Cochrane
Library

Cochrane Database of Systematic Reviews

Hormone replacement therapy for cognitive function in postmenopausal women (Review)

Lethaby A, Hogervorst E, Richards M, Yesufu A, Yaffe K

Lethaby A, Hogervorst E, Richards M, Yesufu A, Yaffe K.
Hormone replacement therapy for cognitive function in postmenopausal women.
Cochrane Database of Systematic Reviews 2008, Issue 1. Art. No.: CD003122.
DOI: [10.1002/14651858.CD003122.pub2](https://doi.org/10.1002/14651858.CD003122.pub2).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	3
OBJECTIVES	3
METHODS	3
RESULTS	6
DISCUSSION	9
AUTHORS' CONCLUSIONS	11
ACKNOWLEDGEMENTS	11
REFERENCES	12
CHARACTERISTICS OF STUDIES	15
DATA AND ANALYSES	25
Analysis 1.1. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 1 CAMCOG after 20 weeks of E2 versus placebo.	26
Analysis 1.2. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 2 3MSE score after CEE + MPA (variable follow up).	27
Analysis 1.3. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 3 3MSE score after CEE alone (variable follow up).	27
Analysis 1.4. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 4 Folstein Mini-Mental State Examination with CEE alone or CEE + MPA after 3 years.	28
Analysis 2.1. Comparison 2 The effect of ERT/HRT versus placebo on prevention of mild cognitive impairment, Outcome 1 Mild cognitive impairment (based on WHIMS protocol).	29
Analysis 3.1. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 1 Paired Associates (intermediate recall) after HRT or placebo.	32
Analysis 3.2. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 2 Paired associate learning (different scoring methods immediate recall).	33
Analysis 3.3. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 3 Paired Associate (delayed recall) after HRT or placebo.	33
Analysis 3.4. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 4 Other Paired Associate tests.	34
Analysis 3.5. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 5 Paragraph recall (immediate) after 2 to 3 months of E2 i.m. or placebo.	34
Analysis 3.6. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 6 Paragraph recall (delayed) after 2 months with E2 i.m. or placebo.	35
Analysis 3.7. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 7 Verbal Fluency (semantic memory).	35
Analysis 3.8. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 8 Word list recall tests.	36
Analysis 3.9. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 9 Wechsler Memory Scale - Revised (logical memory - immediate recall).	37
Analysis 3.10. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 10 Wechsler Memory Scale - Revised (logical memory - delayed recall).	37
Analysis 3.11. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 11 Wechsler Memory Scale - Revised (logical memory - % retained) after 10 weeks E2 transdermal or placebo.	38
Analysis 3.12. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 12 Boston Naming Test after low E transdermal versus placebo.	38
Analysis 3.13. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 13 PMA Vocabulary after 1+ years of HRT versus placebo.	38
Analysis 3.14. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 14 Verbal Fluency - letter after 1+ years of HRT versus placebo.	39
Analysis 3.15. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 15 Verbal Fluency - categories after 1+ years of HRT versus placebo.	39

Analysis 3.16. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 16 Word list recall tests after 1+ years of HRT versus placebo (inverse variance).	39
Analysis 4.1. Comparison 4 The effect of ERT/HRT versus placebo on visual memory, Outcome 1 Visuo(spatial) memory tests. ..	41
Analysis 4.2. Comparison 4 The effect of ERT/HRT versus placebo on visual memory, Outcome 2 Visuospatial memory tests (inverse variance).	43
Analysis 5.1. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 1 Test of simple speed of information processing.	44
Analysis 5.2. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 2 Tests of other information processing.	45
Analysis 5.3. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 3 Attention and working memory tests after 1+ year HRT versus placebo (inverse variance).	45
Analysis 5.4. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 4 Fine Motor Speed after 1+ yrs of HRT versus placebo (inverse variance).	46
Analysis 5.5. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 5 Tests of general information processing speed and recognition.	46
Analysis 7.1. Comparison 7 The effect of ERT/HRT versus placebo on tests of mental rotations, Outcome 1 Mental Rotations time (120 degrees only) after 3 weeks of E2 transdermal or placebo.	47
Analysis 7.2. Comparison 7 The effect of ERT/HRT versus placebo on tests of mental rotations, Outcome 2 Mental Rotations number after 2 weeks of E2 transdermal or placebo.	47
Analysis 7.3. Comparison 7 The effect of ERT/HRT versus placebo on tests of mental rotations, Outcome 3 Card rotations after 1+ year of HRT versus placebo (inverse variance).	47
Analysis 8.1. Comparison 8 The effect of ERT/HRT versus placebo on accuracy, Outcome 1 Clerical speed and accuracy (# correct) after 3 months of E2 i.m. or placebo.	48
Analysis 8.2. Comparison 8 The effect of ERT/HRT versus placebo on accuracy, Outcome 2 Letter Cancellation (# correct) after 9 months of CEE or placebo.	48
Analysis 8.3. Comparison 8 The effect of ERT/HRT versus placebo on accuracy, Outcome 3 Random number generation after 3 weeks of E2 transdermal or placebo.	49
Analysis 9.1. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 1 Wisconsin Card Sorting Test (categories) after 10 weeks E2 transdermal or placebo.	50
Analysis 9.2. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 2 Wisconsin Card Sorting Test (perseverative responses) after 10 weeks E2 transdermal or placebo.	50
Analysis 9.3. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 3 Wisconsin Card Sorting Test (% conceptual level responses) after 10 weeks E2 transdermal or placebo.	50
Analysis 9.4. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 4 Weschler Adult Intelligence Scale - Revised (picture arrangement) after 10 weeks E2 transdermal or placebo.	51
Analysis 9.5. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 5 Tower of London after 3 weeks of E2 transdermal or placebo.	51
Analysis 9.6. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 6 Abstract reasoning (DAT) after 3 months of IM E2 or placebo.	51
Analysis 9.7. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 7 Controlled Word Association Test (COWAT) after 10 weeks E2 transdermal or placebo.	51
Analysis 9.8. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 8 Stroop test (interference time) after 10 weeks of E2 transdermal or placebo.	52
Analysis 9.9. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 9 Stroop (interference errors) after 10 weeks of E2 transdermal or placebo.	52
Analysis 9.10. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 10 Trails B with either CEE + MPA or low E transdermal versus placebo.	52
Analysis 9.11. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 11 Digits Backward after 1+ year HRT versus placebo.	53
ADDITIONAL TABLES	53
WHAT'S NEW	57
HISTORY	57
CONTRIBUTIONS OF AUTHORS	57
DECLARATIONS OF INTEREST	57
SOURCES OF SUPPORT	57
INDEX TERMS	58

[Intervention Review]

Hormone replacement therapy for cognitive function in postmenopausal women

Anne Lethaby¹, Eva Hogervorst², Marcus Richards³, Amina Yesufu², Kristine Yaffe⁴

¹Section of Epidemiology & Biostatistics, School of Population Health, University of Auckland, Auckland, New Zealand. ²Department of Human Sciences, Loughborough University, Loughborough, UK. ³University College London Department of Epidemiology & Public Health, London, UK. ⁴Psychiatry, Neurology and Epidemiology, University of California, San Francisco, San Francisco, California, USA

Contact address: Anne Lethaby, Section of Epidemiology & Biostatistics, School of Population Health, University of Auckland, Private Bag 92019, Auckland, 1142, New Zealand. a.lethaby@auckland.ac.nz.

