blackwell Scientific, 1988.

Sturdee 2003

Sturdee DW, MacLennan AH. HT or HRT that is the question? Time to change the terminology of hormone therapy. *Climacteric* 2003;**6**(1):1.

Suckling 2004

Suckling J, Lethaby A, Kennedy R. Local oestrogen for vaginal atrophy in postmenopausal women. *Cochrane Database of Systematic Reviews* 2004, Issue 3.

Thompson 1999

Thompson SG, Sharp S. Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine* 1999;**18**:2693-708. [MEDLINE: 99453126]

WHI 2002

Writing group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. *JAMA: Journal of the American Medical Association* 2002;**288**:321-33.

WHI 2004

The Women's Health Initiative Steering Committee. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy. The Women's Health Initiative Randomised Controlled Trial. *JAMA: Journal of the American Medical Association* 2004;**294**:1701-12.

Whitehead 1999a

Whitehead A, Bailey AJ, Elbourne D. Combining summaries of binary outcomes with continuous outcomes in a meta-analysis. *Journal of Biopharmaceutical Statistics* 1999;**9**(1):1-16. [MEDLINE: 99190328]

Whitehead 1999b

Whitehead A. Personal Communication 1999.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Archer 1992

Methods	Study Design: parallel, double-blinded, multicentre, placebo-control, randomised clinical trial Randomisation: not stated Number of Centres: 7 Duration of Trial: 12 weeks Power Calculations: none Number of women randomised: 128 to five treatment arms Number of women analysed: 100 Intention-to-treat analysis: no Losses to follow-up /withdrawals from treatment: 28/128 = 22% Compliance: 98% compliance in patients who completed the study (assessed by unused tablets) Source of Funding: not stated
Participants	Menopausal status: peri- and post-menopausal Age : mean 50.6 years, SD 5.9 Location: USA Ethnicity: not stated Source: general and gynaecological practices Inclusion Criteria: healthy natural or surgically post-menopausal women, FSH > 40IU/ml and serum oestradiol < 30 pg/mL (surgical menopause only), age 40 to 60 years (natural menopause only), moderate severe vasomotor symptoms (>5/day of moderate to severe intensity) Exclusion Criteria: significant past or present illness, genitourinary symptoms, psychological symptoms, gastrointestinal conditions, chronic headaches, any contraindications to oestrogen usage, HRT within 3 months, concomitant medications that may affect study parameters, alcohol or drug abuse, laboratory abnormalities, Pap smear dysplasia, endometrial hyperplasia Confirmation of Ovarian Failure: FSH > 40IU/ml and serum oestradiol < 30 pg/mL (surgical menopause only) Baseline Equality: not reported for vasomotor symptoms. Groups equal for age, age at menopause, weight, history of HRT and history of abnormal Pap smear Baseline Symptoms: all women had moderate-severe vasomotor symptoms at baseline (inclusion criteria)
Interventions	Rx1 (E only): micronised oestradiol 1mg/day (Estrace)Rx2 (E only): micronised oestradiol 2 mg/day (Estrace)Rx3 (E): conjugated equine oestrogens (CEE) 0.625 mg/day (Premarin)Rx4 (E): conjugated equine oestrogens (CEE) 1.25 mg/day (Premarin)Rx5: placeboHRT and placebo tablets were identical.Co-interventions: none reported
Outcomes	1. Losses to follow-up. 2. Any adverse events. Mean daily frequency of vasomotor events was not suitable for meta-analysis as no standard deviations or ranges were reported. Similarly, data for side-effects such as breast tenderness, nausea, vomiting and headaches was incomplete for meta-analysis.
Notes	Attempts were made to contact the authors, but no response was obtained.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Baerug 1998

Methods	Study Design: parallel, double-blind, multicentre, placebo-control, clinical trial Randomisation: coded medication provided by manufacturer Number of centres: 5 Duration of Trial: 12 weeks Power Calculations: none

Baerug 1998 *(Continued)*

	Number of women randomised: 119 to three treatment groups Number of women analysed: 108 Intention-to-treat analysis: no Losses to follow-up/withdrawals from treatment: 11/119 = 9%. Breakdown: adverse events (5), lack of effect (3), other (3). Compliance: not stated Source of Funding: Novo Nordisk A/S, Bagsvaerd, Denmark
Participants	Menopausal status: peri- and post-menopausal Age: mean 51 years, SD 4 Range = 45 to 61 years Location: Norway Ethnicity: not stated Source: gynaecological clinics Inclusion Criteria: healthy women aged 45 to 61 years, moderate to severe vasomotor symptoms, at least 3 months amenorrhoea Exclusion Criteria: abnormal bleeding, known or suspected breast or endometrial cancer, liver disease, venous thromboembolism, cardiac dysfunction, diabetes or thyroid disease, porphyria, current treatment with liver inducing medication and use of HRT of any steroids within the past 3 months Confirmation of Ovarian Failure: not stated Baseline Equality: equality reported for vasomotor symptoms, age, age at menopause, time since menopause, weight and previous HRT use Baseline Symptoms: number of participants with vasomotor symptoms at baseline not stated but inclusion criteria required moderate-severe hot flushes. The average severity score at baseline was 2.7 (0 to 3 scale).
Interventions	Rx1 (E+P, CCT): 1mg E2 and 0.25mg NETARx2 (E+P, CCT): 1mg E2 and 0.5mg NETA (Activelle, Novo Nordisk)Rx3: placeboHRT and placebo preparations not reported.Co-interventions: none reported
Outcomes	1. Hot flush frequency2. Vasomotor severity (Kupperman's Index and Vasomotor subscale of Greene's Climacteric Score))3. Hot flush weekly weighted score4. Losses/withdrawals
Notes	The author was contacted and kindly provided further data.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Baumgardner 1978

Methods	Study Design: parallel, double-blind, multicentre, placebo-control, randomised, clinical trial Randomisation: block randomisation with numerically precoded treatment packs Number of centres: 8 Duration of Trial: 24 weeks Power Calculations: none Number of women randomised: 160 Number of women analysed: 156 Intention-to-treat analysis: not clear Losses to follow-up: 4/160 = 2.5% (all exclusions because they did not conform to participant selection criteria) Withdrawals from treatment: 23 = 14.4% Compliance: not stated Source of Funding: Warner-Lambert Research Institute, Bureau of Medicine and Surgery, US Navy
Participants	Menopausal status: Peri- and post-menopausal (surgical or natural) Age: not stated Location: USA

Baumgardner 1978 *(Continued)*

	Ethnicity: not stated Source: gynaecological practice Inclusion Criteria: moderate to severe hot flushes due to oestrogen deficiency, no oral HRT within 8 weeks Exclusion Criteria: organic disease, contraindications to oestrogens including diseases of the breast, neoplasia, abnormal genital bleeding, thromboembolic disorders, impaired liver function Confirmation of Ovarian Failure: not reported Baseline Equality: equal distribution of moderate: severe hot flushes in both groups; other baseline measures not reported Baseline Symptoms: all participants had moderate-severe vasomotor symptoms at baseline (inclusion criteria)
Interventions	Rx1 (E only): quinestrol 0.1 mg/day for 1 week then 0.1 mg weeklyRx2 (E only): quinestrol 0.1 mg/day for 1 week then 0.2 mg weeklyRx3 (E only): conjugated equine oestrogens (CEE) 1.25 mg/day for 21/28 daysRx4: placeboHRT and placebo preparations were identical.Co-interventions: essential non-hormonal medications were permissibleTreatment crossovers: women who discontinued the trial beyond 1 month because of unsatisfactory relief of symptoms were placed on CEE 1.25 mg. day for 21/28 days. The number of these women were not reported.
Outcomes	1. Number of participants with moderate-severe hot flushes. 2. Withdrawals from therapy Outcomes for hot flush frequency were not suitable for meta-analysis as they were expressed as a percentage of baseline and no standard deviations were reported. Side-effect data, included uterine bleeding, nausea, breast discomfort, pruritis vulvae and headaches was not suitable for meta-analysis.
Notes	Attempts were made to contact the author with no response.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Bech 1998

Methods	Study Design: parallel, double-blind, placebo-control, randomised clinical trial Randomisation: not stated Duration of Trial: 12 months Power Calculations: none Number of women randomised: 151 Number of women analysed: 105 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from therapy: 46/151 = 30%. Breakdown: withdrawals from therapy (20), losses not specified (7), invalid data (19) Compliance: not stated Source of Funding: not stated but in part Novo Nordisk
Participants	Menopausal status: peri- and post-menopausal Age: not stated Location: Frederiksborg County, Denmark Ethnicity: not stated Source: population based recruitment Inclusion Criteria: born between 1930 and 1933, last spontaneous bleeding more than 6 and less than 24 months prior to commencement of study Exclusion Criteria: oestrogen dependent neoplasia, thromboembolic disease, liver or pancreatic disease, diabetes mellitus, severe obesity, diseases with high or low bone turnover, medications known to influence bone metabolism or provoke induction of liver enzymes Confirmation of Ovarian Failure: menopausal status evaluated by FSH but criteria not reported

Bech 1998 *(Continued)*

	Baseline Equality: there was a significantly lower hot flush severity score in placebo group at baseline. Baseline equality was reported for time since last spontaneous bleeding, body weight, smoking habits and alcohol intake Baseline Symptoms: number of women with symptoms at baseline not reported
Interventions	Rx1 (E+P CCT): 2mg oestradiol (E2) and 1 mg norethisterone acetate (NETA) for 24/8 days, Kliogest) Rx2 (E+P, cyclic): 2mg E2 days 1 to 12, 2 mg E2 + 1mg NETA days 12 to 22, 1 mg E2 days 23 to 28, Trisequens) Rx3: placebo HRT and placebo preparations were identical Co-interventions: none reported
Outcomes	1. Hot flush severity 2. Losses to follow-up 3. Quality of life (GHQ-11) Data for severity of sweats was reported separately and was not used in the analysis. There was no side-effect data reported.
Notes	In addition to vasomotor symptoms and the Kupperman's Index, the paper by Bech 1998 also reported on psychometric measures such as the Beck Depression Inventory and General Health Questionnaire. The study by Obel 1993 is apparently the same study but reports on biochemical parameters such as FHS, sex hormone-binding globulin (SHBG) and serum E2 over a longer (2 year as opposed to 12 months) follow up.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Blumel 1994

Methods	Study Design: parallel, double-blind, placebo- control, single centre, randomised clinical trial Randomisation: computer generated Duration of Trial: 6 months Power Calculation: no Number of women randomised: 50 Number of women analysed: 48 Intention-to-treat analysis: no Losses to follow-up/withdrawals from treatment: 2 (= 4%). Reasons not stated. Compliance: not stated Source of Funding: hospital
Participants	Menopausal status: peri- and post- menopausal (surgical or natural not specified) Age: mean 52.6 years, SD 4.7 (range 37 to 66) Location: Chile Ethnicity: not reported Source: clinical check-up of hospital workers Inclusion Criteria: symptoms attributable to menopause, amenorrhoea > 6 months, FSH > 40IU/ml, plasma oestradiol < 50pg/ml Exclusion Criteria: chronic illness, hormone dependent malignancies, HRT use within 6 months, use of lipid profile altering medication within 6 months Confirmation of Ovarian Failure: FSH > 40IU/ml, plasma oestradiol < 50pg/ml (inclusion criteria) Baseline Equality: similarity of age, years since menopause, body mass index, arterial systolic and diastolic blood pressure and baseline vasomotor symptoms reported Baseline symptoms: 68% of participants had vasomotor symptoms at baseline (32% mild, 36% moderate-severe)

Blumel 1994 *(Continued)*

Interventions	Rx1 (E+P, CCT): oestradiol valerate 2mg with 2.5mg medroxy progesterone acetate Rx2: placebo HRT and placebo preparations were identical in appearance.Co-interventions: none reported
Outcomes	1. Vasomotor severity score (0-3 scale, summary statistics and breakdown by severity category). 2. Number of women with hot flushes. 3. Number of women with moderate-severe hot flushes. 4. Losses to follow-up. Side-effect data not suitable for meta-analysis.
Notes	Other outcomes included lipid profiles, endometrial thickness and depression. The author was contacted and individual patient data relating to vasomotor outcomes was kindly supplied.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Campbell 1976

Methods	Study Design: cross-over, double-blind, single centre, randomised clinical trial Randomisation: not stated Duration of Trial: 12 months (6 months for each phase) Power Calculations: none Number of women randomised: 68 Number of women analysed: 61 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 7/68 = 10.3%. Reasons not stated. Compliance: not stated Source of Funding: not stated
Participants	Menopausal status: Peri- and post- menopausal Age: not stated Location: London, UK. Ethnicity: not stated Source: menopause clinic Inclusion Criteria: post-menopausal or at least 3 months between menses Exclusion Criteria: severe menopause symptoms Confirmation of Ovarian Failure: not stated Baseline Equality: not reported Baseline Symptoms: the majority of participants had vasomotor symptoms, none severe, at baseline (actual number not reported)
Interventions	Rx1 (E, high dose): conjugated equine oestrogens 1.25 mg/day for 21/28 days (Premarin)Rx2: placeboThere was no statement that HRT and placebo tablets were identical.Co-interventions: none reported
Outcomes	Outcome data for hot flush rating, losses and side-effects were not suitable for meta-analysis as not data was available for the end of the first phase of the cross-over trial.
Notes	This study reported a variety of physical and psychological symptoms attributable to the climacteric in addition to potential side effects. Results for two trials were reported. The first (Study I), consisting of participants with severe menopausal symptoms, was not eligible for this review as each treatment phase was for only 2 months. The study may not have conformed to the review protocol in that there was no stated exclusion of participants who had taken HRT within the month preceding trial commencement. The authors have not replied to our enquiries for further data.

Campbell 1976 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Chung 1996

Methods	Study Design: cross-over, double-blind, placebo-control, single centre, randomised clinical trial Randomisation: performed by the manufacturer of the treatment packs (Novo Nordisk, Bagsvaerd, Denmark). Duration of Trial: 12 months total (6 months for each phase) Power Calculation: not stated Number of women randomised: 100 Number of women analysed: 83 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 17/100 (17.0%) total. Breakdown: 1 withdrew from study (reason not stated), 6 were not compliant and were excluded, 10 failed to attend follow-up (reasons not stated). Compliance: assessed by unused pill counts and serum oestradiol. Non-compliant participants were excluded from analysis. Source of Funding: not stated.
Participants	Menopausal status: post-menopausal (surgical) Age: mean 43.8 years, SD 4.9 Location: The Prince of Wales Hospital, Hong Kong. Ethnicity: Hong Kong Chinese. Source: hospital HRT clinic Inclusion Criteria: total abdominal hysterectomy and bilateral salpingoophorectomy; no contraindications to oestrogen; no current HRT. Exclusion Criteria: lack of compliance (medication taken on average less than 25 days/month) Confirmation of Ovarian Failure: mean baseline FSH level = 78 IU/L Baseline Equality: reported for vasomotor symptoms, age, weight, height, age at menopause, years since menopause, FSH, LH and E2. E2 levels were significantly higher in the group that received HRT first. Baseline Symptoms: 66% of participants had vasomotor symptoms at baseline (37% moderate, 29% moderate-severe)
Interventions	Rx1 (E only) : oestradiol (Estrofem) 2mg/day for 28/28 daysRx2: placeboHRT and placebo tablets were identically packagedCo-interventions: none reported
Outcomes	1. Vasomotor severity score. 2. Number of women with hot flushes. 3. Number of women with moderate-severe hot flushes. Losses to follow-up and side-effect data were not suitable for meta-analysis.
Notes	This trial is a study of HRT in Asian women. Data was used from the first phase (duration = 6 months) of the cross-over trial. The author was contacted and further information supplied. The time since previous HRT was not specified and may have not entirely conformed to the study protocol (HRT within 1 month excluded).

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Conard 1995

Methods	Study Design: parallel, double-blind, placebo-control, randomised clinical trial (number of centres not stated) Randomisation: computer generated list of numbers Duration of Trial: 3 months Power Calculation: no Number of women randomised: 57 Number of women analysed: 50 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from Treatment: 7 (= 12 %). Breakdown: lack of efficacy (2 from placebo group), side effects (2 from placebo group, 2 from HRT groups), not stated (1). Compliance: not stated Source of Funding: pharmaceutical
Participants	Menopausal status: peri- and post- menopausal (all natural) Age: mean 51.8 years, SD 4.1 (range 44 to 61) Location: Paris, France. Ethnicity: not reported. Source: Hospital clinics Inclusion Criteria: amenorrhoea and hot flushes for more than 6 months, serum FSH greater than 30 IU/ml Exclusion Criteria: surgical menopause, symptoms or history of thromboembolic and/or arterial disease, general disorders, diabetes, obesity, weight fluctuations, eating disorders, hyperlipidaemia, more than 10 cigarettes per day, known or suspected cancer, hysterectomy, endometrial hyperplasia, endocrinologic disorders, no steroids or drugs that affect blood pressure, lipid or hepatic metabolism within past 2 months Baseline Equality: similarity of age, age at menopause, years since menopause, body mass index and blood pressure reported. Baseline symptoms: 100% of participants had vasomotor symptoms at baseline (93% moderate-severe).
Interventions	Rx1 (E+P, cyclic): oestradiol (E2) 1mg days 1 to 24 with 2.5mg nomegestrol acetate days 11 to 24 Rx2 (E+P, cyclic): oestradiol (E2) 1.5 mg days 1 to 24 with 3.75 mg nomegestrol acetate days 11 to 24. Rx3: placebo The similarity of HRT and placebo preparations was not reported. Co-interventions: none reported
Outcomes	1. Daily hot flush frequency (diurnal and nocturnal). 2. Vasomotor severity score 3. Number of women with hot flushes. 4. Losses to follow-up. Side-effect data not suitable for meta-analysis.
Notes	The primary outcomes of this study related to cardiovascular risk factors. The exclusion of HRT within one month of study commencement was not explicitly stated, however the exclusion criteria did excludes taking of any lipid altering drugs (presumably including HRT) within 2 months prior to the study. The author was contacted and individual patient data relating to vasomotor outcomes was kindly supplied.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Coope 1975

Methods	Study Design: cross-over, double-blind, single centre, randomised clinical trial Randomisation: precoded treatment packs generated by a random numbers scheme Duration of Trial: 6 months (3 months for each cross-over phase) Power Calculations: none Number of women randomised: 35 Number of women analysed: 30 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 5/35 = 14.3% Compliance: not stated but one patient excluded for lack of compliance Source of Funding: Macclesfield Hospital, Withington Hospital, MRC, Ayerst Laboratories
Participants	Menopausal status: peri- and post-menopausal (surgical and natural) Age: mean 52 years (range 40 to 61) Location: Cheshire UK Ethnicity: not stated Source: general practice in rural and industrial England Inclusion Criteria: menopause symptoms persisting for more than 6 months, Exclusion Criteria: history of thromboembolism, breast and genital cancer Confirmation of Ovarian Failure: FHS levels not measured but the inclusion criteria specified that menopause symptoms must have persisted over 6 months. Baseline Equality: number of hot flushes were higher in placebo group at baseline (mean of 56 compared to 39) but this was not significant (P = 0.26). Baseline equality of other parameters was not reported. Baseline Symptoms: 27/30 (90%) of participants had hot flushes at baseline
Interventions	Rx1 (E only): conjugated equine oestrogens 1.25 mg/day for 21/28 days (Premarin)Rx2: placeboHRT and placebo preparations were identical in appearanceCo-interventions: none reported
Outcomes	1. Hot flush frequency 2. Number of women with hot flushes Side-effects and withdrawals were not reported for first phase of cross-over trial.
Notes	There was no specific exclusion of HRT within one month of study commencement however this is implied in that participants were monitored for persistence of menopausal symptoms over 6 months prior to the trial commencement. The primary outcomes of this trial were climacteric symptoms (Kupperman's Index). Other outcomes reported were vaginal cytology (karyotypic index) and outcomes relating to blood clotting. The author was contacted and kindly provided further data (including some individual patient data).