Editorial group: Cochrane Dementia and Cognitive Improvement Group

Publication status and date: Edited (no change to conclusions), published in Issue 4, 2008.

Citation: Lethaby A, Hogervorst E, Richards M, Yesufu A, Yaffe K. Hormone replacement therapy for cognitive function in postmenopausal women. *Cochrane Database of Systematic Reviews* 2008, Issue 1. Art. No.: CD003122. DOI: [10.1002/14651858.CD003122.pub2](https://doi.org/10.1002/14651858.CD003122.pub2).

Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

As estrogens have been found in animal models to be associated with the maintenance and protection of brain structures, it is biologically plausible that maintaining high levels of estrogens in postmenopausal women by medication could be protective against cognitive decline.

Objectives

To investigate the effect of ERT (estrogens only) or HRT (estrogens combined with a progestagen) in comparison with placebo in RCTs on cognitive function in postmenopausal women.

Search methods

The CDCIG Specialized Register was searched 7 March 2006. Additional searches were made of MEDLINE (1966-2006/02); EMBASE (1985-2006/02); PsycINFO (1967-2006/02) and CINAHL (1982-2006/01).

Selection criteria

All double-blind RCTs trials of the effect of ERT or HRT on cognitive function over a treatment period of at least two weeks in postmenopausal women.

Data collection and analysis

Selection of studies, assessment of quality and extraction of data were undertaken independently by three reviewers with disagreements resolved by discussion.

Main results

In total, 24 trials were included, but only 16 (10,114 women) had analysable data. Meta-analyses showed no effects of either ERT or HRT on prevention of cognitive impairment after five and four years of treatment, respectively (odds ratio 1.34, 95% CI 0.95 to 1.9; odds ratio 1.05, 95% CI 0.72 to 1.54 respectively) (trend favouring control in both instances). Analyses assessing the effects of treatment over time found that both ERT and HRT did not maintain or improve cognitive function and may even adversely affect this outcome (WMD = -0.45, 95% CI -0.99 to 0.09; WMD = -0.16, 95% CI -0.58 to 0.26, respectively at maximum follow up). Negative effects were found for ERT after one year and HRT after three and four years of therapy. Results from smaller trials assessing effects on individual cognitive domains mostly reported no evidence of benefit.

Authors' conclusions

There is good evidence that both ERT and HRT do not prevent cognitive decline in older postmenopausal women when given as short term or longer term (up to five years) therapy. It is not known whether either specific types of ERT or HRT have specific effects in subgroups of women, although there was evidence that combined hormone therapy in similarly aged women was associated with a decrement in a number of verbal memory tests and a small improvement in a test of figural memory. There is insufficient evidence to determine whether subgroups of women using specific types of hormone therapy could benefit from treatment. It remains to be determined whether factors such as younger age (< 60 years of age), type of menopause (surgical or natural) and type of treatment (type of estrogen with or without a progestagen), mode of delivery (transdermal, oral or intramuscular) and dosage have positive effects at a clinically relevant level. In addition, whether the absence or presence of menopausal symptoms can modify treatment effects should be investigated in more detail. Large RCTs currently underway in the USA may be able to provide answers to these uncertainties by the year 2010. In the meantime, based on the available evidence, ERT or HRT cannot be recommended for overall cognitive improvement or maintenance in older postmenopausal women without cognitive impairment.

PLAIN LANGUAGE SUMMARY

There is good evidence that estrogen or combined estrogen and progestagen therapy does not protect against a decline in overall cognitive functioning of older postmenopausal women with normal intellectual ability

Animal studies (performed both in the laboratory and on living animals) suggest that estrogen alone might protect the brain as women get older. After the menopause, levels of estrogens decline in women and estrogen therapy has been claimed to maintain or enhance cognitive function in postmenopausal women. This review found no evidence of a benefit of any types of estrogen on overall cognitive functioning in older postmenopausal women when given either as short term or longer term (up to five years) therapy. There was also no evidence of a beneficial effect of combined estrogen and progestagen therapy overall. There was insufficient evidence to determine whether any type of hormone replacement therapy had beneficial or harmful effects on specific types of cognitive ability, such as verbal or visual memory.

BACKGROUND

Nearly all cognitive functions decline, on average, with age, but there is a large variability which ranges from "successful" aging to dementia (Huppert 1997). The determinants of this variability are uncertain but elderly women seem to have a higher risk of developing Alzheimer's disease (AD) than elderly men (Launer 1999). While this may be because they reach an older age and aging is a risk factor for AD, the age-specific incidence of AD is also higher in women than in men (Fratiglioni 2000). It has been suggested that steroid hormone deficiencies in elderly women may play a role in this difference.

Estrogens (UK spelling oestrogens) are steroid hormones produced in women by their ovaries. The estrogen-producing cells become depleted at menopause and postmenopausal women have much lower estrogen levels than men. In men, and in women after the menopause, the main source of estrogens is from conversion of circulating androgen steroid hormone precursors. In women, the main source of androgen steroids are the ovaries (theca cells) and the adrenal cortex, whereas in men it is the testes. Estrogens have an important role in the female reproductive cycle, but animal and in vivo cell studies have suggested that estrogens can also protect brain structures, including those related to memory, such as the hippocampus and the cholinergic basal forebrain (McEwen 1997). A variety of mechanisms appear to be involved in this process, including anti-amyloidogenic effects, antioxidant effects, dendritic sprouting and effects on various neurotransmitters involved in cognitive function (McEwen 1997; Silva 2001).

It is possible that maintaining high levels of estrogens in women by means of estrogen replacement therapy (ERT) or combined therapy with estrogens and progestagens (hormone replacement therapy - HRT) could be protective against cognitive decline and the development of AD or other dementia syndromes (see Hogervorst 2002 b). Postmenopausal ERT or HRT is usually prescribed to treat menopausal symptoms, such as hot flashes (UK hot flushes) and night sweats, or as treatment or prevention for osteoporosis. Prior to 2002, several large observational studies suggested that ERT or HRT also had a role in the prevention of cerebrovascular disease. However, several recent randomized trials have shown that the harms outweigh the benefits of HRT or ERT; an increased association of HRT with CVD, stroke, VTE and breast cancer and an increased association of ERT with stroke and breast cancer have been reported (Farquhar 2005). Many of these trials had large sample sizes with adequate power, long follow up, clinically relevant endpoints and included largely older women with natural menopause. Thus, HRT or ERT currently only have a role only in the short term treatment of intolerable hot flashes or night sweats in women close to the menopause and these medications are no longer considered first line for the prevention of osteoporosis. For hysterectomized postmenopausal women, replacement therapy is usually given as ERT, but HRT is prescribed for women with intact uteri to reduce the risk of endometrial hypertrophy or cancer.

Several observational studies have found that women who have used HRT or ERT after their menopause have better cognitive function than controls (Hogervorst 2000; Yaffe 1998 a), but, as was the case in the assessment of the relationship between HRT/ERT and cerebrovascular disease, observational studies are subject to bias (Barrett-Connor 1991). Women who choose to use HRT or ERT after the menopause have in general higher education,

healthier lifestyles and are healthier before using ERT or HRT than women who do not (Matthews 1996). Taking ERT or HRT may therefore be associated with a healthier lifestyle in general which in turn could be associated with better cognitive function. In addition, despite estrogen's biologically plausible mechanisms for protecting the aging brain, most reviews have concluded that the human studies have substantial methodological problems and, as a consequence, results have been conflicting (Hogervorst 2000; Yaffe 1998 a). A recent review (Maki 2006) has suggested that the effects could be differential in younger as opposed to older postmenopausal women. This review has presented results separately for women aged less than 65 years and greater than 65 years. The authors found good evidence that long-term hormone treatment in older women did not improve cognitive functioning. They suggested that there could be a benefit of estrogen alone in younger postmenopausal women but the included trials were unconvincing because of their small size and short duration.