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Coope 1981

Methods	Study Design: cross-over, double-blind, single centre, randomised clinical trial Randomisation: precoded treatment packs generated by a random numbers scheme Duration of Trial: 14 months (6 months for each treatment, with an intervening 2 months "washout" period Power Calculations: none Number of women randomised: 66 Number of women analysed: 55 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 11/66 = 16.7% Compliance: not stated

Coope 1981 *(Continued)*

	Source of Funding: Abbot Pharmaceuticals, Withington Hospital Manchester
Participants	Menopausal status: peri-menopausal and post-menopausal (surgical and natural) Age: mean 48 years Location: England Ethnicity: not stated Source: recruited from a semi-rural general practice Inclusion Criteria: menopausal women aged 40 to 60 suffering from depression Exclusion Criteria: contraindications to oestrogen therapy such as thrombosis, breast or genital cancer or ischaemic heart disease, severe depression or suicidal tendencies Confirmation of Ovarian Failure: FSH measured at baseline (mean 45 IU/L) Baseline Equality: baseline equality reported for hot flush frequency in addition to age, natural menopause vs surgical menopause, proportions still menstruating, depression scores, well being assessment, FSH and serum oestrone levels. Baseline Symptoms: not stated
Interventions	Rx1(E only): piperazine oestrone sulphate 1.5 mg/day for 21/28 daysRx2: placeboHRT and placebo preparations were identical in appearance.Co-interventions: psychotropic drugs were prescribed but usage was "kept to a minimum". Women were asked not to take aspirin, phenylbutazone or any drugs that could affect blood clotting tests.
Outcomes	1. Hot flush frequency/week Side-effect data were not reported for first phase of cross-over trial.
Notes	The primary outcomes of this study related to depression as assessed by the Beck depression inventory. Other outcomes (reported more fully in the second study) related to biochemical parameters and coagulation outcomes. The author was contacted and kindly supplied further information.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Davidsen 1974

Methods	Study Design: double-blind, cross-over, single centre, randomised clinical trial Randomisation: not stated Duration of Trial: 6 months (3 months each phase) Power Calculations: none Number of women randomised: 32 women to two treatment groups Number of women analysed: not clear (only analysed patients with each symptom at baseline) Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: not stated Compliance: not stated Source of Funding: not stated
Participants	Menopausal status: peri-menopausal Age: 38 to 56 years (mean 47) Location: Hillerod, Denmark. Ethnicity: not stated Source: gynaecological department Inclusion Criteria: age 38 to 56 years, metrorrhagia, normal gynaecological examination, no HRT within 1 month Exclusion Criteria: not defined Confirmation of Ovarian Failure: not stated Baseline Equality: not stated Baseline Symptoms: 25/32 participants had hot flushes, 21/32 had perspiration and 13/32 had palpitations at baseline

Davidsen 1974 *(Continued)*

Interventions	Rx1 (E+P, cyclic): 2mg oestradiol valerate days 1 to 11, 2 mg oestratiol valerate + 0.5 mg norgestrel days 12 to 21 (Cycloprogynova) Rx2: placebo The similarity of HRT and placebo was not reported. Co-interventions: none reported
Outcomes	Data for number of women with flushes, perspiration and palpitations were not suitable for meta-analysis as there was not data for the end of the first phase of the cross-over and only women with symptoms at baseline were analysed for each of these outcomes. Similarly, side-effect data for nausea, breast tenderness, blood pressure and bleeding patterns was not available for the end of the first cross-over phase and losses to follow-up data was not reported.
Notes	Other outcomes for the trial included psychological symptoms. Several unsuccessful attempts were made to contact the authors.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Dennerstein 1978

Methods	Study Design: double-blind, cross-over, single centre, randomised clinical trial Randomisation: method not stated Duration of Trial: 12 months (3 months for each of the four treatment groups) Power Calculations: none Number of women randomised: 50 to four treatment groups Number of women analysed: 36 women completed the trial of which 32 were analysed for hot flush outcomes Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 14 (= 28%). This includes one woman with remaining functioning ovarian tissue excluded post enrolment. The remaining withdrawals from treatment were predominantly due to psychological symptoms. Compliance: not stated Source of Funding: grant-in-aid from Wyeth Pharmaceuticals Pty Ltd, Allen and Hanburys and Organon (Australia Pty Ltd)
Participants	Menopausal status: post-menopausal (all surgical) Age: mean 46.2 years, SD 8.9 Location: Melbourne, Australia Ethnicity: not stated Source: both private patients involved in a prior study of the sequelae of hysterectomy and bilateral salpingo-oophorectomy and newspaper advertisement Inclusion Criteria: both hysterectomy and bilateral salpingo-oophorectomy, age < 65 years, no contraindications to HRT, stable, heterosexually active relationship Exclusion Criteria: not stated Confirmation of Ovarian Failure: not stated, but all surgical menopause Baseline Equality: not stated Baseline Symptoms: not stated
Interventions	Rx1: (E only): ethinyl oestradiol 50 mg/dayRx2 (E+P, CCT): ethinyl oestradiol 50 mg/day, levenorgestrel 250 mg/day (Nordiol, Wyeth)Rx3: placeboThe tablets were identical in appearance.Co-interventions: none reportedTreatment cross-overs: 12 women changed drug regimens due to intolerable side-effects. The majority of these were in the placebo group due to lack of control of vasomotor symptoms.

Dennerstein 1978 *(Continued)*

Outcomes	Hot flush frequency, intensity and losses and side-effect data were not suitable for meta-analysis as there was no data available for the end of the first cross-over phase.
Notes	The primary outcomes of this study related to psychological symptoms. Participants were requested not to take HRT within 2 weeks of the study commencement and therefore this study may not entirely conform to the review protocol which required no HRT within 1 month of the study commencement. A progestogen only treatment arm was excluded from this review. The author was contacted and kindly supplied further information, but there was no further data to enable inclusion in the meta-analysis.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Derman 1995

Methods	Study Design: parallel, double-blind, multicentre, placebo-control, randomised clinical trial Randomisation: treatments assigned symmetrically in blocks of four according to a computer-generated randomisation scheme Number of Centres: 3 Duration of Trial: 16 weeks Power Calculations: no Number of women randomised: 82 to two treatment groups Number of women analysed: varied according to outcome; 70 for hot flush frequency, and 78 for hot flush severity Intention-to-treat analysis: yes Losses to follow-up: varied according to outcome; 12/82 (= 15% for hot flush frequency) and 4/82 (= 5%) for hot flush severity Withdrawals from treatment: 35/82 (= 43%) did not complete treatment. Breakdown: adverse effects (6), lack of effect (18), unspecified (11). Withdrawal rate was significantly higher in placebo group. Compliance: assessed by unused pill counts but degree of compliance not stated Source of Funding: Novo Nordisk Pharmaceuticals
Participants	Menopausal status: peri- and post-menopausal (all natural) Age: 50 years (mean) Location: USA Ethnicity: not stated Source: not stated Inclusion Criteria: age 40 to 60 years, menopause symptoms with at least 20 vasomotor events/week (and a minimum of five moderate-severe), serum FSH > 40 IU/ml Exclusion Criteria: estrogen therapy within the last month, steroid therapy within the past 3 months, a history of major diseases that would contraindicate oestrogen therapy, long term treatments that would interfere with outcomes Confirmation of Ovarian Failure: serum FSH > 40 IU/ml (inclusion criteria) Baseline Equality: equality for vasomotor symptoms, general climacteric symptoms and Beck Depression Index reported Baseline Symptoms: all participants had vasomotor symptoms (>20/week) at baseline (inclusion criteria)
Interventions	Rx1 (E+P, cyclic): 2mg 17 beta oestradiol days 1 to 12, 2 mg 17 beta oestradiol + 1mg norethisterone acetate days 13 to 22, 1mg 17 beta oestradiol days 23/28 (triphasic sequential therapy) Rx2: placebo HRT and placebo preparations were identical Co-interventions: none reported
Outcomes	1. Hot flush frequency 2. Hot flush severity (vasomotor component of Greene's Climacteric Scale, 0-12 scale)

Derman 1995 *(Continued)*

	3. Losses to follow-up
Notes	Attempts were made to contact the author but were unsuccessful.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Hagen 1982

Methods	Study Design: parallel, double-blind, placebo- control, single centre randomised clinical trial Randomisation: random sampling numbers but concealment of allocation not reported Duration of Trial: 24 months Power Calculation: none Number of women randomised: 119 into four treatment groups (two of which were excluded from this review) Number of women analysed: 97 from four treatment groups (two of which were excluded from this review) Intention-to-treat analysis: no Losses to follow-up: 22/119 = 18% over four groups. Reasons: personal reasons (14), moved (3), withdrawal from treatment (5). Withdrawals from treatment: 5 over the four groups. Reasons: breast cancer (3, two received HRT), cardiac/pulmonary disease (2, neither received HRT). Compliance: not stated Source of Funding: includes Leo Pharmaceuticals, Novo Industri A/S
Participants	Menopausal status: peri- and post-menopausal (all natural) Age: mean 50 years (range 44 to 54) Location: Glostrup, Denmark. Ethnicity: not stated. Source: population sample selected by questionnaire Inclusion Criteria: age 45 to 54 years, spontaneous cessation of menstrual periods within last 6 months to 3 years, no treatment with gonadal hormones, thiazides or other drugs known to influence calcium metabolism after menopause Exclusion Criteria: any gynaecological operation, elevated blood pressure, abnormal blood chemistry, pathological cervical smear, and past or present contra-indications to HRT or thiazides Confirmation of Ovarian Failure: baseline FSH = 78 IU/L and LH = 24 IU/L Baseline Equality: reported for age, time since menopause, FSH, LH Baseline Symptoms: 61% of participants had hot flushes at baseline and overall menopause symptoms were mild to moderate as judged by the Kupperman's Index.
Interventions	Rx1 (E+P, cyclic): Trisequens Forte - 4mg 17 beta oestradiol + 2mg oestriol days 1 to 12; 4mg 17 beta oestradiol + 2mg oestriol + 1mg norethisterone acetate days 13 to 22; 1 mg 17 beta oestradiol + 0.5 mg oestriol days 23 to 28 Rx2: placebo: (the two treatment arms including thiazide were excluded from this review) Identity of HRT and placebo preparations not reported. Co-interventions: 500mg calcium daily (Calcium Sandoz) for all participants
Outcomes	Relevant outcomes such as the number of women with vasomotor symptoms, mean vasomotor severity score, losses to follow-up and side-effect outcomes not suitable for meta-analysis as the data for placebo and HRT groups was combined the comparative thiazide groups.
Notes	This study was part of a larger study with primary outcomes relating to bone loss. Participants were recruited from a population base and inclusion criteria did not include menopause symptoms. Attempts were made to contact the author for further data but no response was obtained.

Hagen 1982 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Jensen J 1983

Methods	Study Design: parallel, placebo-control, single centre, randomised clinical trial Randomisation: random sampling numbers, method of allocation concealment not reported Duration of Trial: 12 months Power Calculation: No Number of women randomised: 100 Number of women analysed: 87 Intention-to-treat analysis: no Losses to follow-up (including withdrawals from treatment): 13/100 = 13%. Reasons: moved or not specified (6), oestrogen related side effects (4) and other illness (3). Compliance: not assessed Source of Funding: grants from Fabrikant Einar Willumsens Mindelegat and Illum Fondet. Medication provided by Novo Research Institute and Sandoz Pharmaceuticals.
Participants	Menopausal status: post-menopausal (all natural) Age: mean 51.5 years (range 46 to 55 years) Source: population sample selected by questionnaire and a medical screening examination Location: Glostrup, Denmark Ethnicity: not stated. Inclusion Criteria: natural menopause 2.5 to 5 years previously, no HRT treatment since menopause, free of diseases and/or medication known to influence study outcomes Exclusion Criteria: not further defined Confirmation of Ovarian Failure: mean FSH level at baseline = 84 IU/L Baseline Equality: equality reported for relevant outcomes, age, menopausal age, Kupperman's Index, S-oestradiol, FSH, cholesterol, bone mass, weight and height Baseline Symptoms: 62% of participants had hot flushes at baseline and overall menopause symptoms were considered mild (as judged by Kupperman's Index).
Interventions	Rx1 (E+P, cyclic): Trisequens Forte - 4mg 17 beta oestradiol + 2mg oestriol days 1 to 12; 4mg 17 beta oestradiol + 2mg oestriol + 1mg norethisterone acetate days 13 to 22; 1 mg 17 beta oestradiol + 0.5 mg oestriol days 23 to 28 Rx2 (E+P, cyclic): Trisequens - 2mg 17 beta oestradiol + 1mg oestriol days 1 to 12; 2mg 17 beta oestradiol + 1mg oestriol + 1mg norethisterone acetate days 13 to 22; 1 mg 17 beta oestradiol + 0.5 mg oestriol days 23 to 28 Rx3 (E+P, cyclic): Trisequens Mite - 1mg 17 beta oestradiol days 1 to 12; 1mg 17 beta oestradiol + 1mg norethisterone acetate days 13 to 22 Rx4: placebo HRT and placebo were identical in appearance. Co-interventions: 500mg calcium daily (Calcium Sandoz) for all participants
Outcomes	1. Number of women with hot flushes. 2. Losses to follow-up. Hot flush severity recorded as part of Kupperman Index but not reported. Side-effect data included "oestrogen related" withdrawal from therapy, any bleeding, irregular bleeding and weight changes but was not suitable for meta-analysis.
Notes	This study was part of a larger study with primary outcomes relating to bone loss. Participants were recruited from a population base and inclusion criteria did not include menopause symptoms. The Jensen 1983 reference reports the results of four studies in which participants overlapped in sequential studies. Only Study I and Study II were eligible for inclusion in this review (Study III had no HRT therapy

Jensen J 1983 *(Continued)*

and Study IV participants had received HRT within 1 month). The Study II participants were the placebo group from Study I. According to the review protocol, only one such study is eligible for inclusion in the review and as vasomotor outcome data was only available for Study II, this was selected for inclusion. Study II was apparently the same study as reported by Christensen 1982. There were some discrepancies between the two papers in terms of the number of patients initially randomised, and subsequently, losses to follow up. The data from Christensen 1982 was used in this instance as this paper contained more detailed information. Neither of the studies contained a clear reference to the study being double blinded although this was inferred.Attempts were made to contact the author(s) but no response was obtained.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Jensen P 1987

Methods	Study Design: parallel, double-blind, placebo-control, randomised clinical trial Randomisation: method not stated Duration of Trial: 2 years Power Calculations: none Number of women randomised: 133 total to four treatment groups; 76 to the two treatment groups relevant to this review Number of women analysed: 57 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 19/76 = 25% Compliance: not stated but serum 17 beta oestradiol and oestrone levels monitored throughout therapy Source of Funding: not stated
Participants	Menopausal status: peri- and post-menopausal (all natural) Age: mean 49.8 years, SD 2.1 Location: Copenhagen County, Denmark Ethnicity: not stated Source: population based recruitment Inclusion Criteria: healthy women who had undergone a natural menopause 6 months to 3 years prior to the start of the study Exclusion Criteria: previous HRT or medications known to influence calcium metabolism, Kupperman's Index of less than 3.7 Confirmation of Ovarian Failure: not stated. Mean serum 17 beta oestradiol and oestrone levels measured at baseline. Baseline Equality: equality of vasomotor symptoms at baseline; equality also reported for age, weight, menopausal age, Kupperman Index and oestrone. Serum levels of 17 beta oestradiol were marginally lower in the placebo group at baseline (P = 0.10). Baseline Symptoms: 89% of participants had hot flushes at baseline.
Interventions	Rx1 (E+P, cyclic): 2 mg oestradiol valerate days 1 to 11, 2mg oestradiol valerate + 1 mg cyproterone acetate days 12 to 21, no treatment days 22 to 28) Rx2: placebo The similarity of HRT and placebo was not reported. Co-interventions: none reported
Outcomes	1. Number of women with hot flushes. 2. Losses/withdrawals Side-effect data was not reported.
Notes	Participants were recruited for a larger study based on outcomes related to osteoporosis. This study reported specifically on climacteric complaints and contained four treatment groups of which only two

Jensen P 1987 *(Continued)*

were oral therapy and therefor included in this review. Attempts made to contact the author were unsuccessful.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Marslew 1992

Methods	Study Design: parallel, double blinded, single centre, placebo controlled, randomised clinical trial Randomisation: method not stated Duration of Trial: 24 months Power Calculations: none Number of women randomised: 50 Number of women analysed: 39 Intention to Treat analysis: no Losses to follow-up/withdrawals from treatment: 11/50 = 22%. Reasons not specified. Compliance: not assessed Source of Funding: Danish Medical Research Council, Schering A/S
Participants	Menopausal status: peri- and post-menopausal (all natural) Age: mean 51 years, SD 2 Location: Glostrup, Denmark Ethnicity: not stated Source: population based recruitment Inclusion Criteria: age 45 to 54 years, natural menopause 6 months to 3 years prior to the study, free from past and present diseases, contraindications to HRT, lack of medications known to influence outcomes studied Exclusion Criteria: not stated Confirmation of Ovarian Failure: not stated Baseline Equality: equality reported for vasomotor symptoms, Kupperman's Index, age, postmenopausal duration and serum E2 at baseline Baseline Symptoms: 90% of participants had hot flushes at baseline
Interventions	Rx1 (E+P, CCT): 2 mg oestradiol valerate and 1 mg cyproterone acetate Rx2: placeboHRT and placebo preparations were identical in appearance.Co-interventions: none reported
Outcomes	1. Number of women with hot flushes. 2. Losses to follow-up Hot flush severity score and side-effect data were not suitable for meta-analysis.
Notes	This study was part of a larger population based study of HRT. Other outcomes included climacteric symptoms, effects on bone, calcium and lipid metabolism, bleeding and adverse effects attributable to progestogen. There was a third treatment group (HRT) in the primary reference which was excluded from this review as it was apparently not double blinded (the tablets were different to the placebo group). The secondary references also report on other treatment groups which were excluded from the review as they were single blinded.Attempts were made to contact the author but no response was obtained.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Martin 1971