Prior to this update, most studies evaluating the effects of ERT or HRT on cognitive abilities in postmenopausal women were of short duration and many had numerous methodological flaws, giving rise to uncertain conclusions. The publication of a number of large studies with adequate periods of follow-up and clinically relevant endpoints has prompted this update.

See Table 1 for a full list of abbreviations used in this review.

OBJECTIVES

To assess the effects of therapy consisting of estrogens alone (ERT) or in combination with a progestagen (HRT) on cognitive function of postmenopausal women.

METHODS

Criteria for considering studies for this review

Types of studies

All double-blind randomized placebo-controlled trials (RCTs) in which ERT or HRT was administered to postmenopausal women without dementia for at least two weeks and cognitive function was measured were eligible for inclusion.

Types of participants

Healthy women of any age who had undergone natural or surgical menopause. Women were considered postmenopausal if they had not menstruated for more than 12 months.

Types of interventions

Interventions containing estrogens alone (ERT) or combined with a progestagen (HRT) at any dose, any dosing schedule and any mode of administration - oral, subdermal, transdermal or intravenous - were eligible for inclusion.

The most commonly used ERTs are:

CEE = conjugated equine estrogens (always orally in dosages 0.625 or 1.25 mg, contains E1-S = estrone sulphate)

E2 = estradiol (usually given transdermally (0.1 to 0.005 mg; Polo-Kantola 1998) or as a bolus injection of 10 mg/month (Phillips 1992; Sherwin 1988; Sherwin 1990) or sometimes orally (1 to 2 mg/daily, Fedor-Freyberg 1977).

Estrogens may be given in combination with a progestagen which is usually MPA (= medroxyprogesterone acetate) or P (= progesterone) (Goebel 1995).

Types of outcome measures

The primary outcome of interest is cognitive function (measured either by validated neuropsychological tests or by strict protocols): The following outcomes were relevant:

A. Tests of overall global cognitive function:

- (1) CAMCOG
- (2) Folstein Mini-Mental State Examination
- (3) Modified Mini-Mental State Examination

B. Cognitive impairment (defined and assessed by strict protocol)

C. Specific tests for individual cognitive domains have been categorized as follows:

- (1) Verbal Memory and Language Tests (Paired Associate Learning and Paragraph Recall or Logical Memory from the Wechsler Memory Scale or WMS; Word List Memory; Word List Recall; California Verbal Learning Test; verbal Fluency with animals, vegetables, fruits or letters; Boston Naming Test; PMA Vocabulary Test)
- (2) Visuospatial (Visual Retention tests or VRT from the WMS; the Benton Visual Retention Test; The Figure of Rey; Block Design; Visuospatial Memory Test)
- (3) Speed Tests, Attention and Manual Dexterity Semantic Memory (Trail Making Test, part A or TMT-A; Digit Symbol Substitution Test or DSST; reaction time (RT) tests; finger tapping; Grooved Pegboard Test; Stroop Tasks A and B)
- (4) Mental Rotation tests and Accuracy (Clerical Speed and Accuracy, Letter Cancellation tests etc.)
- (5) Executive function (Wisconsin Card Sorting Tests; Wechsler Adult Intelligence Scale (picture arrangement); Tower of London; Stroop Tests (interference); TMT-B (Trails B); COWAT; Digits Backward)

Some of the tests have multiple functions and may span other categories to the one in which they have been placed.

Search methods for identification of studies

The trials were identified from a search of the Specialized Register of the Cochrane Dementia and Cognitive Improvement Group on 7 March 2006 using the search terms: HRT, ERT, ORT, PORT, Oestrogen*, estrogen*, progesteron*, progestagen, estradiol.

The Specialized Register at that time contained records from the following databases:

- CENTRAL: July 2005 (issue 3);
- MEDLINE: 1966 to 2005/08, week 2;
- EMBASE: 1980 to 2005/08, week 2;
- PsycINFO: 1887 to 2005/07;
- CINAHL: 1982 to 2004/07;
- SIGLE (Grey Literature in Europe): 1980 to 2004/06;
- ISTP (Index to Scientific and Technical Proceedings): to May 2000;
- INSIDE (BL database of Conference Proceedings and Journals): to June 2000;
- Aslib Index to Theses (UK and Ireland theses): 1970 to March 2003;
- Dissertation Abstract (USA): 1861 to March 2003;

- <http://clinicalstudies.info.nih.gov/>;
- National Research Register (issue 3/2005)
- ClinicalTrials.gov: last searched 1 March 2006;
- LILACS: Latin American and Caribbean Health Science Literature: last searched April 2003
- <http://www.forestclinicaltrials.com/>: last searched 1 September 2005
- ClinicalStudyResults.org: last searched 1 September 2005
- <http://www.lillytrials.com/index.shtml>: last searched 28 August 2005
- ISRCTN Register: last searched 1 September 2005
- IPFMA Clinical trials Register: www.ifpma.org/clinicaltrials.html: last searched September 2005

The search strategies used to identify relevant records in MEDLINE, EMBASE, PsycInfo, CINAHL and LILACS can be found in the Group's module.

In addition MEDLINE (1966-2006/02); EMBASE (1985-2006/02); PsycINFO (1967-2006/02) and CINAHL (1982-2006/01) were searched as the CDCIG Register does not contain all trials with healthy volunteers.

MEDLINE

- #1 "Hormone-Replacement-Therapy"/ all subheadings
- #2 HRT or ERT or ORT or PORT or oestrogen* or estrogen* or progesteron*
- #3 estradiol* or progestagen*
- #4 #1 or #2 or #3
- #5 cognit* or QOL or "quality of life" or mood* or memor*
- #6 #4 and #5
- #7 #6 and (post-menopausal or "post menopausal")
- #8 #7 and (random* or double-blind* or placebo*)

EMBASE

- #1 explode "Hormone-Replacement-Therapy"/ all subheadings
- #2 HRT or ERT or ORT or PORT or oestrogen* or estrogen* or progesteron*
- #3 estradiol* or progestagen*
- #4 #1 or #2 or #3
- #5 cognit* or QOL or "quality of life" or mood* or memor*
- #6 #4 and #5
- #7 #6 and (post-menopausal or "post menopausal")
- #8 #7 and (random* or double-blind* or placebo*)

PsycINFO

- #1 "Hormone-Therapy" in MJ,MN
- #2 HRT or ERT or ORT or PORT or oestrogen* or estrogen* or progesteron*
- #3 cognit* or QOL or "quality of life" or mood* or memor*
- #4 (#1 or #2) and #3
- #5 #4 and (post-menopausal or "post menopausal")

CINAHL

- #1 "Hormone-Replacement-Therapy"/ all subheadings
- #2 HRT or ERT or ORT or PORT or oestrogen* or estrogen* or progesteron*
- #3 estradiol* or progestagen*
- #4 #1 or #2 or #3
- #5 cognit* or QOL or "quality of life" or mood* or memor*
- #6 #4 and #5

#7 #6 and (post-menopausal or "post menopausal")
#8 #7 and (random* or double-blind* or placebo*)

Studies were eligible for inclusion if they were published in the English language.