Methods	Study Design: parallel, double-blind, placebo-control, randomised clinical trial Randomisation: not stated Number of Centres: 2 Duration of Trial: 3 months Power Calculations: none Number of women randomised: 165 Number of women analysed: 120 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 45/165 = 27%. Eight women withdrew in initial washout phase and 37 during the double-blind phase. The reasons during the double-blind phase were lack of symptom control: 26 (24 placebo group); bleeding problems: 2; personal reasons: 9. Compliance: not stated Source of Funding: not stated
Participants	Menopausal status: peri-menopausal and post-menopausal (both surgical and natural) Age: majority in the age range 45 to 55 years Location: California, USA. Ethnicity: not stated Source: private practice Inclusion Criteria: hot flushes with and without other climacteric symptoms Exclusion Criteria: no vasomotor symptoms, gynaecologic or breast carcinoma, cardiovascular disease, thrombophlebitis, blood dyscrasias or other serious illnesses Confirmation of Ovarian Failure: not stated Baseline Equality: equality of baseline severity of vasomotor symptoms; other baseline equality measures not reported Baseline Symptoms: 100% of women had vasomotor symptoms at baseline (inclusion criteria)
Interventions	Rx1 (E+P, graded low dose): 12.5 mcg mestranol (MEE) days 1 to 5, 25mcg MEE days 6 to 13, 50mcg MEE days 14 to 15, 25 mcg MEE + 1mg norethindrone (NET) days 16 to 18, 30 mcg MEE + 1.5 mg NET days 19 to 24, 20 mcg MEE + 0.75 MEE + 0.75 mg NET days 25 to 28 Rx2 (E+P, graded high dose): 25 mcg mestranol (MEE) days 1 to 5, 50mcg MEE days 6 to 13, 100mcg MEE days 14 to 15, 50 mcg MEE + 1mg norethindrone (NET) days 16 to 18, 60 mcg MEE + 1.5 mg NET days 19 to 24, 40 mcg MEE + 0.75 MEE + 0.75 mg NET days 25 to 28 Rx3: placebo HRT and placebo were identical in appearance. Co-interventions: none reported
Outcomes	1. Losses to follow-up Hot flush frequency and severity data was not suitable for meta-analysis as no standard deviations were supplied. Side-effect data included nervousness, headaches, breast soreness and bleeding patterns but was also not suitable for meta-analysis.
Notes	The participants in this study were selected for vasomotor symptoms. The majority of participants had been treated with HRT prior to the study commencement and therefore the study design included an initial 8 week placebo "wash-out" phase for all women. There was a high withdrawal rate from therapy in the placebo group attributed to lack of symptom control. Analysis confirmed that the women who withdrew from the placebo group had more severe hot flushes at baseline than those who remained, introducing a bias against the treatment effect. Attempts to contact the author were unsuccessful.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Unclear risk	B - Unclear

Notelovitz 2000a

Methods	Study Design: parallel, double-blind, multicentre, placebo-controlled, randomised clinical trial Randomisation: block randomisation within each centre Number of Centres: 15 Duration of Trial: 3 months Power Calculations: yes (based on moderate to severe hot flush outcome) Number of women randomised: 333 into 1 of 5 treatment groups Number of women analysed: 280 for moderate/severe hot flushes & 324 for hot flush weekly weighted score Intention-to-treat analysis: no for moderate/severe hot flushes, yes for hot flush weekly weighted score. Losses to follow-up/withdrawals from treatment: 53/333 = 16% for moderate/severe hot flushes and 4/333 = 1.2% for hot flush weekly weighted score Compliance: not stated Source of Funding: supported by grants to participating institutions and Novo Nordisk Pharmaceuticals Inc
Participants	Menopausal status: peri- and post-menopausal Age: mean 54.12 ± 4.14 years (mean ± SD) (range 40 to 60 years) Location: USA Ethnicity: Race (n, %) in placebo, 0.25 mg E2, 0.5 mg E2, 1.0 mg E2, 2.0 mg E2 respectively: White - 60 (91%), 62 (91%), 61 (95%), 57 (85%), 62 (91%) Black - 3 (5%), 3 (4%), 2 (3%), 5 (7%), 1 (1%) Hispanic - 1 (2%), 2 (3%), 1 (2%), 4 (6%), 3 (4%) Asian/Pacific - 2 (3%), 0 (0%), 0 (0%), 1 (1%), 2 (3%) Other - 0 (0%), 1 (1%), 0 (0%), 0 (0%), 0 (0%) Source: study population was obtained from the investigators´ sites or through local advertising (i.e. a mixture of clinical and general population) Inclusion Criteria: menopause symptoms persisting for more than 6 months, healthy menopausal women with an intact uterus, 40-60 years old, at least 56 moderate-severe hot flushes/week, at least 6 months amenorrhoea, E2 levels <= 20pg/mL, FSH >= 50 IU/L Exclusion Criteria: history of endometrial hyperplasia, abnormal bleeding of unknown origin, endometrial thickness at least 5mm, history of estrogen-dependent tumours, gallbladder, liver kidney or endocrine diseases except controlled thyroid disease, venous thromboembolism, cerebrovascular accidents, myocardial infarction or ischaemic heart disease, history of severe headache or migraines, high blood pressure, alcohol or drug abuse, smoking > 15 cigarettes/day, weight increased more than 20% over ideal body weight, use of steroid hormones/drugs known to influence estrogen metabolism & use of HRT within 2 months prior to randomisation Confirmation of Ovarian Failure: at least 6 months amenorrhoea, E2 levels <= 20pg/mL, FSH >= 50 IU/L Baseline Equality: matched for age, time of amenorrhoea, weight, baseline hot flush symptoms Baseline Symptoms: At least 56, with 72 ± 21 (mean ± SD) moderate-severe hot flushes/week & mean hot flush weekly weighted score 183 ± 61(mean ± SD).
Interventions	Rx1 (E, low dose): micronized 17b-oestradiol 0.25 mg/day Rx2 (E, low dose): micronized 17b-oestradiol 0.5 mg/day Rx3 (E, moderate dose): micronized 17b-oestradiol 1.0 mg/day Rx4 (E, high dose): micronized 17b-oestradiol 2.0 mg/day Rx5: placebo The HRT and placebo preparations were identical in appearance Co-interventions: none reported
Outcomes	1. Weekly hot flush frequency 2. Hot flush weekly weighted score 3. Withdrawals from therapy 4. Adverse event frequency
Notes	The author was contacted and supplied further information.
Risk of bias	

Notelovitz 2000a *(Continued)*

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Paterson 1982a

Methods	Study Design: cross-over, double-blind, single centre, placebo-control, randomised clinical trial Randomisation: random numbers with numerically pre coded treatment packs Duration of Trial: 6 months (3 months each phase) Power Calculations: none Number of women randomised: 23 Number of women analysed: 20 Intention-to-treat analysis: no Losses to follow-up/Withdrawals from treatment: 3/23 = 25%. Compliance: not stated Source of Funding: Syntex
Participants	Menopausal status: peri- and post- menopausal (post-menopausal both surgical and natural) Age: 47 years (range 34 to 59) Location: Birmingham UK Ethnicity: not stated Source: menopause clinic Inclusion Criteria: climacteric symptoms, hysterectomy at least one year previously, FSH > 20 IU/ml Exclusion Criteria: no previous uterine malignance Confirmation of Ovarian Failure: FSH > 20 IU/ml Baseline Equality: baseline equality reported for vasomotor and overall menopausal symptoms, age Baseline Symptoms: the majority of participants had moderate-severe hot flushes at baseline (mean severity score = 2.7)
Interventions	Rx1 (E + P, graded sequential): mestranol 5 mg/day for 5 days, mestranol 25 mg/day for 8 days, mestranol 50 mg/day for 2 days, mestranol 25 mg/day + norethisterone 1 mg/day for 3 days, mestranol 30 mg/day + norethisterone 1.5 mg/day for 6 days, mestranol 20 mg/day + norethisterone 0.75 mg/day for 4 days (Syntex Menophase)Rx2: placeboHRT and placebo preparations were identical.Co-interventions: none
Outcomes	1. Hot flush severity score Incomplete losses data not suitable for meta-analysis. Side-effects data summarised over both phases of cross-over trial.
Notes	The author was contacted and supplied further information.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

PEPI 1998

Methods	Study Design: parallel, double-blind, multicentre, placebo- control trial Randomisation: computer generated, variable length, blocked randomisation scheme in which treatment assignment was stratified by clinical centre and hysterectomy status. Number of Centres: 7 Duration of Trial: 36 months Power Calculation: yes (based on HDL-cholesterol primary outcome)

PEPI 1998 *(Continued)*

	Number of women randomised: 875 to five treatment arms; 175 to Arm 1 and 174 to Arm 5 (Placebo) Number of women analysed: 846 (Total); 170 for Arm 1 and 166 for Arm 5 (Placebo) Intention-to-treat analysis: yes Losses to follow-up: 29 (= 3.3%) Total; Arm1 = 5/175 (2.9%); Arm5 (placebo) = 8/174 (4.6%) Withdrawals from Treatment: 210 Total. Breakdown by Rx not reported. Reasons were: protocol mandated (51), symptoms (127), concerns regarding health risks (11) and personal circumstances (21). Compliance: assessed by unused pill counts. 612 (=70%) women (Total) were classified as adherent as defined by taking at least 80% of pills during the 6 months before each annual visit. Breakdown by Rx not reported. Source of Funding: NIH, NHLBI, NIHCD, NIAMSD, NIDDK, NIA. Medication for the trial was supplied by Wyeth-Ayerst Laboratories, Schering-Plough Research Institute and the Upjohn Company.
Participants	Menopausal status: post-menopausal (both surgical and natural) Age: 45 to 64 years at time of first visit. Overall mean age 56.1 years, SD 4.3. Breakdown by Rx not reported. Location: USA. Sites were: George Washington University, Washington DC; The Johns Hopkins University, Baltimore; Stanford University, California; The University of California, Los Angeles, California; The University of California, San Diego, California; The University of Iowa, Ames, Iowa; The University of Texas Health Science Centre, San Antonio, Texas. Source: population sample obtained through media and community based approaches Ethnicity: overall 89% white, 5% Hispanic, 4% black, 2% Asian. Breakdown by Rx not reported. Inclusion Criteria: aged 45 to 64 years at recruitment; good health; last menstrual period between 1 to 10 yrs ago (natural menopause only); more than 2 months post-hysterectomy and FSH more than 40 IU/L (surgical menopause only). Exclusion Criteria: severe menopause symptoms; any HRT within 3 months; serious illness; if on thyroid hormone, medication stable for more than 3 months and normal TSH levels; contraindications to estrogen. Further exclusions based on LDL-C, triglycerides, BMI, blood pressure and fasting glucose. Criteria for confirming menopausal status: not reported (apart from Inclusion Criteria). Baseline Equality: women randomised to the five trial arms had similar sociodemographic, lifestyle and menopause-related characteristics and most primary outcome variables and important secondary outcome variables. Some significant differences were observed between the treatment arms in LDL-C and fibrinogen at baseline. Baseline symptoms: 52.5% of participants had vasomotor symptoms at baseline
Interventions	Rx1 (E only): CEE (Premarin) 0.625mg/day for 28/28 days Rx2 (E + P, cyclic): conjugated equine oestrogen (Premarin) 0.625mg/day for 28/28 days; medoxyprogesterone acetate 10mg/day for days 1 to 12 Rx3 (E + P, CCT): conjugated equine oestrogen (Premarin) 0.625mg/day + medoxyprogesterone acetate 2.5 mg/day for 28/28 days Rx4 (E + P, cyclic): conjugated equine oestrogen (Premarin) 0.625mg/day for 28/28 days; micronized progesterone 200 mg/day for days 1 to 12 Rx5: placebo The HRT and placebo preparations were identical in appearance Co-interventions: by the end of the study, 19 women in the placebo group (11%) had begun taking privately prescribed hormones
Outcomes	1. Number of women with any vasomotor symptoms (defined as hot flashes, night sweats, cold sweats). 2. Withdrawals from therapy. 3. Side-effects
Notes	The PEPI Investigators were contacted and kindly supplied further data for the analysis.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Symons 2000 Study 1

Methods	Study Design: parallel, double-blind, multicentre, placebo-controlled, randomised clinical trial Randomisation: computer generated randomisation and treatment allocation codes Number of Centres: 11 Duration of Trial: 4 months Power Calculations: yes (based hot flush frequency outcome) Number of women randomised: 219 into 5 treatment groups Number of women analysed: 187 for hot flush frequency Intention-to-treat analysis: yes, 'modified' ITT Losses to follow-up/withdrawals from treatment: not clear (31 women did not complete the study) Compliance: assessed by medication return Source of Funding: Parke-Davis Pharmaceutical Research Division, Warner Lambert Company.
Participants	Menopausal status: post-menopausal Age: mean 51.58 ± 0.58 years (mean ± SD) (range 41 to 65 years) Location: USA Ethnicity: Not provided when requested Source: not clear if from a clinical or general population Inclusion Criteria: ammenorrheic for at least one year, but not more than 5 years Exclusion Criteria: history of any chronic disease, current vaginal bleeding or endometrial hyperplasia, any contraindications for HRT & use of HRT within 3 months prior to randomisation Confirmation of Ovarian Failure: 6-12 months amenorrhoea, E2 levels <= 25pg/mL, FSH >= 50 mIU/mL Baseline Equality: matched for age, time since last menstrual period, hot flush frequency and smoking history Baseline Symptoms: at least 10 hot flushes/week & an average of 48 hot flushes in the week prior to randomisation
Interventions	Rx1 (E + P, low dose): ethinyl estradiol 1 µg + norethindrone acetate 0.2 mg Rx2 (E + P, low dose): ethinyl estradiol 2.5 µg + norethindrone acetate 0.5 mg Rx3 (E + P, medium dose): ethinyl estradiol 5 µg + norethindrone acetate 1 mg Rx4 (E + P, high dose): ethinyl estradiol 10 µg + norethindrone acetate 1 mg Rx5: placebo The HRT and placebo preparations were identical in appearance Co-interventions: none reported
Outcomes	1. Weekly hot flush frequency 2. Withdrawals from therapy 3. Adverse event frequency
Notes	The author was contacted and supplied further information.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Allocation concealment?	Low risk	A - Adequate

Symons 2000 Study 2

Methods	Study Design: parallel, double-blind, multicentre, placebo-controlled, randomised clinical trial Randomisation: computer generated randomisation and treatment allocation codes Number of Centres: 24 Duration of Trial: 3 months Power Calculations: yes (based on hot flush frequency outcome) Number of women randomised: 266 into 4 treatment groups Number of women analysed: 261 for hot flush frequency

Symons 2000 Study 2 *(Continued)*

	Intention-to-treat analysis: yes Losses to follow-up/withdrawals from treatment: not clear (36 women did not complete the study) Compliance: subjects recorded tablet intake in a diary (non-compliance was defined as 3 consecutive days where medication was missed) Source of Funding: Parke-Davis Pharmaceutical Research Division, Warner Lambert Company.
Participants	Menopausal status: peri- and post-menopausal Age: mean 51.12 ± 4.14 years (mean ± SD) (range 40 to 62 years) Location: USA Ethnicity: not provided when requested Source: not clear if from a clinical or general population Inclusion Criteria: within 5 years of surgical or natural menopause, at least 56 moderate to severe hot flushes in the week prior to randomisation Exclusion Criteria: amenorrhoeic for less than 6 months, history of any chronic disease, any contraindications for HRT & use of HRT within 2 months prior to randomisation Confirmation of Ovarian Failure: amenorrhoea for at least 1 year or amenorrhoea for at least 6-12 months & serum estradiol <= 25 pg/mL, FSH >= 50 mIU/mL Baseline Equality: matched for age, time since last menstrual period, hot flush frequency and smoking history Baseline Symptoms: at least 56 hot flushes/week & an average of over 80 hot flushes in the week prior to randomisation
Interventions	Rx1 (E + P, low dose): ethinyl estradiol 2.5 µg + norethindrone acetate 0.5 mg Rx2 (E + P, medium dose): ethinyl estradiol 5 µg + norethindrone acetate 1 mg Rx3 (E + P, high dose): ethinyl estradiol 10 µg + norethindrone acetate 1 mg Rx4: placebo The HRT and placebo preparations were identical in appearance Co-interventions: none reported
Outcomes	1. Weekly hot flush frequency 2. Withdrawals from therapy 3. Adverse event frequency
Notes	The author was contacted and supplied further information.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Viklylaeva 1997

Methods	Design of Study: parallel, double-blind, multicentre, placebo-control trial Randomisation: block randomisation with random block size between 2 and 8. The initial degree of general symptoms and centres were used as stratification factors. Number of Centres: 3 Duration of Trial: 24 weeks Power Calculation: not stated Number of women randomised: 64 Total, 33 in Rx group and 31 in placebo group Number of women analysed: 60 Total, 32 in Rx group and 28 in placebo group Intention-to-treat analysis: not clear Losses to follow-up: 4 (= 6.3%) Total, 1 in Rx group and 3 in placebo group. Reasons not stated. Withdrawals from Treatment: 2 in Rx group due to moderate adverse effects (nature not specified). Relationship between withdrawals from treatment and losses to follow up not clarified. Compliance: not stated Source of Funding: Norwegian Ministry of FA, hospital budget , pharmaceutical (Novo Nordisk)

Viklylaeva 1997 *(Continued)*

Participants	Menopausal status: peri-menopausal Age: range 39 to 56 years Location: Moscow, Russia Source of patients: not stated Ethnicity: not stated Inclusion Criteria: moderate-severe menopausal symptoms with desire for treatment, informed consent, no regular vaginal bleeding during last 3 months, nor HRT during last 2 months, intact normal uterus, general good health, no planned major surgery during treatment period, diastolic blood pressure <= 90mmHg, BMI <= 30 kg/m2, willingness and ability to comply with visit schedule Exclusion Criteria: known or suspected estrogen dependent neoplasia, acute or chronic liver disease, deep venous thrombosis, thromboembolic disorders, cerebrovascular accidents or past history associated with estrogen use, pregnancy, haemoglobinopathies, porpyria, chronic medication, cardiac diseases not stable for >= 3 months, diabetes mellitus, abnormal genital bleeding of unknown aetiology, abnormal blood tests Criteria for confirming menopausal status: not reported Baseline Equality: not specified but statement that Rx and placebo groups "appeared identical according to pre randomisation data" Baseline symptoms: mean weekly hot flush frequency = 39.5±14.5
Interventions	Rx1 (E+P, cyclic): 2mg 17 beta oestradiol days 1 to 12; 2mg 17 beta oestradiol + 1mg norethisterone acetate days 13 to 22, 1mg 17 beta oestradiol days 23 to 28 (Trisequens) Rx2: placebo HRT and placebo preparations were identical in appearance. Co-interventions: none reported
Outcomes	1. Hot flush frequency. Side-effects data not suitable for meta-analysis.
Notes	The author was contacted and kindly supplied further information. The full paper is published in a Russian journal: J. Akusherstvo i Ginecologia 1997; 5: 76-80. Further data on the mean severity score for vasomotor symptoms yet to be obtained.

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment?	Low risk	A - Adequate

Study Design: parallel, double blinded, single centre, placebo controlled, randomised clinical trial
Randomisation: method not stated
Duration of Trial: 24 months
Power Calculations: none
Number of women randomised: 50
Number of women analysed: 39
Intention to Treat analysis: no
Losses to follow-up/withdrawals from treatment: 11/50 = 22%. Reasons not specified.
Compliance: not assessed
Source of Funding: Danish Medical Research Council, Schering A/S

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Agnusdei 1995	All groups (including placebo) given medroxyprogesterone acetate (MPA) for 15/30 days
Al-Azzawi 1999	Not double blinded; no placebo

Study	Reason for exclusion
Archer 1999	No oral HRT, no placebo
Aslaksen 1982	Oral MPA only, no oestrogen
Aylward 1974	Cross-over trial with no placebo group
Bakke 1965	Three phase cross-over study with each treatment period = two months
Barret-Connor 1999	No placebo
Blahey 1953	Cross-over study with treatment period two months
Blatt 1953	No specific hot flush outcome. Patient group may overlap with that of Kupperman 1953.
Cano 1999	No oral HRT, no placebo
CHART 1996	No hot flush outcomes (confirmed by author)
Cheng 1993	Not double-blind, no hot flush outcomes for placebo group
Cooper 1974	Treatment period = eight weeks
Crona 1988	Cross-over study with each treatment period = six weeks
Dennerstein 1980	Not a randomised controlled trial
Ditkoff 1991	No hot flush outcomes. Women with vasomotor symptoms at baseline excluded from the trial.
Fedor-Freybergh 1977	No specific hot flush outcomes
Fritsch 1997	Not clear if RCT, no placebo
Furuhjelm 1976	No vasomotor symptoms
Furuhjelm 1984	Treatment period = two months
Geola 1980	No hot flush outcomes, not double blinded
Greenblatt 1959	Medication changed at monthly intervals
Hailes 1981	Cross-over study with each treatment period = six weeks
Hammar 1998	No placebo
Hammar 1999	No placebo, not blinded
Hovic 1989	No placebo, not double-blind
Jarvinen 1971	Not randomised
Jones 1977	Not randomised control trial, no placebo
Khoo 1998	No hot flush outcomes
Kirkham 1991	No oral HRT, no hot flush outcomes

Study	Reason for exclusion
Klaiber 1996	Cross-over trial with each treatment phase two months, no hot flush outcomes
Kupperman 1953	No specific hot flush outcomes. Patient group may overlap with that of Blatt 1953. Not clear if participants randomised.
Lagrelius 1986	No placebo
Leonetti 1999	No oral HRT
Lind 1979	Not double-blind
Martin 1972	First phase- single-blind; second phase- no placebo and blinding unclear
Mattsson 1999 Study1	Transdermal oestrogen
Mattsson 1999 Study2	No placebo
Morrison 1980	Treatment = MPA (medroxyprogesterone acetate) - no oestrogen
Myers 1990	Only eight weeks HRT treatment
Nand 1998	No placebo
Natchigall 1979	No hot flush symptoms recorded
Nordin 1980	Only three weeks HRT treatment
Notelovitz 2000b	No oral HRT - transdermal patches only
Paterson 1982b	Oral progestogen only, no oestrogen. Apparently different study group to Paterson 1982a.
Pattison 1989	No placebo
Place 1985	No placebo
Polo-Kantola 1998	No oral HRT (patch or gel treatment)
Pratt 1937	No specific hot flush data; oestrogen content not clear and doses varied, duration of treatment not clear ("several months")
Rauramo 1975	No placebo, no HRT treatment
Rohr 1999	No oral HRT, no placebo, treatment phase less than three months
Saarikoski 1981	No hot flush outcomes
Schiff 1979	Cross-over trial with individual treatment phases of only 28 days
Schubert 1977	Treatment period of only 30 days
Shargil 1985	Not double blind, no placebo therapy. HRT dose outside usual therapeutic range for menopause symptoms (assessed contraceptive therapy)
Sheffrey 1969	Crossover trial - treatment period = two months

Study	Reason for exclusion
Sherwin 1989	No placebo
Spencer 1999	No placebo, not double-blind
Strickler 1977	Both HRT and placebo group medication supplemented with gestagen
Studd 1995	No placebo
Sulak 1999	No placebo
Taga 1999	No placebo, single blind
Thomson 1976	Cross-sectional study (no randomisation); treatment = 8 weeks
Utian 1971	Single-blind
Volpe 1986	Not double-blind

DATA AND ANALYSES

Comparison 1. Quality score of included studies

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Risk of bias assessment of included studies by assessment criteria in Table 1			Other data	No numeric data

Analysis 1.1. Comparison 1 Quality score of included studies, Outcome 1 Risk of bias assessment of included studies by assessment criteria in Table 1.

Risk of bias assessment of included studies by assessment criteria in Table 1

Study	Allocation	Treatment Blinding	Outcome Assessment	Baseline Equality	Losses to follow-up	Analysis Basis
Archer 1992	B	A	A	B	C	C
Baerug 1998	A	A	A	A	A	C
Baumgardner 1978	A	A	A	A	A	B
Bech 1998	B	A	A	C	C	C
Blumel 1994	A	A	A	A	A	C
Campbell 1976	B	B	A	B	C	C
Chung 1996	A	A	A	A	C	C
Conard 1995	A	B	A	A	C	C
Coope 1975	A	A	A	A	C	C
Coope 1981	A	A	A	A	C	C
Davidsen 1974	B	B	A	B	B	C
Dennerstein 1978	B	A	A	B	C	C
Derman 1995	A	A	A	A	C	A
Hagen 1982	B	B	A	A	C	C
Jensen J 1983	B	A	A	A	C	C

Risk of bias assessment of included studies by assessment criteria in Table 1

Study	Allocation	Treatment Blinding	Outcome Assessment	Baseline Equality	Losses to follow-up	Analysis Basis
Jensen P 1987	B	B	A	A	C	C
Marslew 1992	A	A	A	A	C	C
Martin 1971	B	A	A	A	C	C
Notelovitz 2000a	A	A	A	A	B	B
Paterson 1982a	A	A	A	A	C	C
PEPI 1998	A	A	A	C	A	A
Symons 2000 Study 1	A	A	A	A	C	A
Symons 2000 Study 2	A	A	A	A	A	A
Viklylaeva 1997	A	A	A	A	A	B

Comparison 2. Any HRT versus placebo: vasomotor outcomes at end of study

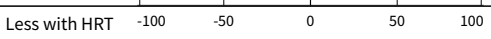
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Hot flush frequency/week	9	1104	Mean Difference (IV, Random, 95% CI)	-17.92 [-22.86, -12.99]
2 Hot flush frequency - log transformed	9	1104	Mean Difference (IV, Random, 95% CI)	-1.40 [-1.73, -1.07]
3 Hot Flush Frequency Weekly Weighted Score (HFWWS)	2	432	Mean Difference (IV, Random, 95% CI)	-48.14 [-64.22, -32.07]
4 HFWWS - log transformed	2	432	Mean Difference (IV, Random, 95% CI)	-1.55 [-2.92, -0.19]
5 Hot flush severity (all scales, continuous) - SMD	7	503	Std. Mean Difference (IV, Random, 95% CI)	-1.36 [-1.81, -0.90]
6 Hot flush severity (dichotomous)	9		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
6.1 Presence versus absence of hot flushes	8	1240	Odds Ratio (M-H, Random, 95% CI)	0.13 [0.06, 0.27]
6.2 Moderate-severe versus mild-absent hot flushes	4	337	Odds Ratio (M-H, Random, 95% CI)	0.26 [0.10, 0.66]
7 Hot flush severity (proportional odds ratios)	13	1724	Odds Ratio (Random, 95% CI)	0.13 [0.07, 0.23]
8 Hot flush outcomes not used in the meta-analysis			Other data	No numeric data

Analysis 2.1. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 1 Hot flush frequency/week.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	3.7 (6.9)	33	33.5 (31)		10.7%	-29.8[-40.49,-19.11]
Conard 1995	35	1 (2.9)	15	32.2 (35.6)		5.56%	-31.22[-49.24,-13.2]
Coope 1975	15	2.9 (6.4)	15	21.9 (26.8)		7.92%	-19.06[-33,-5.12]
Coope 1981	29	4.9 (14.4)	26	16.3 (26.9)		9.83%	-11.4[-22.99,0.19]
Derman 1995	34	0.9 (16.6)	36	12.6 (16.6)		14.01%	-11.79[-19.57,-4.01]
Notelovitz 2000a	225	13 (24)	55	28 (29)		13.39%	-15[-23.28,-6.72]
Symons 2000 Study 1	149	7.6 (13.9)	38	25 (29)		11.98%	-17.4[-26.89,-7.91]
Symons 2000 Study 2	199	14.1 (26.2)	67	39.4 (32.7)		12.97%	-25.3[-33.93,-16.67]
Viklylaeva 1997	32	13.8 (15.3)	28	23.7 (16.5)		13.63%	-9.84[-17.93,-1.75]
Total ***	**791**		**313**			**100%**	**-17.92[-22.86,-12.99]**

Heterogeneity: Tau²=29.47; Chi²=17.38, df=8(P=0.03); I²=53.97%

Test for overall effect: Z=7.12(P<0.0001)

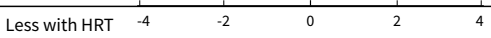


Analysis 2.2. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 2 Hot flush frequency - log transformed.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	1.5 (1.1)	33	3.9 (0.8)		12.14%	-2.42[-2.79,-2.05]
Conard 1995	35	0.1 (0.3)	15	1.4 (0.9)		11.2%	-1.29[-1.75,-0.83]
Coope 1975	15	0.6 (1.1)	15	2.1 (1.7)		5.88%	-1.53[-2.57,-0.49]
Coope 1981	29	0.8 (1.4)	26	2.2 (1.1)		9.05%	-1.43[-2.09,-0.77]
Derman 1995	34	-0.7 (1.3)	36	0.8 (0.7)		10.77%	-1.5[-2,-1]
Notelovitz 2000a	225	2 (1.2)	55	3 (0.8)		13.09%	-1.07[-1.34,-0.8]
Symons 2000 Study 1	149	1.5 (1.1)	38	2.9 (0.9)		12.43%	-1.34[-1.68,-1]
Symons 2000 Study 2	199	2 (1.2)	67	3.5 (0.7)		13.34%	-1.43[-1.67,-1.19]
Viklylaeva 1997	32	2.3 (0.9)	28	3 (0.6)		12.11%	-0.68[-1.05,-0.31]
Total ***	**791**		**313**			**100%**	**-1.4[-1.73,-1.07]**

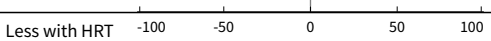
Heterogeneity: Tau²=0.2; Chi²=50.11, df=8(P<0.0001); I²=84.03%

Test for overall effect: Z=8.37(P<0.0001)

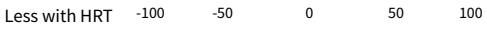


Analysis 2.3. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 3 Hot Flush Frequency Weekly Weighted Score (HFWWS).

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	6.2 (3.3)	35	62.6 (13.2)		49.66%	-56.4[-60.84,-51.96]
Notelovitz 2000a	260	36.9 (22)	64	76.9 (9.6)		50.34%	-40[-43.56,-36.44]
Total ***	**333**		**99**			**100%**	**-48.14[-64.22,-32.07]**



Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Heterogeneity: Tau²=130.27; Chi²=31.91, df=1(P<0.0001); I²=96.87%							
Test for overall effect: Z=5.87(P<0.0001)							

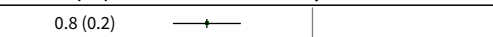


Analysis 2.4. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 4 HFWWS - log transformed.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	1.9 (0.4)	35	4.1 (0.2)		49.93%	-2.25[-2.37,-2.13]
Notelovitz 2000a	260	3.5 (0.5)	64	4.4 (0.1)		50.07%	-0.86[-0.93,-0.79]
Total ***	**333**		**99**			**100%**	**-1.55[-2.92,-0.19]**
Heterogeneity: Tau²=0.96; Chi²=367.42, df=1(P<0.0001); I²=99.73%							
Test for overall effect: Z=2.24(P=0.03)							

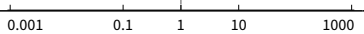
Analysis 2.5. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 5 Hot flush severity (all scales, continuous) - SMD.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Conard 1995	35	0.1 (0.3)	15	0.8 (0.2)		12.21%	-2.44[-3.22,-1.65]
Bech 1998	68	0.1 (0.4)	37	0.8 (0.9)		16.32%	-1.1[-1.53,-0.68]
Chung 1996	43	0.4 (0.7)	40	0.8 (0.9)		16.2%	-0.55[-0.99,-0.11]
Blumel 1994	25	0 (0.2)	23	0.6 (0.8)		14.36%	-0.92[-1.52,-0.32]
Paterson 1982a	11	0.3 (0.2)	9	0.6 (0.1)		9.1%	-1.88[-2.97,-0.78]
Baerug 1998	78	0.4 (0.7)	41	2 (1.1)		16.1%	-1.86[-2.31,-1.42]
Derman 1995	39	4.5 (3.7)	39	9.4 (4.5)		15.71%	-1.18[-1.66,-0.7]
Total ***	**299**		**204**			**100%**	**-1.36[-1.81,-0.9]**
Heterogeneity: Tau²=0.28; Chi²=28.72, df=6(P<0.0001); I²=79.11%							
Test for overall effect: Z=5.82(P<0.0001)							

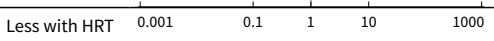


Analysis 2.6. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 6 Hot flush severity (dichotomous).

Study or subgroup	HRT n/N	Placebo n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
2.6.1 Presence versus absence of hot flushes					
Blumel 1994	1/25	10/23		6.48%	0.05[0.01,0.47]
Chung 1996	11/43	22/40		15.93%	0.28[0.11,0.71]
Conard 1995	4/35	12/15		9.39%	0.03[0.01,0.17]

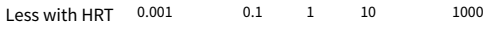


Study or subgroup	HRT n/N	Placebo n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
Coope 1975	3/15	10/15		9.24%	0.13[0.02,0.66]
Jensen J 1983	8/64	15/23		13.7%	0.08[0.02,0.24]
Jensen P 1987	6/27	23/30		12.64%	0.09[0.03,0.3]
Marslew 1992	5/17	17/22		10.86%	0.12[0.03,0.52]
PEPI 1998	101/680	49/166		21.76%	0.42[0.28,0.62]
Subtotal (95% CI)	**906**	**334**		**100%**	**0.13[0.06,0.27]**
Total events: 139 (HRT), 158 (Placebo)					
Heterogeneity: Tau²=0.68; Chi²=23.15, df=7(P=0); I²=69.77%					
Test for overall effect: Z=5.47(P<0.0001)					
2.6.2 Moderate-severe versus mild-absent hot flushes					
Baumgardner 1978	30/114	24/42		48.32%	0.27[0.13,0.56]
Blumel 1994	0/25	2/23		9.87%	0.17[0.01,3.71]
Chung 1996	4/43	6/40		31.23%	0.58[0.15,2.23]
Conard 1995	0/35	6/15		10.58%	0.02[0,0.4]
Subtotal (95% CI)	**217**	**120**		**100%**	**0.26[0.1,0.66]**
Total events: 34 (HRT), 38 (Placebo)					
Heterogeneity: Tau²=0.29; Chi²=4.31, df=3(P=0.23); I²=30.47%					
Test for overall effect: Z=2.83(P=0)					



Analysis 2.7. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 7 Hot flush severity (proportional odds ratios).

Study or subgroup	HRT N	Placebo N	log[Odds Ratio] (SE)	Odds Ratio IV, Random, 95% CI	Weight	Odds Ratio IV, Random, 95% CI
Baerug 1998	74	41	-3.6 (0.689)		6.83%	0.03[0.01,0.11]
Baumgardner 1978	114	42	-1.3 (0.378)		9.45%	0.27[0.13,0.56]
Bech 1998	68	37	-2.1 (0.48)		8.58%	0.13[0.05,0.33]
Blumel 1994	25	23	-1.8 (0.615)		7.43%	0.17[0.05,0.56]
Chung 1996	43	40	-1.2 (0.38)		9.43%	0.31[0.15,0.66]
Conard 1995	45	15	-3.1 (0.568)		7.82%	0.05[0.02,0.14]
Coope 1975	15	15	-2.1 (0.847)		5.68%	0.13[0.02,0.66]
Derman 1995	39	39	-2.2 (0.522)		8.22%	0.11[0.04,0.31]
Jensen J 1983	64	23	-2.4 (0.55)		7.98%	0.09[0.03,0.27]
Jensen P 1987	27	30	-2.4 (0.633)		7.28%	0.09[0.03,0.3]
Marslew 1992	17	22	-2.1 (0.736)		6.47%	0.12[0.03,0.52]
Paterson 1982a	11	9	-3.4 (1.267)		3.51%	0.03[0,0.38]
PEPI 1998	680	166	-0.9 (0.022)		11.31%	0.42[0.4,0.43]
Total (95% CI)					**100%**	**0.13[0.07,0.23]**
Heterogeneity: Tau²=0.72; Chi²=68.48, df=12(P<0.0001); I²=82.48%						
Test for overall effect: Z=7.17(P<0.0001)						



Analysis 2.8. Comparison 2 Any HRT versus placebo: vasomotor outcomes at end of study, Outcome 8 Hot flush outcomes not used in the meta-analysis.