Data collection and analysis

Selection of studies

Abstracts of the references retrieved by the search were read by three reviewers (EH, AL and KY) in order to discard those that were clearly not eligible for inclusion (such as those that were not randomized, placebo-controlled or double blinded, or where cognitive function was not measured). The reviewers studied the full text of the remaining references and independently selected studies for inclusion. Any disparity in the resulting lists was resolved by discussion with all reviewers in order to arrive at the final list of included studies.

Quality assessment

Two reviewers assessed the quality of the studies according to the Cochrane Collaboration guidelines which focus on the concealment of allocation of treatment.

Category A (adequate) is where the report describes allocation of treatment by: (1) some form of centralized randomized scheme, such as having to provide details of an enrolled participant to an office by phone to receive the treatment group allocation; (2) some form of randomization scheme controlled by a pharmacy; (3) numbered or coded containers, such as in a pharmaceutical trial in which capsules from identical-looking numbered bottles are administered sequentially to enrolled participants; (4) an on-site or coded computer system, given that the allocations were in a locked, unreadable file that could be accessed only after inputting the characteristics of an enrolled participant; or (5) if assignment envelopes were used, the report should at least specify that they were sequentially numbered, sealed, and opaque; (6) other combinations of described elements of the process that provides assurance of adequate concealment.

Category B (unclear) is where the report does not include a description of how allocation to randomized groups is concealed and authors do not respond to requests for this information. Examples include: (1) assignment envelopes without description of adequate safeguards; (2) use of a "list" or "table"; (3) flip of a coin; (4) a trial in which the description suggests adequate concealment but other features are suspicious - for example, markedly unequal control and treatment groups, a trial described as random but unable to obtain further details.

Category C (inadequate) is where the report describes allocation of treatment by: (1) alternation; (2) reference to case record numbers, dates of birth, day of week, or any other such approach; (3) any allocation procedure that is entirely transparent before assignment, such as an open list of random numbers or assignments.

Empirical research has shown that lack of adequate allocation concealment is associated with bias. Trials with unclear concealment measures have been shown liable to yield more pronounced estimates of treatment effects than trials that have

taken adequate measures to conceal allocation schedules, albeit to a less extent than inadequately concealed trials.

Other aspects of the trial quality (methodology, statistics) were assessed and included in an additional table of quality features.

Data collection

Data for the meta-analyses were based on reported summary statistics for each study. To test cognitive change after treatment, the main outcome of interest was the change from baseline to final assessment (mean difference in performance and SD of the mean difference in performance). The baseline assessment was defined as the last available assessment prior to randomization. In the case of large trials with long follow-up, the difference between the treatment and control groups was extracted at follow-up.

Data analysis

Clinical heterogeneity was assessed through evaluation of potential differences between participants, interventions and outcomes within each study. Where studies appeared to be clinically comparable, data were pooled in meta-analyses. Odds ratios using the fixed effects model were obtained for dichotomous data (rates of cognitive impairment) and Weighted Mean Differences (WMD) using the fixed effects model were obtained for continuous data (differences between groups in the psychometric test scores). In some cases, the mean difference was calculated as the difference between post-treatment and baseline performance. The standard deviation (SD) of this difference was calculated as the square root of the variance of the baseline plus the variance of the final assessment where the SD of the difference was not reported. For the large trials with long follow-up, the treatment effect and standard error were calculated. The WMD was used as the outcome measure where trials were measuring identical outcomes. The Standardized Mean Difference (SMD) with a random effects model was calculated where trials were not strictly comparable; either different tests within the same cognitive domain, different types of participants or different interventions.

Statistical heterogeneity was assessed by inspecting the scatter in the data points on the graphs and the overlap in their confidence intervals and, more formally, by checking the I^2 square quantity (Higgins 2003). This quantity describes the percentage of total variation across studies that is due to heterogeneity rather than chance. Interpretation of a given degree of heterogeneity will differ according to whether the estimates show the same direction of effect but it was planned to tentatively assign descriptions of low, moderate and high heterogeneity to I^2 values of 0 to 25%, 26 to 74% and 75 to 100% respectively. A priori, it was planned to look at the possible contribution of differences in trial design to any meta-analyses with an I^2 value of 50% or more.

The null hypothesis tested was that, for any of the above outcomes, treatment had no effect compared with placebo.

The following sensitivity analyses were planned to examine the stability of the results. Analyses were to be repeated excluding studies which differed in certain ways:

- (1) Estrogen therapy alone versus combined estrogen plus progestagen therapy
- (2) Surgically menopausal women versus women who had natural menopause

- (3) Early postmenopausal women (<60 years) versus older postmenopausal women (>60 years)
 - (4) Studies with adequate allocation concealment versus studies without adequate allocation concealment
 - (5) Studies with short follow up (<6 months) versus studies with longer follow up (>6 months)
- However, there were insufficient trials assessing the same types of cognitive tests, except in one case, to conduct these analyses.

RESULTS

Description of studies

Twenty four trials were identified that met the inclusion criteria for the review. Six earlier trials (Ditkoff 1991; Fedor-Freyberg 1977; Goebel 1995; Polo-Kantola 1998; Shaywitz 1999; Vanhulle 1976) had insufficient outcome data for analysis. We had tried to contact all the authors, but could not reach the above, save for Fedor-Freyberg and Polo-Kantola. However, neither of these could provide us with the necessary data for analysis. The authors of a further two trials (Dumas 2006; Schiff 2005) have been contacted for clarification of outcome data and are in the 'Studies awaiting assessment' section of the review. The reviewers included data from the tables of Almeida 2006; Binder 2001; Dunkin 2005; Duka 2000; Greenspan 2005; Hackman 1977; Linzmayer 2001; Phillips 1992; Resnick 2006; Shumaker 2003; Shumaker 2004; Wolf 1999; Yaffe 2006 and data from the graphs of Janowsky 2000; Sherwin 1988 and Sherwin 1990. For only three (Almeida 2006; Hackman 1977; Wolf 1999) the data of the mean difference and the SD of that difference were available. For all the others, these had to be calculated using the method described above or where there was long follow up and natural attrition, the treatment effect and standard error were calculated. The descriptive details that follows is a summary of only the 16 trials that had data suitable for analysis.

Study design

Fifteen of the 16 studies were of parallel group design. In the crossover study that was included (Sherwin 1988) we used data from only the first treatment period which is comparable to the parallel group design. Treatment lasted from two weeks (Wolf 1999) to 5.4 years (Shumaker 2004). Five studies had a duration of greater than one year.

Participants

Screening and selection

The total number of participants randomized in the trials ranged from 12 to 4532. Prior to 2002, most of the trials included fewer than 100 women. The three largest trials (Resnick 2006; Shumaker 2003; Shumaker 2004) included 4532, 2947 and 1416 women respectively. In total, 10,114 postmenopausal women have had assessments of their cognitive function in randomized trials. A wide variety in the type of subjects were included. While some studies had rigorous health screenings, others had few exclusion criteria. In addition, one study included only postmenopausal women with sleep difficulties (Linzmayer 2001), while others investigated more heterogeneous groups of women from the community (e.g. Almeida 2006; Dunkin 2005). Some studies included a mix of naturally and surgically menopausal women but others investigated relatively young women who were tested prior to and immediately after surgically induced menopause (e.g. Phillips 1992;

Sherwin 1988; Sherwin 1990). There was insufficient information in the studies to do split analyses on these aspects.

Age and control for confounding factors (education, mood)

One study included women of a wide age-range (Hackman 1977: 29 to 68 years) without controlling for age-effects in the analyses. Eight trials, including all of the large studies, only included older postmenopausal women (> 60 years in one study, > 65 years in four studies, > 70 years in one study, > 75 years in one study and 'several years after the menopause' in one study). Eleven studies (Almeida 2006; Binder 2001; Duka 2000; Dunkin 2005; Janowsky 2000; Phillips 1992; Resnick 2006; Shumaker 2003; Shumaker 2004; Wolf 1999; Yaffe 2006) had comparable age and education status between randomized groups at baseline. For four studies there was no information on education so separate analyses based on this variable could not be undertaken. Only a few studies ensured that mood was taken into account (Almeida 2006; Dunkin 2005; Hackman 1977; Phillips 1992; Resnick 2006; Wolf 1999; Yaffe 2006); they either demonstrated that groups were comparable in terms of rate of depression at baseline or they controlled for mood in their analyses.

Interventions

The mode, administration, dose and type of hormone therapy varied widely.

Three RCTs prescribed unopposed oral CEE to postmenopausal women. The other studies used transdermal (four), intramuscular (three) or oral (two) estradiol (E2) treatments. In addition, one RCT prescribed estrone (E1). Hence, most (nine) studies employed an estradiol (E2) preparation. All studies from the group led by Sherwin (Phillips 1992; Sherwin 1988; Sherwin 1990) gave a bolus injection of 10 mg of estradiol once monthly, which was shown to give very high blood levels of E2 (e.g. 1837 SD 11 pmol/L). In contrast, other studies used a transdermal patch containing E2 (Duka 2000; Dunkin 2005; Wolf 1999; Yaffe 2006) which gave much lower blood levels of E2. In addition, while all of the transdermal studies supplied the same amount of E2 in the transdermal patch, namely 0.1 mg/24 h, the average E2 levels reached in the women investigated by Duka 2000 were higher (average 220 pmol/L) than those of Wolf 1999 (average 124 pmol/L). One study (Almeida 2006) used a relatively high escalating dosage of unopposed estradiol in elderly women to minimize side effects. Two studies (Binder 2001; Greenspan 2005) added a progestagen (MPA) for women who had not undergone hysterectomy but no separate analyses were done for cases with versus without MPA. One study (Linzmayer 2001) compared the effects of estradiol alone with a new progestagen regimen (Dienogest). Two large studies assessed the effects of combined oral CEE + MPA on cognitive function in women with an intact uterus (Resnick 2006; Shumaker 2003).

Compliance checks were done using serum or plasma estrogen levels in 10 studies, five did a pill count and one study did not have a compliance check (Hackman 1977).

Outcomes

Cognitive assessments

Most of the studies assessed cognitive function using specific cognitive tests for different cognitive domains but few of the studies used similar cognitive tests, so making the pooling of data difficult. In some instances, different aspects of the same tests were

used (only immediate or delayed recall, or reaction time but not accuracy), or different versions. Analyses were split for tests of verbal (e.g. Paragraph Recall, Paired Associate test, Guild Memory test, Fluency) and visual (e.g. Visual Retention test) memory and for those which tested other aspects of cognitive function, such as speed of information processing (TMT-A/B, Stroop), mental rotations, accuracy of performance, manual dexterity and motor speed and executive function.

Five studies (Almeida 2006; Greenspan 2005; Shumaker 2003; Shumaker 2004; Yaffe 2006) used more global measures to assess cognitive function: either CAMCOG (Cambridge Cognitive Examination for Mental Disorders of the Elderly) (one study), Folstein Mini Mental State Examination (one study) or 3MSE (Modified Mini-Mental State Examination) all of which are highly correlated. Two of these studies (Shumaker 2003; Shumaker 2004) also compared the rates of "mild cognitive impairment" between groups after follow up of four or five years. Mild cognitive impairment was distinguished from "probable dementia" in the study and was defined according to a strict protocol with four phases of ascertainment. The diagnosis was based on the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) criteria and included poor performance (10th or lower percentile based on CERAD norms) on at least one CERAD test, a report of some functional impairment (but not sufficient to interfere with basic activities of daily living) from the designated informant, no evidence of a psychiatric disorder or medical condition that could account for the decline in cognitive function and an absence of adjudicated dementia.

Statistics

Some studies (e.g. Hackman 1977; Linzmayer 2001) made separate comparisons for treated subjects and the subjects using placebo, or made separate analyses for baseline and post-treatment effects (Linzmayer 2001; Sherwin 1990). Without adjustment, this strategy enhances the risk of chance findings due to multiple comparisons. For some of the larger studies, with long follow up, there was attrition of participants and so the change score could not be used because the number of participants at baseline differed from the number of participants at later follow-up. Where this occurred, the treatment effect at follow-up was calculated by the authors and the data entered using an inverse variance method, where the weight given to each study is chosen to be the inverse of the variance of the effect estimate. This method minimizes the uncertainty of the pooled effect estimate.

Risk of bias in included studies

In general, some methodological problems could be found in most of the earlier smaller studies (Hogervorst 2000; Yaffe 1998 a). The quality features of all of the included studies with available data are summarized in Additional Table 1, Quality of Included Studies (Table 2). Blinding is not included in this table as all trials were required to have at least double blinding before they could be considered eligible for inclusion.

Randomization method

Seven studies provided full details of their randomization procedure, two provided incomplete details and seven did not describe the randomization procedure.

Allocation concealment

Seven studies provided evidence of adequate allocation concealment and received a score of A. For the remaining trials, it was not clear whether there were processes in place to conceal the allocation to randomized groups; nine studies received an allocation concealment score of B.

Attrition

The proportion of dropouts varied widely. Some small trials were short term and had no or very few dropouts. Five trials (Almeida 2006; Binder 2001; Dunkin 2005; Phillips 1992; Sherwin 1988) had more than 10% of randomized participants not included in the analyses. Dunkin 2005 lost 33% and 36% in the ERT and placebo groups respectively after randomization.

Intention to treat analyses

Two trials (Almeida 2006; Greenspan 2005) had true intention to treat analysis and also provided completer's analyses of outcomes. The two largest studies (Shumaker 2003; Shumaker 2004) undertook true intention to treat analysis for the primary outcome of mild cognitive impairment but only modified intention to treat analyses for the scores on the 3MSE test (inclusion in the analyses of the participants in the groups to which they had been initially randomized but no imputation for participants who dropped out during follow up). Two studies (Resnick 2006; Yaffe 2006) claimed that they had performed intention to treat analyses but there was no correction made for those who dropped out during follow up. Another small trial (Duka 2000) did not provide any information in the publication about dropouts and intention to treat analysis is thus presumed. The remaining studies, n = 9, did not report intention to treat analysis.

Power calculation for sample size

The five largest studies (Greenspan 2005; Resnick 2006; Shumaker 2003; Shumaker 2004; Yaffe 2006) undertook power calculations for sample size. The remaining studies did not report whether this had been undertaken but given their size, it is unlikely.