Hot flush outcomes not used in the meta-analysis

Study	Outcome	Reason not used	Result
Archer 1992	Mean daily hot flush frequency	No standard deviations (SD) or range	Hot flushes significantly decreased in all three HRT groups compared to placebo
Baumgardner 1978	Percent reduction (from baseline) in daily hot flushes	Baseline reduction, no SD or range	Significantly greater reduction in daily hot flushes in 3 HRT groups compared to placebo group
Campbell 1976	Hot flush graphic rating scale	No data from end of first cross-over	Significant reduction in hot flush rating in HRT group compared to placebo
Davidsen 1974	Patient preference for HRT or placebo for vasomotor relief	Not all participants analysed, data combined for both cross-over phases	More patients preferred HRT for vasomotor symptom relief
Dennerstein 1978	Hot flush frequency & intensity	No data for first phase of cross-over	E, P and E+P more effective than placebo in controlling hot flush frequency & intensity
Hagen 1982	Hot flush severity score	Data combined with thiazide groups	Significant decrease in hot flush severity in HRT groups not observed in placebo group
Martin 1971	Percent reduction from baseline for hot flush frequency & intensity	No SD or range for end of study	Significantly greater reduction in HRT group compared to placebo for both hot flush frequency & intensity

Comparison 3. Any HRT versus placebo: vasomotor outcomes at 3 months

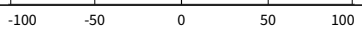
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Hot flush frequency/week	8	1036	Mean Difference (IV, Random, 95% CI)	-19.38 [-27.50, -11.25]
2 Hot flush frequency - log transformed	8	1036	Mean Difference (IV, Random, 95% CI)	-1.20 [-1.63, -0.77]
3 HFWWS	2	432	Mean Difference (IV, Random, 95% CI)	-48.14 [-64.22, -32.07]
4 HFWWS - log transformed	2	432	Mean Difference (IV, Random, 95% CI)	-1.55 [-2.92, -0.19]
5 Hot flush severity (0-3 scale, continuous) - WMD	4	237	Mean Difference (IV, Random, 95% CI)	-0.67 [-1.09, -0.26]
6 Hot flush severity (dichotomous)	6		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
6.1 Presence versus absence of hot flushes	5	272	Odds Ratio (M-H, Random, 95% CI)	0.11 [0.05, 0.24]
6.2 Moderate-severe versus mild-absent hot flushes	3	254	Odds Ratio (M-H, Random, 95% CI)	0.17 [0.05, 0.65]
7 Hot flush severity (proportional odds ratios)	8	554	odds ratios (Random, 95% CI)	0.09 [0.05, 0.17]

Analysis 3.1. Comparison 3 Any HRT versus placebo: vasomotor outcomes at 3 months, Outcome 1 Hot flush frequency/week.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	7.4 (13.7)	33	67 (62)		7.83%	-59.6[-80.99,-38.21]
Conard 1995	35	1 (2.9)	15	32.2 (35.6)		9.28%	-31.22[-49.24,-13.2]
Coope 1975	15	2.9 (6.4)	15	21.9 (26.8)		11.37%	-19.06[-33,-5.12]
Coope 1981	29	10.9 (12.6)	26	18 (25.1)		13.19%	-7.1[-17.78,3.58]
Notelovitz 2000a	225	13 (24)	55	28 (29)		14.51%	-15[-23.28,-6.72]
Symons 2000 Study 1	150	7.4 (12.4)	39	21 (25)		14.61%	-13.6[-21.69,-5.51]
Symons 2000 Study 2	199	14.1 (26.2)	67	39.4 (32.7)		14.32%	-25.3[-33.93,-16.67]
Viklylaeva 1997	32	15.5 (15.3)	28	21.8 (14.4)		14.9%	-6.22[-13.74,1.3]
Total * **	**758**		**278**			**100%**	**-19.38[-27.5,-11.25]**

Heterogeneity: Tau^2=100.67; Chi^2=32.89, df=7(P<0.0001); I^2=78.72%

Test for overall effect: Z=4.67(P<0.0001)



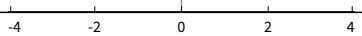
Less with HRT — Less with Placebo

Analysis 3.2. Comparison 3 Any HRT versus placebo: vasomotor outcomes at 3 months, Outcome 2 Hot flush frequency - log transformed.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	1.5 (1.1)	33	3.9 (0.8)		13.14%	-2.42[-2.79,-2.05]
Conard 1995	35	0.1 (0.3)	15	1.4 (0.9)		12.48%	-1.29[-1.75,-0.83]
Coope 1975	15	0.6 (1.1)	15	2.1 (1.7)		7.85%	-1.53[-2.57,-0.49]
Coope 1981	29	2.1 (0.9)	26	2.4 (1)		12.16%	-0.34[-0.84,0.16]
Notelovitz 2000a	225	2 (1.2)	55	3 (0.8)		13.77%	-1.07[-1.34,-0.8]
Symons 2000 Study 1	150	1.6 (1.1)	39	2.7 (0.9)		13.37%	-1.13[-1.46,-0.8]
Symons 2000 Study 2	199	2 (1.2)	67	3.5 (0.7)		13.94%	-1.43[-1.67,-1.19]
Viklylaeva 1997	32	2.5 (0.8)	28	3 (0.6)		13.28%	-0.46[-0.81,-0.11]
Total * **	**758**		**278**			**100%**	**-1.2[-1.63,-0.77]**

Heterogeneity: Tau^2=0.33; Chi^2=76.04, df=7(P<0.0001); I^2=90.79%

Test for overall effect: Z=5.49(P<0.0001)



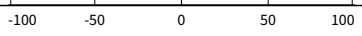
Less with HRT — Less with Placebo

Analysis 3.3. Comparison 3 Any HRT versus placebo: vasomotor outcomes at 3 months, Outcome 3 HFWWS.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	6.2 (3.3)	35	62.6 (13.2)		49.66%	-56.4[-60.84,-51.96]
Notelovitz 2000a	260	36.9 (22)	64	76.9 (9.6)		50.34%	-40[-43.56,-36.44]
Total * **	**333**		**99**			**100%**	**-48.14[-64.22,-32.07]**

Heterogeneity: Tau^2=130.27; Chi^2=31.91, df=1(P<0.0001); I^2=96.87%

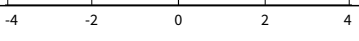
Test for overall effect: Z=5.87(P<0.0001)



Less with HRT — Less with Placebo

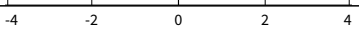
Analysis 3.4. Comparison 3 Any HRT versus placebo: vasomotor outcomes at 3 months, Outcome 4 HFWWS - log transformed.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	73	1.9 (0.4)	35	4.1 (0.2)		49.93%	-2.25[-2.37,-2.13]
Notelovitz 2000a	260	3.5 (0.5)	64	4.4 (0.1)		50.07%	-0.86[-0.93,-0.79]
Total ***	**333**		**99**			**100%**	**-1.55[-2.92,-0.19]**
Heterogeneity: Tau^2=0.96; Chi^2=367.42, df=1(P<0.0001); I^2=99.73%							
Test for overall effect: Z=2.24(P=0.03)							



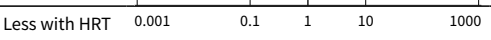
Analysis 3.5. Comparison 3 Any HRT versus placebo: vasomotor outcomes at 3 months, Outcome 5 Hot flush severity (0-3 scale, continuous) - WMD.

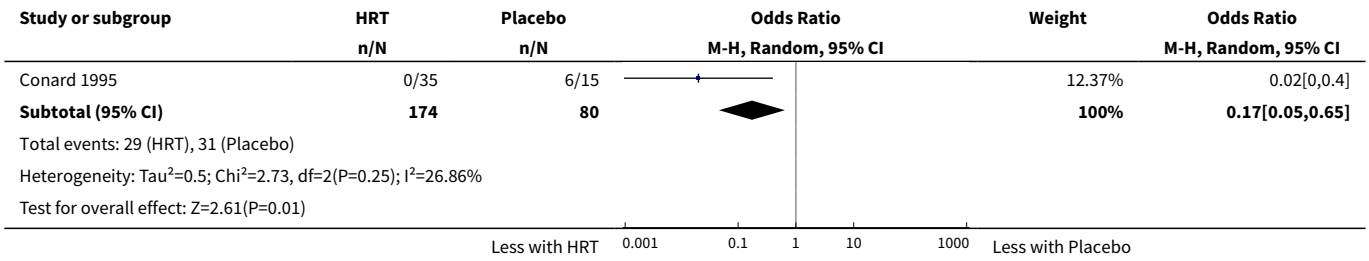
Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	78	0.4 (0.7)	41	2 (1)		22.98%	-1.57[-1.91,-1.23]
Blumel 1994	25	0.2 (0.4)	23	0.4 (0.6)		24.25%	-0.19[-0.46,0.08]
Conard 1995	35	0.1 (0.3)	15	0.8 (0.2)		26.36%	-0.7[-0.83,-0.57]
Paterson 1982a	11	0.3 (0.2)	9	0.6 (0.1)		26.4%	-0.31[-0.44,-0.18]
Total ***	**149**		**88**			**100%**	**-0.67[-1.09,-0.26]**
Heterogeneity: Tau^2=0.17; Chi^2=59.33, df=3(P<0.0001); I^2=94.94%							
Test for overall effect: Z=3.16(P=0)							



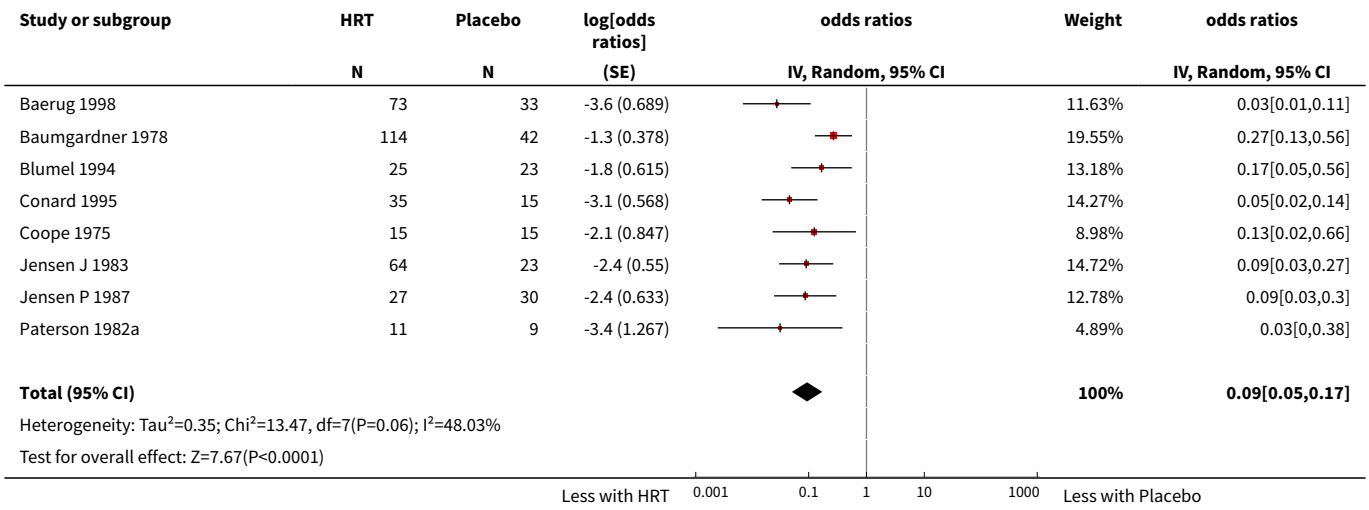
Analysis 3.6. Comparison 3 Any HRT versus placebo: vasomotor outcomes at 3 months, Outcome 6 Hot flush severity (dichotomous).

Study or subgroup	HRT n/N	Placebo n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
3.6.1 Presence versus absence of hot flushes					
Blumel 1994	4/25	7/23		19.82%	0.44[0.11,1.75]
Conard 1995	4/35	12/15		15.79%	0.03[0.01,0.17]
Coope 1975	3/15	10/15		15.51%	0.13[0.02,0.66]
Jensen J 1983	11/64	16/23		26.22%	0.09[0.03,0.27]
Jensen P 1987	9/27	25/30		22.66%	0.1[0.03,0.35]
Subtotal (95% CI)	**166**	**106**		**100%**	**0.11[0.05,0.24]**
Total events: 31 (HRT), 70 (Placebo)					
Heterogeneity: Tau^2=0.25; Chi^2=6.08, df=4(P=0.19); I^2=34.17%					
Test for overall effect: Z=5.63(P<0.0001)					
3.6.2 Moderate-severe versus mild-absent hot flushes					
Baumgardner 1978	29/114	24/42		77.16%	0.26[0.12,0.54]
Blumel 1994	0/25	1/23		10.47%	0.29[0.01,7.59]





Analysis 3.7. Comparison 3 Any HRT versus placebo: vasomotor outcomes at 3 months, Outcome 7 Hot flush severity (proportional odds ratios).



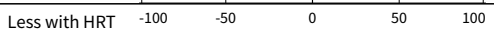
Comparison 4. Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Hot flush frequency/week	9		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 Unopposed oestrogen (E only) only versus placebo	3	365	Mean Difference (IV, Random, 95% CI)	-14.78 [-20.85, -8.72]
1.2 Oestrogen + Progestogen (E+P) versus Placebo	6	739	Mean Difference (IV, Random, 95% CI)	-19.65 [-26.74, -12.56]
2 Hot flush frequency (log transformed)	9		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 E only versus placebo	3	365	Mean Difference (IV, Random, 95% CI)	-1.14 [-1.38, -0.90]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.2 E+P versus Placebo	6	739	Mean Difference (IV, Random, 95% CI)	-1.44 [-1.88, -1.01]
3 HFWWS	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.1 E only versus placebo	1	324	Mean Difference (IV, Random, 95% CI)	-40.00 [-43.56, -36.44]
3.2 E+P versus Placebo	1	108	Mean Difference (IV, Random, 95% CI)	-56.4 [-57.26, -55.54]
4 HFWWS - log transformed	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.1 E only versus placebo	1	324	Mean Difference (IV, Random, 95% CI)	-0.86 [-0.93, -0.79]
4.2 E+P versus Placebo	1	108	Mean Difference (IV, Random, 95% CI)	-2.25 [-2.37, -2.13]
5 Hot flush severity score (all scales, continuous) - SMD	7		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
5.1 E only versus placebo	1	83	Std. Mean Difference (IV, Random, 95% CI)	-0.55 [-0.99, -0.11]
5.2 E+P versus Placebo	6	420	Std. Mean Difference (IV, Random, 95% CI)	-1.50 [-1.93, -1.07]
6 Hot flush severity (dichotomous)	9		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
6.1 Presence versus absence of hot flushes - E only versus Placebo	3	959	Odds Ratio (M-H, Random, 95% CI)	0.35 [0.22, 0.55]
6.2 Presence versus absence of hot flushes - E+P versus Placebo	6	957	Odds Ratio (M-H, Random, 95% CI)	0.10 [0.04, 0.27]
6.3 Moderate-severe versus mild-absent hot flushes - E only versus placebo	2	239	Odds Ratio (M-H, Random, 95% CI)	0.32 [0.17, 0.61]
6.4 Moderate-severe versus mild-absent hot flushes - E+ P versus placebo	2	98	Odds Ratio (M-H, Random, 95% CI)	0.06 [0.01, 0.48]
7 Hot flush severity (proportional odds ratios)	13		Odds Ratios (Random, 95% CI)	Subtotals only
7.1 E only vs Placebo	4	605	Odds Ratios (Random, 95% CI)	0.35 [0.22, 0.56]
7.2 E+P vs Placebo	10	1268	Odds Ratios (Random, 95% CI)	0.10 [0.06, 0.19]

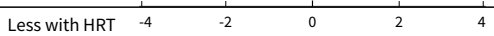
Analysis 4.1. Comparison 4 Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes, Outcome 1 Hot flush frequency/week.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
4.1.1 Unopposed oestrogen (E only) only versus placebo							
Coope 1975	15	2.9 (6.4)	15	21.9 (26.8)		25.45%	-19.06[-33,-5.12]
Coope 1981	29	4.9 (14.4)	26	16.3 (26.9)		31.58%	-11.4[-22.99,0.19]
Notelovitz 2000a	225	13 (24)	55	28 (29)		42.97%	-15[-23.28,-6.72]
Subtotal ***	**269**		**96**			**100%**	**-14.78[-20.85,-8.72]**
Heterogeneity: Tau²=0; Chi²=0.69, df=2(P=0.71); I²=0%							
Test for overall effect: Z=4.78(P<0.0001)							
4.1.2 Oestrogen + Progestogen (E+P) versus Placebo							
Baerug 1998	73	3.7 (6.9)	33	33.5 (31)		15.55%	-29.8[-40.49,-19.11]
Conard 1995	35	1 (2.9)	15	32.3 (35.6)		8.09%	-31.32[-49.34,-13.3]
Derman 1995	34	0.9 (16.6)	36	12.6 (16.6)		20.34%	-11.75[-19.53,-3.97]
Symons 2000 Study 1	149	7.6 (13.9)	38	25 (29)		17.4%	-17.4[-26.89,-7.91]
Symons 2000 Study 2	199	14.1 (26.2)	67	39.4 (32.7)		18.83%	-25.3[-33.93,-16.67]
Viklylaeva 1997	32	13.8 (15.3)	28	23.7 (16.5)		19.78%	-9.84[-17.93,-1.75]
Subtotal ***	**522**		**217**			**100%**	**-19.65[-26.74,-12.56]**
Heterogeneity: Tau²=51.69; Chi²=15.92, df=5(P=0.01); I²=68.6%							
Test for overall effect: Z=5.43(P<0.0001)							



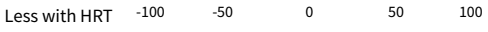
Analysis 4.2. Comparison 4 Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes, Outcome 2 Hot flush frequency (log transformed).

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
4.2.1 E only versus placebo							
Coope 1975	15	0.6 (1.1)	15	2.1 (1.7)		21%	-1.53[-2.57,-0.49]
Coope 1981	29	0.8 (1.4)	26	2.2 (1.1)		32.29%	-1.43[-2.09,-0.77]
Notelovitz 2000a	225	2 (1.2)	55	3 (0.8)		46.71%	-1.07[-1.34,-0.8]
Subtotal ***	**269**		**96**			**100%**	**-1.14[-1.38,-0.9]**
Heterogeneity: Tau²=0; Chi²=1.54, df=2(P=0.46); I²=0%							
Test for overall effect: Z=9.29(P<0.0001)							
4.2.2 E+P versus Placebo							
Baerug 1998	73	1.5 (1.1)	33	3.9 (0.8)		16.86%	-2.42[-2.79,-2.05]
Conard 1995	35	0.1 (0.3)	15	1.4 (0.9)		15.56%	-1.29[-1.75,-0.83]
Derman 1995	34	-0.7 (1.3)	36	0.8 (0.7)		14.96%	-1.5[-2,-1]
Symons 2000 Study 1	149	1.5 (1.1)	38	2.9 (0.9)		17.27%	-1.34[-1.68,-1]
Symons 2000 Study 2	199	2 (1.2)	67	3.5 (0.7)		18.53%	-1.43[-1.67,-1.19]
Viklylaeva 1997	32	2.3 (0.9)	28	3 (0.6)		16.82%	-0.68[-1.05,-0.31]
Subtotal ***	**522**		**217**			**100%**	**-1.44[-1.88,-1.01]**
Heterogeneity: Tau²=0.26; Chi²=44.11, df=5(P<0.0001); I²=88.66%							
Test for overall effect: Z=6.48(P<0.0001)							



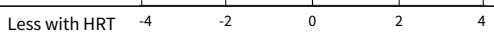
Analysis 4.3. Comparison 4 Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes, Outcome 3 HFWWS.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
4.3.1 E only versus placebo							
Notelovitz 2000a	260	36.9 (22)	64	76.9 (9.6)		100%	-40[-43.56,-36.44]
Subtotal ***	**260**		**64**			**100%**	**-40[-43.56,-36.44]**
Heterogeneity: Not applicable							
Test for overall effect: Z=22.01(P<0.0001)							
4.3.2 E+P versus Placebo							
Baerug 1998	73	6.2 (3.3)	35	62.6 (1.2)		100%	-56.4[-57.26,-55.54]
Subtotal ***	**73**		**35**			**100%**	**-56.4[-57.26,-55.54]**
Heterogeneity: Not applicable							
Test for overall effect: Z=128.58(P<0.0001)							



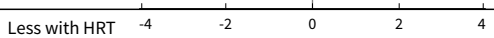
Analysis 4.4. Comparison 4 Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes, Outcome 4 HFWWS - log transformed.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
4.4.1 E only versus placebo							
Notelovitz 2000a	260	3.5 (0.5)	64	4.4 (0.1)		100%	-0.86[-0.93,-0.79]
Subtotal ***	**260**		**64**			**100%**	**-0.86[-0.93,-0.79]**
Heterogeneity: Not applicable							
Test for overall effect: Z=23.44(P<0.0001)							
4.4.2 E+P versus Placebo							
Baerug 1998	73	1.9 (0.4)	35	4.1 (0.2)		100%	-2.25[-2.37,-2.13]
Subtotal ***	**73**		**35**			**100%**	**-2.25[-2.37,-2.13]**
Heterogeneity: Not applicable							
Test for overall effect: Z=35.97(P<0.0001)							

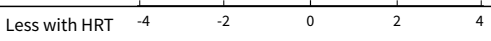


Analysis 4.5. Comparison 4 Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes, Outcome 5 Hot flush severity score (all scales, continuous) - SMD.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
4.5.1 E only versus placebo							
Chung 1996	43	0.4 (0.7)	40	0.8 (0.9)		100%	-0.55[-0.99,-0.11]
Subtotal ***	**43**		**40**			**100%**	**-0.55[-0.99,-0.11]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.45(P=0.01)							
4.5.2 E+P versus Placebo							

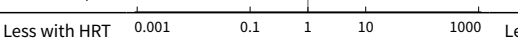


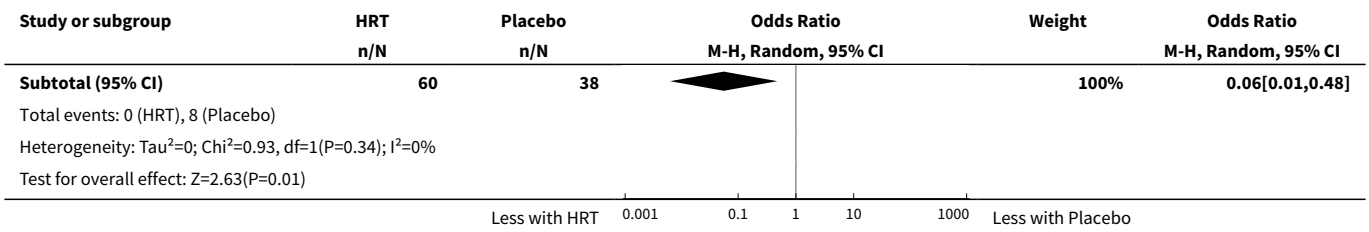
Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Baerug 1998	78	0.4 (0.7)	41	2 (1.1)		19.21%	-1.86[-2.31,-1.42]
Bech 1998	68	0.1 (0.4)	37	0.8 (0.9)		19.47%	-1.1[-1.53,-0.68]
Blumel 1994	25	0 (0.2)	23	0.6 (0.8)		17.14%	-0.92[-1.52,-0.32]
Conard 1995	35	0.1 (0.3)	15	0.8 (0.2)		14.57%	-2.44[-3.22,-1.65]
Derman 1995	39	4.5 (3.7)	39	9.4 (4.5)		18.75%	-1.18[-1.66,-0.7]
Paterson 1982a	11	0.3 (0.2)	9	0.6 (0.1)		10.86%	-1.88[-2.97,-0.78]
Subtotal * **	**256**		**164**			**100%**	**-1.5[-1.93,-1.07]**
Heterogeneity: Tau²=0.19; Chi²=16.61, df=5(P=0.01); I²=69.89%							
Test for overall effect: Z=6.79(P<0.0001)							



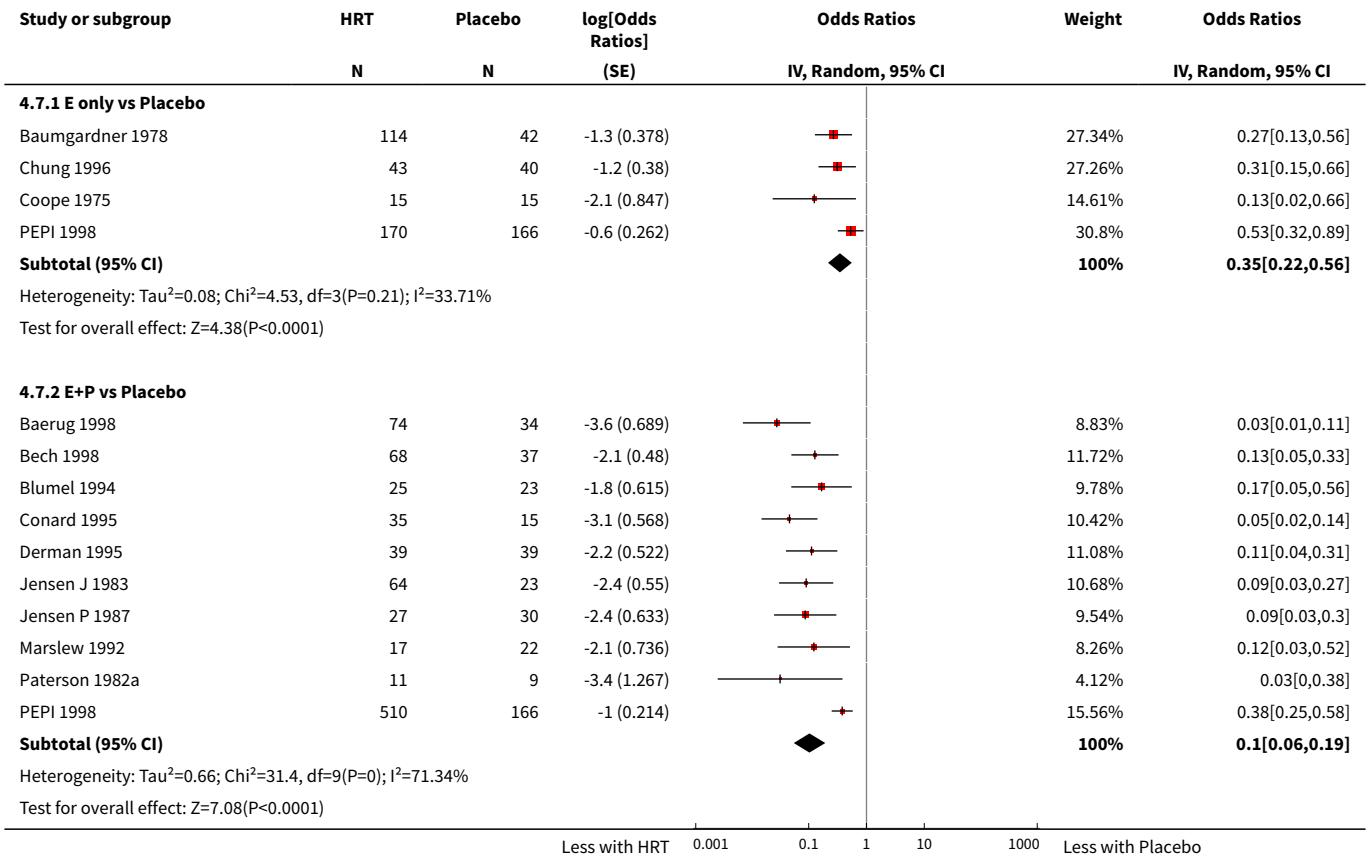
Analysis 4.6. Comparison 4 Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes, Outcome 6 Hot flush severity (dichotomous).

Study or subgroup	HRT n/N	Placebo n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
4.6.1 Presence versus absence of hot flushes - E only versus Placebo					
Chung 1996	11/43	22/40		33.01%	0.28[0.11,0.71]
Coope 1975	3/15	10/15		17.31%	0.13[0.02,0.66]
PEPI 1998	101/680	49/166		49.67%	0.42[0.28,0.62]
Subtotal (95% CI)	**738**	**221**		**100%**	**0.35[0.22,0.55]**
Total events: 115 (HRT), 81 (Placebo)					
Heterogeneity: Tau²=0.03; Chi²=2.33, df=2(P=0.31); I²=14.26%					
Test for overall effect: Z=4.51(P<0.0001)					
4.6.2 Presence versus absence of hot flushes - E+P versus Placebo					
Blumel 1994	1/25	10/23		7.71%	0.05[0.01,0.47]
Conard 1995	4/35	12/15		11.63%	0.03[0.01,0.17]
Jensen J 1983	8/64	15/23		18.1%	0.08[0.02,0.24]
Jensen P 1987	6/27	23/30		16.44%	0.09[0.03,0.3]
Marslew 1992	5/17	17/22		13.75%	0.12[0.03,0.52]
PEPI 1998	70/510	49/166		32.37%	0.38[0.25,0.58]
Subtotal (95% CI)	**678**	**279**		**100%**	**0.1[0.04,0.27]**
Total events: 94 (HRT), 126 (Placebo)					
Heterogeneity: Tau²=0.94; Chi²=19.62, df=5(P=0); I²=74.51%					
Test for overall effect: Z=4.68(P<0.0001)					
4.6.3 Moderate-severe versus mild-absent hot flushes - E only versus placebo					
Baumgardner 1978	30/114	24/42		63.1%	0.27[0.13,0.56]
Chung 1996	4/43	6/40		36.9%	0.58[0.15,2.23]
Subtotal (95% CI)	**157**	**82**		**100%**	**0.32[0.17,0.61]**
Total events: 34 (HRT), 30 (Placebo)					
Heterogeneity: Tau²=0; Chi²=0.98, df=1(P=0.32); I²=0%					
Test for overall effect: Z=3.44(P=0)					
4.6.4 Moderate-severe versus mild-absent hot flushes - E+ P versus placebo					
Blumel 1994	0/25	2/23		48.17%	0.17[0.01,3.71]
Conard 1995	0/35	6/15		51.83%	0.02[0,0.4]





Analysis 4.7. Comparison 4 Unopposed oestrogen versus placebo and oestrogen + progestogen versus placebo for vasomotor outcomes, Outcome 7 Hot flush severity (proportional odds ratios).



Comparison 5. Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes)

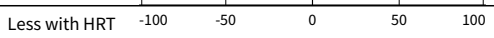
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Hot flush frequency/week	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
1.1 Perimenopausal participants	1	60	Mean Difference (IV, Random, 95% CI)	-9.84 [-17.93, -1.75]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.2 Postmenopausal participants	1	187	Mean Difference (IV, Random, 95% CI)	-17.4 [-26.89, -7.91]
2 Hot flush frequency (log transformed)	2		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 Perimenopausal participants	1	60	Mean Difference (IV, Random, 95% CI)	-0.68 [-1.05, -0.31]
2.2 Postmenopausal participants	1	187	Mean Difference (IV, Random, 95% CI)	-1.34 [-1.68, -1.00]
3 HFWWS	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
3.1 Perimenopausal participants	1	45	Mean Difference (IV, Random, 95% CI)	-35.5 [-41.81, -29.19]
3.2 Postmenopausal participants	1	63	Mean Difference (IV, Random, 95% CI)	-70.1 [-76.21, -63.99]
4 HFWWS - log transformed	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
4.1 Perimenopausal participants	1	45	Mean Difference (IV, Random, 95% CI)	-2.47 [-2.68, -2.26]
4.2 Postmenopausal participants	1	63	Mean Difference (IV, Random, 95% CI)	-2.17 [-2.31, -2.03]
5 Hot flush severity score (all scales, continuous) - WMD	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
5.1 Perimenopausal participants	0	0	Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
5.2 Postmenopausal participants	1	83	Mean Difference (IV, Random, 95% CI)	-0.43 [-0.77, -0.09]
6 Hot flush severity (dichotomous)	3		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
6.1 Presence versus absence - perimenopausal participants	0	0	Odds Ratio (M-H, Random, 95% CI)	0.0 [0.0, 0.0]
6.2 Presence versus absence - postmenopausal participants	3	1016	Odds Ratio (M-H, Random, 95% CI)	0.23 [0.09, 0.58]
6.3 Moderate-severe versus mild-absent - perimenopausal participants	0	0	Odds Ratio (M-H, Random, 95% CI)	0.0 [0.0, 0.0]
6.4 Moderate-severe versus mild-absent - postmenopausal participants	1	83	Odds Ratio (M-H, Random, 95% CI)	0.58 [0.15, 2.23]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7 Hot flush severity (proportional odds ratios)	3		Odds Ratios (Random, 95% CI)	Subtotals only
7.1 Perimenopausal participants	0	0	Odds Ratios (Random, 95% CI)	0.0 [0.0, 0.0]
7.2 Postmenopausal participants	3	1016	Odds Ratios (Random, 95% CI)	0.26 [0.12, 0.56]

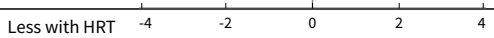
Analysis 5.1. Comparison 5 Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes), Outcome 1 Hot flush frequency/week.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
5.1.1 Perimenopausal participants							
Viklylaeva 1997	32	13.8 (15.3)	28	23.7 (16.5)		100%	-9.84[-17.93,-1.75]
Subtotal ***	**32**		**28**			**100%**	**-9.84[-17.93,-1.75]**
Heterogeneity: Not applicable							
Test for overall effect: Z=2.38(P=0.02)							
5.1.2 Postmenopausal participants							
Symons 2000 Study 1	149	7.6 (13.9)	38	25 (29)		100%	-17.4[-26.89,-7.91]
Subtotal ***	**149**		**38**			**100%**	**-17.4[-26.89,-7.91]**
Heterogeneity: Not applicable							
Test for overall effect: Z=3.59(P=0)							



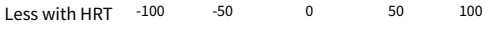
Analysis 5.2. Comparison 5 Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes), Outcome 2 Hot flush frequency (log transformed).

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
5.2.1 Perimenopausal participants							
Viklylaeva 1997	32	2.3 (0.9)	28	3 (0.6)		100%	-0.68[-1.05,-0.31]
Subtotal ***	**32**		**28**			**100%**	**-0.68[-1.05,-0.31]**
Heterogeneity: Not applicable							
Test for overall effect: Z=3.59(P=0)							
5.2.2 Postmenopausal participants							
Symons 2000 Study 1	149	1.5 (1.1)	38	2.9 (0.9)		100%	-1.34[-1.68,-1]
Subtotal ***	**149**		**38**			**100%**	**-1.34[-1.68,-1]**
Heterogeneity: Not applicable							
Test for overall effect: Z=7.75(P<0.0001)							



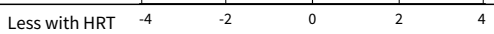
Analysis 5.3. Comparison 5 Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes), Outcome 3 HFWWS.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
5.3.1 Perimenopausal participants							
Baerug 1998	31	2.5 (1.5)	14	38 (12)		100%	-35.5[-41.81,-29.19]
Subtotal ***	**31**		**14**			**100%**	**-35.5[-41.81,-29.19]**
Heterogeneity: Not applicable							
Test for overall effect: Z=11.03(P<0.0001)							
5.3.2 Postmenopausal participants							
Baerug 1998	42	8.9 (4.1)	21	79 (14)		100%	-70.1[-76.21,-63.99]
Subtotal ***	**42**		**21**			**100%**	**-70.1[-76.21,-63.99]**
Heterogeneity: Not applicable							
Test for overall effect: Z=22.47(P<0.0001)							



Analysis 5.4. Comparison 5 Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes), Outcome 4 HFWWS - log transformed.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
5.4.1 Perimenopausal participants							
Baerug 1998	31	1.2 (0.4)	14	3.6 (0.3)		100%	-2.47[-2.68,-2.26]
Subtotal ***	**31**		**14**			**100%**	**-2.47[-2.68,-2.26]**
Heterogeneity: Not applicable							
Test for overall effect: Z=22.94(P<0.0001)							
5.4.2 Postmenopausal participants							
Baerug 1998	42	2.2 (0.4)	21	4.4 (0.2)		100%	-2.17[-2.31,-2.03]
Subtotal ***	**42**		**21**			**100%**	**-2.17[-2.31,-2.03]**
Heterogeneity: Not applicable							
Test for overall effect: Z=31.27(P<0.0001)							



Analysis 5.5. Comparison 5 Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes), Outcome 5 Hot flush severity score (all scales, continuous) - WMD.