Effects of interventions

Twenty four studies met inclusion criteria, but data of six RCTs (Ditkoff 1991; Fedor-Freyberg 1977; Goebel 1995; Polo-Kantola 1998; Shaywitz 1999; Vanhulle 1976) were not available for analysis and two other studies are waiting assessment. The following results are thus based on data from the remaining 16 studies. The findings of the studies without suitable data for analysis are also discussed, where relevant. All studies had been published in peer-reviewed journals.

Overall measures of cognitive function

Global cognition scores

There was no evidence of a difference in the scores on the global cognition test, CAMCOG after 20 weeks of E2 oral or placebo (Forest plot 1-1). However, the large WHIMS trial of combined estrogen plus progestagen therapy found significant differences in the 3MSE scores in favour of placebo at three and four years follow up (P value unadjusted) (WMD = -0.36, 95% CI -0.61 to -0.11; WMD = -0.52, 95% CI -0.81 to -0.23 respectively) (Forest plot 1-2). The other summary estimates at different follow up times did not differ significantly between treatment groups, although estimates were consistently

to the left of the line of no effect, favouring placebo. Two trials ([Shumaker 2004](#); [Yaffe 2006](#)) assessed 3MSE scores after one and two years of unopposed estrogen therapy and one trial assessed this outcome at longer follow up ([Greenspan 2005](#)). Women in the estrogen groups had significantly lower 3MSE scores at one year follow up (almost identical estimates in both trials) (Forest plot 1-3). There was no evidence of a significant difference between 3MSE scores at two, three, four, five or six years follow up, although estimates were consistently to the left of the line of no effect, favouring placebo.

Mild cognitive impairment

There was no evidence of a difference in the rates of mild cognitive impairment either after a mean of five years with unopposed oral estrogen therapy or after a mean of four years with combined estrogen plus progestagen therapy (Forest plot 2-1).

Measures of specific cognitive domains

Verbal memory and language tests

Only two studies, by the same group ([Phillips 1992](#); [Sherwin 1990](#)), had employed the same type of participants (surgically menopausal women), the same treatment (a 10 mg E2 bolus injection intramuscularly monthly for 2 to 3 months), the same tests and the same outcome measures. In these studies, the Paired Associate learning immediate recall was significantly in favour of treatment (SMD = 1.09, 95% CI 0.12 to 2.07) but no significant effects were found between groups in a small trial using transdermal E2 with participants with unknown type of menopause (Forest plot 3-1). The same test with a different scoring system for the immediate recall showed no evidence of an effect after nine months of CEE + MPA in older women (SMD = 0.3, 95% CI -0.27 to 0.88) but continued to show a statistically significant difference in the two smaller trials with surgically menopausal women (SMD = 1.02, 95% CI 0.19 to 1.85) (Forest plot 3-2). However, the delayed recall of the Paired Associate test did not reach significance after treatment with 10 mg E2 i.m./month for two months in relatively young surgically menopausal women, in the older women who were treated with CEE + MPA for nine months or in the women treated with E2 transdermal (Forest plot 3-3). The Paragraph recall (immediate and delayed recall) also showed no overall significant effect after 2 to 3 months of 10 mg E2 i.m./month in the relatively young surgically menopausal women (Forest plots 3-5 and 3-6). The study by [Wolf 1999](#) had used a different Paired Associates test (with only six instead of 10 word pairs) and the study by [Linzmayr 2001](#)(*) had employed a different test for "associative verbal memory". Results of treatment with E2 + P oral for two months or E2 transdermal for two weeks were non-significant on both tests (Forest plot 3-4).

None of the other verbal memory or language tests used in the other studies showed evidence of an effect of treatment (Forest plots 3-7, 3-8, 3-9, 3-10, 3-11, 3-12, 3-13, 3-14, 3-15). There was no evidence of an effect on semantic memory (verbal Fluency), word list recall tests, Wechsler Memory Scale tests, Boston Naming Test or PMA Vocabulary after two weeks E2 transdermal; or after nine months of CEE; or on short-term memory (Digit Span forward) after 2 to 3 months of 10 mg E2 i.m./month; or on episodic memory (Guild memory test) after six months of piperazine E1. The trials included in these comparisons used various doses and modes of administration of estrogen, except for one small trial that used CEE + MPA ([Binder 2001](#)), and different durations of follow up.

However, some significant differences were found between groups in a larger trial of combined estrogen and progestagen therapy (CEE + MPA) after 4 to 5 years of therapy ([Resnick 2006](#)) (Forest plot 3-16). In this study, women treated with HRT had significantly lower scores on a number of word list recall tests (CVLT total, CVLT short delay and CVLT long delay) (WMD = -0.52, 95% CI -0.091, to -0.13; WMD = -0.24, 95% CI -0.44 to -0.04; WMD = -0.23, 95% CI -0.43 to -0.03 respectively). There were no significant differences found between the scores on the other verbal memory or language tests assessed in this trial: PMA vocabulary or verbal fluency.

Visual memory tests

There was no evidence for an overall effect of treatment on different tests of visuospatial memory using SMD with a random-effects model (Forest plot 4-1). Only one of the studies employing a number of different regimens of estrogen (10 mg E2 i.m. monthly for 2 to 3 months, E2 transdermal for 2 to 10 weeks, E2 oral for 2 to 4 months, CEE oral for one month or low E transdermal for two years) showed evidence of a significant effect. In the [Almeida 2006](#) trial, women randomized to E2 oral had a significantly higher mean score for the visual memory for faces test than those randomized to placebo. Overall, there was no significant heterogeneity between studies ($I^2 = 13.8\%$) which suggests that it was reasonable to pool the studies. A number of other studies had employed visual memory tests ([Fedor-Freyberg 1977](#); [Phillips 1992](#); [Polo-Kantola 1998](#); [Vanhulle 1976](#)) but did not have data available for the meta-analyses. However, none of the studies, except [Fedor-Freyberg 1977](#), reported evidence for an effect of HRT on visual memory functions.

However, a significant but small difference on the Benton Visual Retention Test was found in the WHISCA trial comparing combined estrogen plus progestagen therapy (CEE + MPA) with placebo after 4 to 5 years follow up (Forest plot 4-2). A difference of -0.27 (95% CI -0.49 to -0.05) errors per year was reported between groups.

Speed tests, attention and manual dexterity

There was no evidence of an effect of nine months of CEE + MPA on simple speed of information processing (TMT-A, SMD = -0.42, 95% CI -1.0 to 0.15) (Forest plot 5-1). There was also no evidence of a significant difference in any of the other individual speed or attention tests or in the overall computed estimates for subgroups of tests with unopposed estrogen or combined therapy over variable follow up times (Forest plots 5-2, 5-3, 5-4, 5-5). In addition, most of the other studies with data not available for analysis also did not report any evidence for differences between HRT and placebo on measures of speed of information processing ([Ditkoff 1991](#); [Goebel 1995](#); [Polo-Kantola 1998](#) but see [Fedor-Freyberg 1977](#)).

Mental rotations and accuracy

There was no significant effect on mental rotation test aspects (time to complete or number correct) after 2 to 3 weeks of E2 transdermal or after one year of CEE + MPA (Forest plots 7-1, 7-2, 7-3).