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
5.5.1 Perimenopausal participants							
Subtotal ***	**0**		**0**				**Not estimable**
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							
5.5.2 Postmenopausal participants							
Chung 1996	43	0.4 (0.7)	40	0.8 (0.9)		100%	-0.43[-0.77,-0.09]
Subtotal ***	**43**		**40**			**100%**	**-0.43[-0.77,-0.09]**

Study or subgroup	HRT N	HRT Mean(SD)	Placebo N	Placebo Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Heterogeneity: Not applicable							
Test for overall effect: Z=2.5(P=0.01)							

Less with HRT -4 -2 0 2 4 Less with Placebo

Analysis 5.6. Comparison 5 Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes), Outcome 6 Hot flush severity (dichotomous).

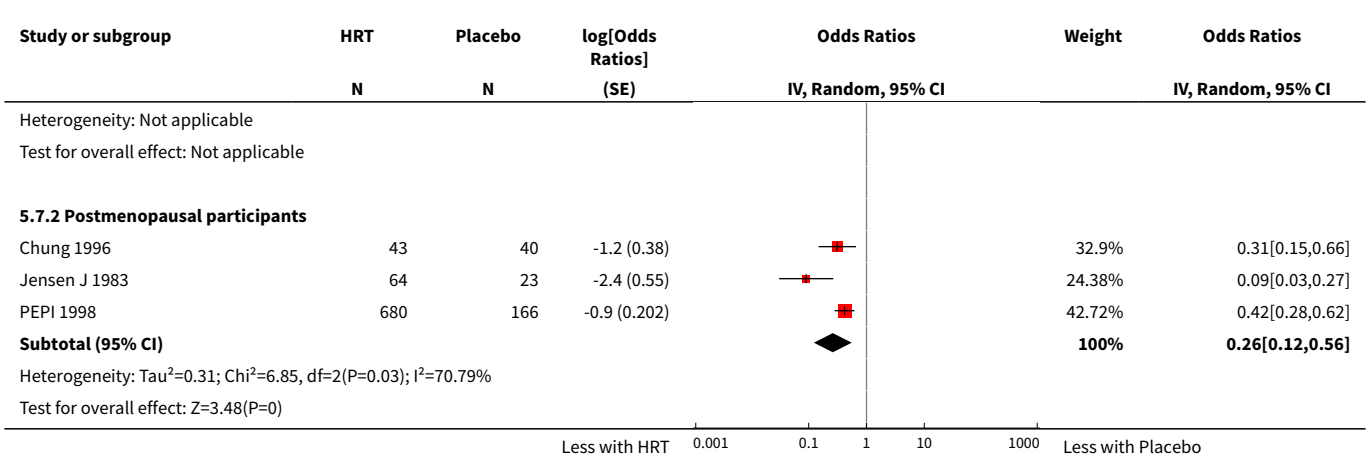
Study or subgroup	HRT n/N	Placebo n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
5.6.1 Presence versus absence - perimenopausal participants					
Subtotal (95% CI)	**0**	**0**			**Not estimable**
Total events: 0 (HRT), 0 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Not applicable					
5.6.2 Presence versus absence - postmenopausal participants					
Chung 1996	11/43	22/40		30.28%	0.28[0.11,0.71]
Jensen J 1983	8/64	15/23		25.36%	0.08[0.02,0.24]
PEPI 1998	101/680	49/166		44.36%	0.42[0.28,0.62]
Subtotal (95% CI)	**787**	**229**		**100%**	**0.23[0.09,0.58]**
Total events: 120 (HRT), 86 (Placebo)					
Heterogeneity: Tau^2=0.48; Chi^2=7.87, df=2(P=0.02); I^2=74.59%					
Test for overall effect: Z=3.11(P=0)					
5.6.3 Moderate-severe versus mild-absent - perimenopausal participants					
Subtotal (95% CI)	**0**	**0**			**Not estimable**
Total events: 0 (HRT), 0 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Not applicable					
5.6.4 Moderate-severe versus mild-absent - postmenopausal participants					
Chung 1996	4/43	6/40		100%	0.58[0.15,2.23]
Subtotal (95% CI)	**43**	**40**		**100%**	**0.58[0.15,2.23]**
Total events: 4 (HRT), 6 (Placebo)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.79(P=0.43)					

Less with HRT 0.001 0.1 1 10 1000 Less with Placebo

Analysis 5.7. Comparison 5 Peri-menopausal and post-menopausal women: any HRT versus placebo (vasomotor outcomes), Outcome 7 Hot flush severity (proportional odds ratios).

Study or subgroup	HRT N	Placebo N	log[Odds Ratios] (SE)	Odds Ratios IV, Random, 95% CI	Weight	Odds Ratios IV, Random, 95% CI
5.7.1 Perimenopausal participants						
Subtotal (95% CI)						**Not estimable**

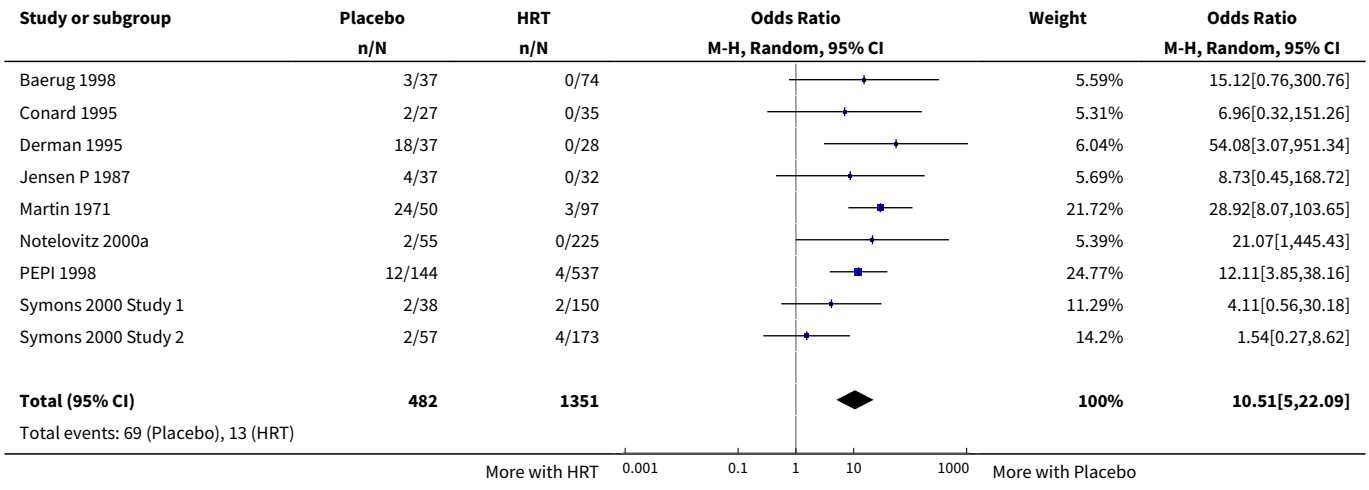
Less with HRT 0.001 0.1 1 10 1000 Less with Placebo



Comparison 6. Other outcomes: any HRT vs placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Withdrawals due to lack of effect	9	1833	Odds Ratio (M-H, Random, 95% CI)	10.51 [5.00, 22.09]
2 Withdrawals due to adverse events (any)	9	1882	Odds Ratio (M-H, Random, 95% CI)	1.25 [0.83, 1.90]
3 Adverse Events (any)	3	818	Odds Ratio (M-H, Random, 95% CI)	1.41 [1.00, 1.99]
4 Side-efects			Other data	No numeric data
5 Quality of life	1	105	Mean Difference (IV, Random, 95% CI)	-1.20 [-2.76, 0.36]

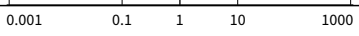
Analysis 6.1. Comparison 6 Other outcomes: any HRT vs placebo, Outcome 1 Withdrawals due to lack of effect.



Study or subgroup	Placebo n/N	HRT n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
Heterogeneity: Tau²=0.24; Chi²=9.86, df=8(P=0.28); I²=18.86%					
Test for overall effect: Z=6.21(P<0.0001)					

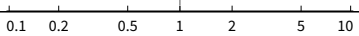
Analysis 6.2. Comparison 6 Other outcomes: any HRT vs placebo, Outcome 2 Withdrawals due to adverse events (any).

Study or subgroup	HRT n/N	Placebo n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
Baerug 1998	3/77	2/36		5.02%	0.69[0.11,4.32]
Conard 1995	2/37	2/17		4.03%	0.43[0.06,3.33]
Derman 1995	6/34	0/19		1.99%	8.89[0.47,167.17]
Jensen J 1983	7/71	0/23		2.03%	5.47[0.3,99.46]
Jensen P 1987	2/34	0/33		1.81%	5.15[0.24,111.52]
Notelovitz 2000a	21/225	5/55		15.39%	1.03[0.37,2.86]
PEPI 1998	138/671	21/132		52.66%	1.37[0.83,2.26]
Symons 2000 Study 1	16/150	1/38		4.03%	4.42[0.57,34.41]
Symons 2000 Study 2	10/173	5/57		13.03%	0.64[0.21,1.95]
Total (95% CI)	**1472**	**410**		**100%**	**1.25[0.83,1.9]**
Total events: 205 (HRT), 36 (Placebo)					
Heterogeneity: Tau²=0.02; Chi²=8.32, df=8(P=0.4); I²=3.87%					
Test for overall effect: Z=1.07(P=0.29)					



Analysis 6.3. Comparison 6 Other outcomes: any HRT vs placebo, Outcome 3 Adverse Events (any).

Study or subgroup	HRT n/N	Placebo n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
Notelovitz 2000a	182/267	44/66		36.28%	1.07[0.6,1.9]
Symons 2000 Study 1	120/176	25/43		25.46%	1.54[0.78,3.06]
Symons 2000 Study 2	116/199	30/67		38.26%	1.72[0.99,3.01]
Total (95% CI)	**642**	**176**		**100%**	**1.41[1,1.99]**
Total events: 418 (HRT), 99 (Placebo)					
Heterogeneity: Tau²=0; Chi²=1.45, df=2(P=0.48); I²=0%					
Test for overall effect: Z=1.95(P=0.05)					



Analysis 6.4. Comparison 6 Other outcomes: any HRT vs placebo, Outcome 4 Side-efects.

Study	Side-efects									
	Atypical bleeding	Nau-sea/Vom-iting	Breast tenderness	Headaches	Weight changes	Dizziness	Throm-bosis	Rash	Pruritis	Other
Archer 1992	-	Few women	Higher dos-es of oe-	Most women	-	-	-	Three women re-	No marked changes	Non-sig-nificant

Study	Atypical bleeding	Side-efects: Nausea/Vomiting	Breast tenderness	Headaches	Weight changes	Dizziness	Thrombosis	Rash	Pruritis	Other
		with symptoms at baseline, no consistent trends during the study	strogen were associated with higher rates of breast tenderness and pain. However, of women who completed the study, 78% given 2 mg E2 and 70% given 1.25mg CEE reported no discomfort	reported symptoms at baseline which decreased modestly in all groups (including placebo) during course of the study				ported adverse skin events: 2/25 (8%) placebo and 1/102 (1%) HRT	in vaginal itching of dryness	trend towards higher rate of any adverse events with higher oestrogen (20-22% in placebo, 1 mg E2 and 0.625mg CEE groups compared to 31-35% in 2 mg E2 and 1.25 mg CEE groups). Predominant adverse events were classified as digestive, nervous/psychiatric and urogenital.
Baerug 1998	Higher incidence of bleeding in E+P women compared to placebo but there was no difference between HRT and placebo group in the incidence of severe bleeding.	One woman (placebo group) withdrew due to nausea	One women (from HRT group) withdrew due to breast tenderness	One women (from placebo group) withdrew due to headaches	-	-	-	-	-	Additional withdrawals due to oedema and emotional lability
Baumgardner 1978	No significant differences between HRT and placebo groups	One woman given CEE withdrew from therapy due to nausea	Low prevalence, no significant differences between HRT and placebo	Low prevalence, no significant differences between HRT and placebo	No significant weight gain	Low prevalence, no significant differences between HRT and placebo	No events	No skin eruptions	1-2 women in each group complained of symptoms at baseline - all improved (including placebo group) as the study progressed	Oedema and visual symptoms were recorded but were of low prevalence with no difference between HRT and placebo groups. Reasons for withdrawals from therapy not itemised but there was a stated difficulty in getting women in the placebo group to continue due to lack of efficacy.

Study	Side-efects									
	Atypical bleeding	Nausea/Vomiting	Breast tenderness	Headaches	Weight changes	Dizziness	Thrombosis	Rash	Pruritis	Other
Bech 1998	Four women withdrew from trial in HRT groups due to bleeding irregularities compared to none in placebo group.	Two women in sequential therapy group, 2 in placebo group and none in the combined group suffered from nausea in the first 6 cycles.	Mastodynia was experienced significantly more often in the HRT groups compared to the placebo groups in the first 6 months. Over the 2 years study (reported in the secondary reference), 4 women on HRT withdrew due to mastodynia compared to none in the placebo group.	-	One women in the sequential therapy group withdrew due to weight gain	-	-	-	-	Two women in the sequential group, two in the combined group and four in the placebo group suffered from oedema in the first 6 cycles. Over the two year study (reported in the secondary reference), two women in the HRT group withdrew due to breast cancer.
Blumel 1994	Significantly more women in HRT group had irregular bleeding (12/25) compared to placebo group (3/23) at 6 months: p = 0.009. Results were similar at three months	-	At 3 months 7/25 (= 28%) of HRT group complained of mastodynia compared to 3/23 placebo (ns). At 6 months only 1/25 HRT participants and 0/23 placebo participants complained of mastodynia.	Marginally significant improvement (p = 0.05) in headaches in HRT group compared to placebo	-	Decreased significantly from baseline in HRT and placebo groups, but no significant difference	-	-	-	No gastric intolerance
Campbell 1976	The number of menopausal women having either breakthrough or withdrawal bleeding increased during the first three cycles of Premarin. Thereafter, 32% experienced withdrawal bleeding and 28% experienced breakthrough bleeding.	7% of women experienced nausea during HRT phase and 3% during placebo phase (x-over study)	13% of women experienced breast pain during HRT phase and 10% during placebo phase (x-over study)	Non-significant improvement during HRT phase	There was an overall decrease in mean weight during the study (not significant) and no significant change in weight between the first and second treatment phases (x-over study)	-	None : 125I-labelled fibrinigen test was performed on all participants on three occasions - no evidence of venous thrombosis of the legs on any occasion	-	-	The most common adverse events were leg cramps, breast tenderness, limb pains, fluid retention, eye irritation, nausea and vaginal discharge. Frequency of all were slightly higher during HRT phase than placebo phase but none significant. (x-over study).

Study	Side-efects									
	Atypical bleeding	Nausea/Vomiting	Breast tenderness	Headaches	Weight changes	Dizziness	Thrombosis	Rash	Pruritis	Other
Chung 1996	-	-	-	No significant difference at end of 1st cross-over phase	-	No significant difference at end of 1st cross-over phase	-	-	-	-
Conard 1995	Rare - only in 2 women (both HRT), resulted in 1 withdrawal from therapy	-	Significant increase in incidence (31.6% HRT vs 5.3% placebo, p = 0.036) at 3 months. Two withdrawals (both in 1mg E2 group) due to breast discomfort.	-	No significant modification of weight or BMI	-	-	-	-	7 withdrawals due to adverse events. Reasons: headaches (one each in 1.5 mg E2 group and placebo), abdominal pain (1 only in placebo) and metrorrhagia (1 only in 1.5 mg E2 group). Adverse events at 3 months were: hot flushes (1 placebo), breast discomfort (2 in 1 mg E2 group) and metrorrhagia (1 in 1.5 mg E2 group)
Coope 1975	No breakthrough bleeding in any women but withdrawal bleeding in majority of perimenopausal women	Two women reported nausea during placebo phase (x-over study)	Three women reported breast pain: 2 during placebo phase and 1 during HRT phase (x-over study)	One woman reported headaches during HRT phase (x-over study)	Overall, no significant increases in weight. Four women reported weight gain of more then 3 kg: 2 during placebo and 2 during HRT phase (x-over study.	-	-	-	-	Predominant side effects (reported more than twice) were nausea, breast discomfort, urinary infection, nasal stuffiness, rise in blood pressure and weight gain
Coope 1981	Regular withdrawal bleeding and 1/36 women with a uterus reported breakthrough bleeding.	-	One case of severe breast swelling attributable to oestrogen therapy	-	-	-	One case of small deep vein thrombosis.	-	-	Side effects possibly attributable to oestrogen therapy were breast swelling (1 case), fluid retention and left ventricular failure with gallop rhythm, basal crepitations and sacral oedema (1 case which

Study	Side-efects									
	Atypical bleeding	Nausea/Vomiting	Breast tenderness	Headaches	Weight changes	Dizziness	Thrombosis	Rash	Pruritis	Other
										resolved on withdrawal from E2), severe depression (2 women), and small deep vein thrombosis (1 women). Two deaths occurred - one case of recurrent gastric Ca and one from epileptic seizure (treatment groups not reported).
Davidsen 1974	Significantly increased frequency of regular bleeding in HRT group and increased frequency of no bleeding in placebo group, but no significant difference in irregular bleeding	No difference between HRT and placebo	No difference between HRT and placebo	-	Four women had weight changes of more than 2 kg (three weighed most in the HRT phase of x-over)	-	-	-	-	No difference in blood pressure between HRT and placebo phases of x-over
Dennerstein 1978	-	-	One woman (during HRT phase of x-over) stopped or changed tablets due to mastalgia.	The two oestrogens preparations were associated with less headaches in the first months of use compared to the two non-oestrogen preparations. This was not significant for the second and third months.	-	-	-	-	-	Predominant reasons for withdrawing or changing tablets included hot flushes during placebo phase and depression/anxiety during HRT phase.
Derman 1995	Resulted in withdrawals from therapy in HRT group (number not specified)	-	-	-	Resulted in withdrawals from therapy in HRT group (number not specified)	-	-	-	-	Predominant reasons for withdrawals from therapy were weight gain, palpitations and bleeding in HRT group and lack

Study	Side-efects									
	Atypical bleeding	Nausea/Vomiting	Breast tenderness	Headaches	Weight changes	Dizziness	Thrombosis	Rash	Pruritis	Other
										of effect in placebo group
Hagen 1982	Most women on HRT developed regular bleeding, but incidence of irregular bleeding was actually (non-significantly) less in HRT group compared to placebo	Slight increase in women with nausea in HRT group, no change in women with nausea in placebo group	Significant increase in breast tenderness from baseline to end of study in both HRT and placebo groups and incidence of breast tenderness significantly higher in HRT group at end of study compared to placebo groups (data from combined HRT and thiazide arms)	-	No significant changes in body weight in relevant HRT and placebo groups (but a significant decrease in the excluded HRT + thiazide group	-	No cases	-	-	Three women (two in placebo groups and one in HRT groups withdrew due to CA of the breast (data includes thiazide groups).
Jensen J 1983	Increased incidence of regular and irregular bleeding in HRT groups compared to placebo	-	-	-	No difference between HRT and placebo groups	-	-	-	-	Reasons for withdrawals due to adverse events included breast tenderness, vaginal bleeding, nervousness and nausea, depression, Rectal Ca and Bronchitis. Not specified by treatment group or number.
Jensen P 1987	-	-	-	-	One woman (HRT therapy) withdrew from treatment due to weight gain	-	-	-	-	Reasons for withdrawals from therapy were weight gain, varicose veins
Marslew 1992	Three women (HRT therapy) withdrew due to regular bleeding	-	Increase in breast tenderness severity rating (significance not stated) in HRT group compared to placebo	-	-	-	-	-	-	Statement of relatively high incidence of adverse events in HRT groups (only one included in this review), but not itemised.