There was also no significant effect on accuracy as measured with Letter Cancellation after nine months of CEE + MPA or as measured with random number generation after three weeks of E2 transdermal (Forest plots 8-2, 8-3). However, again a highly significant effect was seen on the Clerical Speed and Accuracy test after three months of 10 mg E2 i.m. monthly in relatively young

surgically menopausal women (WMD = 6.0, 95% CI 4.72 to 7.28) (Forest plot 8-1).

Executive function

The cognitive tests that assessed executive function included Wisconsin Card Sorting Test (categories, perseverative responses, % conceptual level responses), Wechsler Adult Intelligence Scale (picture arrangement), Tower of London, abstract reasoning, Trails B and the Stroop Interference Time and Errors tests. There was a significantly better performance after three months of 10 mg E2 i.m. monthly in relatively young surgically menopausal women on an abstract reasoning test (WMD = 6.80, 95% CI 5.52 to 8.08) (Forest plot 9-6). There was no evidence of a significant effect of any of the regimens (3 to 10 weeks or 1 to 2 years of E2 transdermal) on any of the other tests (Forest plots 9-1, 9-2, 9-3, 9-4, 9-5, 9-7, 9-8, 9-9, 9-10).

*For the [Linzmayr 2001](#) study, we compared only the E2+P versus placebo regimen only (see also the visual memory tests) as the E2 treatment was never shown to be significant compared with placebo.

DISCUSSION

Previous animal studies, epidemiologic observational studies and clinical trial data have suggested that estrogen therapy could protect against cognitive aging, particularly verbal memory, in healthy postmenopausal women ([Sherwin 2003](#)). There is good evidence from the studies in this review that cognitive abilities in older postmenopausal women (> 65 years), as measured by global cognition tests, are not improved or maintained, and may even be adversely affected by hormone therapy, either unopposed estrogen or combined estrogen plus progestagen. In particular, significant declines in scores on the 3MSE test, a measure of global cognition function, were found for older postmenopausal women on combined CEE + MPA after three and four years and after one year of unopposed estrogen (4532 and 2947 women respectively). Rates of mild cognitive impairment were also not reduced when women are treated with hormone therapy. The two WHIMS trials, one with ERT and the other with HRT, both found similar effects on 3MSE scores, although women in the two trials differed according to many factors at baseline, including cognitive function. The effects also appeared to be consistent across numerous subgroups. However, baseline 3MSE score was associated with differential treatment effects. In both the ERT and HRT WHIMS trials, among women whose baseline scores exceeded 95 at baseline, the reduction in scores over the follow up of the trials did not differ statistically from zero. Women with relatively low baseline cognitive function (as defined by scores below the cutpoints that were defined according to education) had larger decreases in their 3MSE scores over time. These findings strongly suggest that hormone therapy, either ERT or HRT, should not be initiated in older postmenopausal women for the purpose of protecting cognitive function.

The 3MSE scale is a reliable and valid measure of global cognitive function but may be too coarse an instrument to adequately detect benefits or harms in specific cognitive domains, such as verbal memory. There was large variation in the other included studies in this review in the individual tests used, the design and size of the included studies, the participants (menopausal symptoms, age, type of menopause) and/or the treatment (type of hormone regimen, mode of administration, dose and duration). The question

remains whether it is possible to tease out more specific effects of various types of hormone therapy on individual cognitive domains in specific subgroups of postmenopausal women.

Results from the trials assessing effects on individual cognitive domains showed no clear pattern. In summary, the meta-analyses showed a significant beneficial effect of 10 mg E2 i.m./month for 2 to 3 months in relatively young surgically menopausal women on only one aspect of a test of verbal memory (Paired Associate learning immediate recall) (two trials with a total of 31 women), on a test of abstract reasoning (one trial, 20 women) and a test of speed and accuracy (one trial, 20 women). A significant benefit was found for women on E2 oral for 20 weeks in their scores on a specific visual memory test, the Visual Memory for Faces test, but there was no difference overall when all visuospatial tests were combined. Treatment with CEE + MPA only showed a beneficial effect on one test of complex information processing (TMT-B) after nine months, which was not reported by the authors (one trial, 52 women) but this was not found in two other studies assessing this test. A deleterious effect after 4 to 5 years of combined CEE plus MPA in older postmenopausal women was found on a number of verbal memory tests (CVLT total and CVLT short and long delay). By contrast, a small beneficial effect after the same regimen was found for a test of short-term figural memory and visuoconstructional abilities (1416 women). Significant declines in scores on the 3MSE test, a measure of global cognition function, were also found for older postmenopausal women on combined CEE + MPA after three and four years and after one year of unopposed estrogen (4532 and 2947 women respectively). Many studies did not find any evidence of effects of treatment on verbal and visual memory tests, simple and complex speed of information processing, accuracy and executive function, but most of these studies had fewer than 100 participants and were likely to be underpowered.

Given that this review has found no benefit of either ERT or HRT for cognitive functioning overall, it has not been possible to determine whether the significant results found in some tests of individual cognitive domains are due to bias from methodological shortcomings in the studies or to other factors such as the type of test used, type of participants, type of intervention or duration of follow up.

Type of individual tests

In general, verbal memory has been considered to be most susceptible to the effects of estrogens ([Phillips 1992](#); [Sherwin 1990](#)). However, there was only one memory subtest, the immediate Paired Associates test (but not the delayed recall) of the Wechsler Memory Scale, which showed a significant beneficial effect of estrogen therapy. By contrast, a larger study found a deleterious effect of HRT on tests of verbal memory, the California Verbal Learning Test (total, short and long delay). This finding was not reported for other smaller trials with a shorter duration of therapy. There was no evidence of a significant difference on any of the other verbal memory or language tests. The verbal memory tests that were not affected by HRT may have been conceptually different from the Paragraph and Paired Associates test or the CVLT tests and were probably slightly more difficult.

No significant differences were reported between groups for any of the visual memory tests, except for a small benefit for women randomized to CEE + MPA on the Benton Visual Retention Test (BVRT), a test of short-term figural memory and visuoconstructional abilities. It is not clear whether this represents a clinically important

benefit. Other smaller trials assessing this test did not find a significant difference between groups.

Of the speeded tests, there was no evidence of a difference between treatment and control groups except for one small study. [Sherwin 1988](#) reported significant beneficial effects for women on estrogen therapy compared to placebo on a test of clerical speed and accuracy.

There was no evidence of a significant difference in any of the tests measuring executive function, except for a significant benefit for one small study measuring abstract reasoning ([Sherwin 1988](#)).

Subjects-screening and selection

There was a wide variation in the type of participants included. One of the studies that investigated asymptomatic women ([Ditkoff 1991](#)) did not report any effects of treatment and others ([Resnick 2006](#); [Shumaker 2003](#); [Shumaker 2004](#)) found deleterious effects of hormone treatment. This raises the issue of whether the reduction in menopausal or climacteric symptoms is responsible for the effects of estrogen on cognitive function (see [LeBlanc 2001](#)). [Ditkoff 1991](#) used two tests which were not found to be sensitive to the effects of estrogen overall (Digit Span and DSST, a speed test). In addition, studies that directly tested this hypothesis (e.g. [Phillips 1992](#); [Wolf 1999](#)) did not find that cognitive effects were mediated through an alleviation in depression or improved sleep. However, [Phillips 1992](#) did report a significant correlation between the reduction in hot flushes and performance on the immediate recall of the Paired Associate learning test, the only verbal memory test which showed a significant overall effect of treatment, which would be in favour of the 'symptom reduction causes cognition improvement' hypothesis. In line with this, [Fedor-Freyberg 1977](#) who reported a large improvement on several measures of speed of performance, had included only women who were highly symptomatic

Age, education and type of menopause of participants

Some studies included women of a wide age-range (e.g. [Hackman 1977](#): 29 to 68 years; [Shaywitz 1999](#): 33 to 61 years). The age differences between and within studies make comparison difficult as many of the tests are sensitive to the effects of aging. In fact, both studies did not find any effects which may have been caused by the large age-range, obscuring smaller effects of HRT.