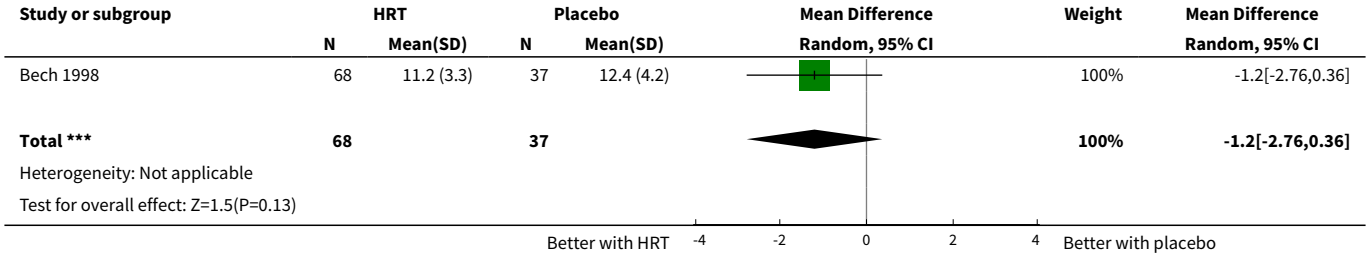
Study	Side-efects									
	Atypical bleeding	Nausea/Vomiting	Breast tenderness	Headaches	Weight changes	Dizziness	Thrombosis	Rash	Pruritis	Other
Martin 1971	Irregular bleeding resulted in two withdrawals from therapy (1 placebo, 1 Rx). Withdrawal bleeding common and the incidence of uterine bleedings at other times very low. Heavy and/or prolonged bleeding was reported by a few women in each group.	-	Breast soreness increased with high E2	Improvement in headaches with high E2, but also improvement in placebo in cycle 3	-	-	-	-	-	Withdrawal rate significantly higher in placebo group ($p < 0.05$) and these women had more severe vasomotory symptoms than those who remained in study (author's analysis)
Notelovitz 2000a	Bleeding and breast pain were the most frequently reported adverse events, both of which tended to be more prevalent in the 2mg E2 group, although there were no significant differences between treatment groups. Bleeding was noted by 14% of women in the placebo group, 10% in the 0.25mg E2 group, 6% in the 0.5mg group, 21% in the 1mg group and 37% in the 2 mg group. Significantly more women discontinued therapy because of bleeding in the 2mg E2 group than	-	Bleeding and breast pain were the most frequently reported adverse events, both of which tended to be more prevalent in the 2mg E2 group, although there were no significant differences between treatment groups. Breast pain was reported by 12% of women treated with 2mg E2, but only by 3-6% of women who received placebo or in the 0.25-1mg E2 groups.	-	-	-	-	-	-	-

Study	Side-efects									
	Atypical bleeding	Nau-sea/Vom-iting	Breast tenderness	Headaches	Weight changes	Dizziness	Throm-bosis	Rash	Pruritis	Other
	in any other group.									
Paterson 1982a	-	New symp-toms oc-cured mar-ginally more fre-quently during HRT phase com-pared to placebo phase (x-over study). No differ-ence in severity score be-tween HRT and place-bo at end of first cross over phase	Occurred only in HRT phase and not in placebo phase (x-over study)	Occurred equally during HRT and place-bo phas-es (x-over study). No difference in severity score be-tween HRT and place-bo at end of first cross-over phase.	No differ-ence in weight change be-tween HRT and place-bo at end of first cross-over phase.	No differ-ence in severity score be-tween HRT and place-bo at end of first cross-over	-	-	-	Other new symp-toms/side effects were de-pres-sion/anxi-ety (slight-ly more fre-quent dur-ing HRT phase) and aching limbs/joints (occurred only dur-ing placebo phase).
PEPI 1998	-	-	Breast dis-comfort was more common with E+P compared to E alone and place-bo	If headaches were ab-sent at baseline, the CEE (E only) group were more likely to develop headache than place-bo or E+P groups. Con-versely, if headaches were present at baseline, the E only group re-ported less headaches than the placebo or E+P groups	The CEE + micro-nised prog-esterone treated group were more likely to report a perceived weight loss com-pared to the placebo group. The was no dif-ference for the other HRT groups compared to placebo.	-	There were two cas-es of deep vein throm-bophlebitis in the CEE only group and one case of su-perficial phlebitis in the CEE + micro-nised prog-esterone group	-	-	The most common side effects associat-ed with withdrawal from treat-ment in the placebo group were headaches and joint pain com-pared to adenoma-tous hy-perplasia & atypia (E only), vaginal bleeding, headaches, depres-sion , and breast cancer in the HRT groups. There was difficulty with lack of effect in the place-bo group and by the end of the study, 11% of the women in the place-bo group had begun taking pri-vately pre-scribed HRT
Symons 2000 Study 1	Overall, the incidence of bleed-ing and spotting	-	One of 10 women who with-drew from the study	Three of 10 women who with-drew from the study	One of 10 women who with-drew from the study	-	-	-	-	Two of 10 women who with-drew from the study

Study	Atypical bleeding	Nausea/Vomiting	Side-efects Breast tenderness	Headaches	Weight changes	Dizziness	Thrombosis	Rash	Pruritis	Other
	tended to be greater with increasing doses of E+P and decreased over time within each treatment group. The percentage of women with bleeding and spotting in the placebo group ranged from 3-5% and was as high as 30% in the 1mg NA/10ug EE treatment group in the first 8 weeks of the study. Three of 10 women who withdrew from the study because of medication related adverse events did so because of bleeding.		because of medication related adverse events did so because of breast tenderness.	because of medication related adverse events did so because of headaches.	because of medication related adverse events did so because of increased appetite.					because of medication related adverse events did so because of insomnia & depression respectively.
Symons 2000 Study 2	Overall, the incidence of bleeding and spotting tended to be greater with increasing doses of E+P and decreased over time within each treatment group. The percentage of women with bleeding and spotting in the placebo group was approximately 5% and greater than 33% in the 1mg NA/10ug EE treatment group by week 4 of the study.	-	Two of nine women who withdrew from the study because of medication related adverse events did so because of breast tenderness.	Two of nine women who withdrew from the study because of medication related adverse events did so because of headaches.	One of nine women who withdrew from the study because of medication related adverse events did so because of bloating.	-	One of nine women who withdrew from the study because of medication related adverse events did so because of superficial thrombophlebitis.	-	-	Two of 9 women who withdrew from the study because of medication related adverse events did so because of nervousness & palpitations respectively.

Study	Side-efects									
	Atypical bleeding	Nau-sea/Vom-iting	Breast tenderness	Headaches	Weight changes	Dizziness	Throm-bosis	Rash	Pruritis	Other
	One of 9 women who with-drew from the study because of medica-tion relat-ed adverse events did so because of bleeding.									
Viklylaeva 1997	Regular bleeding with Trise-quens but no exces-sive or un-controlled bleeding	-	-	Marginal-ly signifi-cant (p = 0.1) greater reduction from base-line in HRT group com-pared to placebo	-	No signif-icant dif-ference in change from base-line be-tween HRT and place-bo groups	-	-	-	-

Analysis 6.5. Comparison 6 Other outcomes: any HRT vs placebo, Outcome 5 Quality of life.

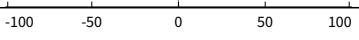


Comparison 7. Comparison of end of study and baseline vasomotor outcomes for women randomised to placebo therapy

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Hot flush frequency/week	6	367	Mean Difference (IV, Random, 95% CI)	-25.37 [-37.55, -13.19]
2 Hot flush frequency - log transformed	6	367	Mean Difference (IV, Random, 95% CI)	-0.86 [-1.13, -0.60]
3 HFWWS	2	205	Mean Difference (IV, Random, 95% CI)	-77.95 [-126.35, -29.55]
4 HFWWS - log transformed	2	205	Mean Difference (IV, Random, 95% CI)	-0.71 [-0.88, -0.55]
5 Hot flush severity (all scales, continuous) - SMD	6	332	Std. Mean Difference (IV, Random, 95% CI)	-0.66 [-1.25, -0.08]

Analysis 7.1. Comparison 7 Comparison of end of study and baseline vasomotor outcomes for women randomised to placebo therapy, Outcome 1 Hot flush frequency/week.

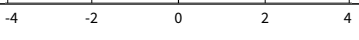
Study or subgroup	End of Study N	End of Study Mean(SD)	Baseline N	Baseline Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	33	33.5 (31)	41	59.5 (30.5)		17.56%	-26[-40.11,-11.89]
Conard 1995	15	32.2 (35.6)	19	51.9 (21.6)		13.95%	-19.74[-40.18,0.7]
Coope 1975	15	21.9 (26.8)	15	56.8 (46.1)		10.79%	-34.87[-61.86,-7.88]
Coope 1981	26	16.3 (26.9)	26	27.8 (30.5)		16.66%	-11.5[-27.13,4.13]
Notelovitz 2000a	55	28 (29)	66	72 (21)		20.31%	-44[-53.19,-34.81]
Viklylaeva 1997	28	23.7 (16.5)	28	40.3 (15.3)		20.73%	-16.57[-24.9,-8.24]
Total ***	**172**		**195**			**100%**	**-25.37[-37.55,-13.19]**
Heterogeneity: Tau^2=168.19; Chi^2=23.72, df=5(P=0); I^2=78.92%							
Test for overall effect: Z=4.08(P<0.0001)							



Less at end of study — Less at baseline

Analysis 7.2. Comparison 7 Comparison of end of study and baseline vasomotor outcomes for women randomised to placebo therapy, Outcome 2 Hot flush frequency - log transformed.

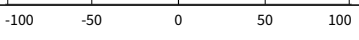
Study or subgroup	End of Study N	End of Study Mean(SD)	Baseline N	Baseline Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	33	3.9 (0.8)	41	4.7 (0.5)		20.98%	-0.75[-1.05,-0.45]
Conard 1995	15	1.4 (0.9)	19	2.1 (0.4)		14.82%	-0.69[-1.17,-0.21]
Coope 1975	15	2.1 (1.7)	15	3.5 (1.4)		4.73%	-1.42[-2.54,-0.3]
Coope 1981	26	2.2 (1.1)	26	3 (0.9)		13.11%	-0.74[-1.28,-0.2]
Notelovitz 2000a	55	3 (0.8)	66	4.3 (0.3)		23.8%	-1.23[-1.46,-1]
Viklylaeva 1997	28	3 (0.6)	28	3.7 (0.4)		22.57%	-0.64[-0.9,-0.38]
Total ***	**172**		**195**			**100%**	**-0.86[-1.13,-0.6]**
Heterogeneity: Tau^2=0.06; Chi^2=14.54, df=5(P=0.01); I^2=65.62%							
Test for overall effect: Z=6.36(P<0.0001)							



Less at end of study — Less at baseline

Analysis 7.3. Comparison 7 Comparison of end of study and baseline vasomotor outcomes for women randomised to placebo therapy, Outcome 3 HFWWS.

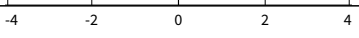
Study or subgroup	End of Study N	End of Study Mean(SD)	Baseline N	Baseline Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	35	62.6 (13.2)	40	116.3 (13.6)		50.91%	-53.7[-59.77,-47.63]
Notelovitz 2000a	64	76.9 (9.6)	66	180 (59)		49.09%	-103.1[-117.53,-88.67]
Total ***	**99**		**106**			**100%**	**-77.95[-126.35,-29.55]**
Heterogeneity: Tau^2=1188.29; Chi^2=38.26, df=1(P<0.0001); I^2=97.39%							
Test for overall effect: Z=3.16(P=0)							



Less at end of study — Less at baseline

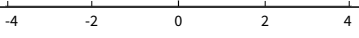
Analysis 7.4. Comparison 7 Comparison of end of study and baseline vasomotor outcomes for women randomised to placebo therapy, Outcome 4 HFWWS - log transformed.

Study or subgroup	End of Study N	End of Study Mean(SD)	Baseline N	Baseline Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Baerug 1998	35	4.1 (0.2)	40	4.8 (0.1)		50.27%	-0.63[-0.71,-0.55]
Notelovitz 2000a	64	4.4 (0.1)	66	5.2 (0.3)		49.73%	-0.8[-0.88,-0.72]
Total ***	**99**		**106**			**100%**	**-0.71[-0.88,-0.55]**
Heterogeneity: Tau^2=0.01; Chi^2=8.51, df=1(P=0); I^2=88.25%							
Test for overall effect: Z=8.41(P<0.0001)							



Analysis 7.5. Comparison 7 Comparison of end of study and baseline vasomotor outcomes for women randomised to placebo therapy, Outcome 5 Hot flush severity (all scales, continuous) - SMD.

Study or subgroup	End of Study N	End of Study Mean(SD)	Baseline N	Baseline Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Bech 1998	37	0.8 (0.9)	37	0.7 (0.8)		18.65%	0.12[-0.34,0.57]
Blumel 1994	23	0.6 (0.8)	25	1.1 (1)		17.44%	-0.57[-1.15,0.01]
Chung 1996	40	0.8 (0.9)	40	1 (1)		18.8%	-0.22[-0.65,0.22]
Conard 1995	15	0.8 (0.2)	19	2.3 (0.7)		13.09%	-2.87[-3.86,-1.88]
Derman 1995	39	9.4 (4.5)	39	10.9 (4.7)		18.73%	-0.32[-0.77,0.13]
Paterson 1982a	9	2.3 (0.9)	9	2.9 (0.3)		13.3%	-0.82[-1.79,0.16]
Total ***	**163**		**169**			**100%**	**-0.66[-1.25,-0.08]**
Heterogeneity: Tau^2=0.42; Chi^2=30.58, df=5(P<0.0001); I^2=83.65%							
Test for overall effect: Z=2.22(P=0.03)							



Comparison 8. Investigations of assumptions

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Comparison of logistic regression (LR) and approximated log odds ratios (LOR) assuming proportional odds (PO)			Other data	No numeric data

Analysis 8.1. Comparison 8 Investigations of assumptions, Outcome 1 Comparison of logistic regression (LR) and approximated log odds ratios (LOR) assuming proportional odds (PO).

		Comparison of logistic regression (LR) and approximated log odds ratios (LOR) assuming proportional odds (PO)				
Study	Time	LOR (LR)	SE (LR)	Fit of PO Model (LR)	LOR (approximation)	SE (approximation)
Blumel 1994	Baseline	0.223	0.387	p = 0.89	0.223	0.567
Blumel 1994	3 months	-0.806	0.607	p = 0.78	-0.739	0.588
Blumel 1994	6 months	-1.791	0.615	p = 0.85	-1.738	0.634
Chung 1996	Baseline	0.224	0.310	p = 0.52	0.229	0.440

Comparison of logistic regression (LR) and approximated log odds ratios (LOR) assuming proportional odds (PO)

Study	Time	LOR (LR)	SE (LR)	Fit of PO Model (LR)	LOR (approximation)	SE (approximation)
Chung 1996	6 months	-1.167	0.380	p = 0.47	-1.009	0.453
Chung 1996	-	-	-	-	-	-
Conard 1995	Baseline	-0.479	0.538	p = 0.19	-0.439	0.567
Conard 1995	3 months	-3.284	0.633	p = 0.82	-3.387	0.945
Conard 1995	-	-	-	-	-	-

ADDITIONAL TABLES

Table 1. Quality Assesment Criteria

Assessment	A	B	C
Allocation concealment	Adequate e.g. central randomisation / allocation, sealed envelopes etc	Not reported/unclear	Inadequate
Treatment blinding	Statement that containers were identical, drugs were identical in appearance etc	Not reported/unclear	HRT and placebo not identical
Outcome assessment	Blinded, standardised assessment	Assessment procedures not stated	Assessment not blinded or standardised
Baseline equality of treatment groups	Groups balanced in terms of age, menopause status, and menopause symptoms	Balance not reported	Groups not balanced
Losses to follow-up (not including early cessation of therapy, followed up)	Losses of 10% or less	Not reported/Unclear	Losses of more than 10%
Basis for analysis	Intention-to-treat analysis	Unclear	Not intention-to-treat

WHAT'S NEW

Date	Event	Description
12 November 2008	Amended	Converted to new review format.

HISTORY

Protocol first published: Issue 4, 1997
Review first published: Issue 1, 2000

Date	Event	Description
17 August 2004	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

MacLennan A: wrote the protocol, searched the literature, visited the Menstrual Disorders and Subfertility Review Group centre in Auckland, New Zealand for Cochrane literature, sought data from authors and other sources e.g. pharmaceutical companies, performed the quality analysis of the trials, co-authored the text of the review.
Broadbent J: contributed to the 2004 update - reviewed studies for inclusion and exclusion, corresponded with and sought data from authors and other sources, performed the quality analysis of the trials, performed the data extraction, edited the text of the review.
Lester S: performed the data extraction and data analysis, co-authored the text of the review.
Moore V: checked data extraction, provided statistical advice on the data analysis, edited the text of the review.

DECLARATIONS OF INTEREST

Alastair MacLennan is the Co-Editor-in-Chief of Climacteric, the Journal of the International Menopause Society. He has been the recipient of research grants awarded by various pharmaceutical companies for clinical trials of new products and epidemiological studies on menopause and grants for educational videos on menopause.

SOURCES OF SUPPORT

Internal sources

- Department of Obstetrics & Gynaecology, The University of Adelaide, Australia.

External sources

- Australian Menopause Society grants, 1999, 2000, Australia.

NOTES

Original review title (no longer in use): Oral oestrogen replacement therapy versus placebo for hot flushes. Title changed with this update.

INDEX TERMS

Medical Subject Headings (MeSH)

Administration, Oral; Estrogen Replacement Therapy [*methods]; Hot Flashes [*drug therapy]; Placebo Effect; Randomized Controlled Trials as Topic

MeSH check words

Female; Humans