It has not been possible in this review to answer the recent suggestion that there may be a critical window in the early postmenopause where treatment with estrogen therapy may protect against cognitive decline ([Henderson 2004](#); [Sherwin 2003](#)). The studies by Sherwin's group ([Phillips 1992](#); [Sherwin 1988](#)) all included relatively young (on average 47 years of age) women who had just undergone surgical menopause, suggesting that there may be only a limited window of time in which effects of HRT on cognitive function are apparent. Compatible with this interpretation is the observation that RCTs investigating much older women ([Binder 2001](#); [Goebel 1995](#)) found no evidence of benefit from treatment. However, studies including relatively young women ([Shaywitz 1999](#): 51 years of age, [Polo-Kantola 1998](#): 56 years of age) also found no evidence of positive effects on cognitive function. Possibly, no effects were found because these groups had all included a mixed group of natural and surgical menopausal women. According to some studies, a surgical

menopause at a younger age accelerates cognitive aging ([Nappi 1999](#)) and estrogen could perhaps prevent this.

Epidemiological studies also suggest that the effect of estrogen may be stronger in women with less education ([Hogervorst 2000](#)). However, education of participants was not always given and we could not control for differential effects of treatment on participants with a low education. In general, age, type of menopause, education and depression have rarely been controlled for in the analyses.

Treatment

The studies that found positive effects on the immediate Paired Associates, the abstract reasoning and accuracy test were all from one group. Their subjects were given bolus injections of E2 i.m. per month which gave very high plasma levels of E2 (> 1800 pmol/L [Phillips 1992](#)). In comparison, [Wolf 1999](#) had given their subjects a transdermal patch of E2 that gave an overall low plasma E2 level (one tenth of those reached in the Sherwin studies employing a bolus injection). Interestingly, Wolf et al.'s post hoc analyses showed that the women who were treated and had higher plasma E2 levels (> 106 pmol/L) had a better immediate and delayed Paired Associates performance than the women who were also treated but had lower E2 levels. [Polo-Kantola 1998](#) also used a transdermal E2 patch (but with varying dosages) and no effect of treatment or the level of E2 was seen on cognitive functions. [Duka 2000](#) used a transdermal patch of E2 but their effects seemed to be largely explained by baseline differences. [Yaffe 2006](#) used a very low dose of transdermal estrogen (delivering 0.014 mg of estradiol per day) and reported no change in cognitive function. Thus, possibly the type of administration (transdermal patch versus intramuscular injections) and the E2 level reached determined the effect size.

In addition, the type of estrogen could be important. Inspection of the table of included studies show that those which used CEE, the most widely prescribed estrogen, either did not show any positive result as reported by the authors (see effect of TMT-B) or suggested potential harm (in global cognitive scores). The study by [Hackman 1977](#) employed an estrone (E1) treatment and there was only a weak trend in the Guild Memory test in favour of treatment. E1 has less receptor affinity than E2 and could thus be expected to be less biologically active. In a very early study ([Vanhulle 1976](#)) estriol (E3) was employed which has an even lower receptor affinity than E1. In this study, only very few and small effects were reported on vigilance and concentration.

Longitudinal studies that have assessed estrogen levels are in line with this theory. [Yaffe 2000](#) found that bioavailable E2 was associated with a slower decrement in the MMSE performance over time. However, [Yaffe 1998 b](#) also showed that high levels of E1 had detrimental effects on some speed measures over time. No distinction was made here between natural and surgical menopause, and age at menopause was not taken into account in the analyses. However, consistent with the discussion on different regimens, it is possible that different types of estrogens have different or even counteractive effects. This hypothesis remains to be tested; the review contains insufficient evidence to determine whether different types of estrogen have differential effects on cognition.

Another important questions is whether adding a progestagen is detrimental. This has often been postulated, since animal studies have shown that progestagens can counteract several of the actions

of estrogens (on monoaminergic neurotransmitter systems and possibly on the vascular system (Hogervorst 2000)). A progestagen is added in HRT for women who have not been hysterectomized to reduce the risk of endometrial hyperplasia. Five studies (Binder 2001; Goebel 1995; Greenspan 2005; Linzmayer 2001; Resnick 2006) used a combined regimen (for two of these, the data was not analysed separately from women having unopposed estrogen). Resnick 2006 reported that a combined regimen was associated with worsening scores for verbal memory tests and a small benefit for a test of figural memory when compared with placebo. The other trials did not find any effects of HRT on cognitive function but all of these, except for one, had fewer than 100 participants and are likely to be underpowered. It was not possible to determine in these trials whether age of the participants had an effect on the results. All but one of the trials were conducted with older women. Duration of therapy could explain the differences between the trials; the Resnick 2006 trial investigated effects after four or five years of therapy. There was insufficient evidence to determine whether different types of progestagen have differential effects on cognition.

Heterogeneity and sensitivity analyses

There was no significant statistical heterogeneity ($I^2 > 50\%$) in any of the analyses. There were insufficient trials to undertake the planned sensitivity analyses, except in two cases.

(1) The results of three trials were pooled to assess effects of ERT/HRT on paired associate learning (immediate recall, with different scoring methods) (comparison 1/2). When the two trials with inadequate allocation concealment were excluded from the meta-analysis, the effects of hormone treatment were no longer significantly different from placebo.

(2) It is not clear whether some of the participants in two of the included studies in this review are identical, although they have been analysed separately (Sherwin 1990; Phillips 1992). The authors are the same in both studies and the intervention and outcomes are identical. When the smaller of these studies is excluded in a sensitivity analysis (Sherwin 1990) (Forest plots 3-1 and 3-2), there is no longer a significant effect of ERT on the Paired Associate Learning test results (immediate recall).

This calls into question the suggestion that ERT may have a differential effect on a particular domain of cognitive function in specific types of participants.

Summary

It is clear that ERT or HRT does not protect against cognitive aging in older postmenopausal women, as represented by their scores on global cognition tests. The 3MSE test has good sensitivity and specificity for the detection of mild cognitive impairment (Bland 2001) and thus provides an overall indication of whether hormones protect against cognitive decline. Moreover, the findings in the same trials that both ERT or HRT increase the risk of probable dementia and do not prevent mild cognitive impairment supports the conclusion that older postmenopausal women should not use hormones to attenuate cognitive decline as they age. We have not been able to determine, because of insufficient randomized data, whether any type of hormone regimen has more specific effects on individual domains of cognition or whether different subgroups of women are more susceptible to the harmful or beneficial effects of

hormones. There is an indication that HRT causes decrements in some verbal memory tests and a small but significant benefit for a particular figural memory test when given for an average of four years. In addition, it has been considered possible that the effect of unopposed estrogen is very specific and that there may be a critical period for benefit (Henderson 2004; Maki 2005; Maki 2006; Sherwin 2003), notably symptomatic and recently menopausal women. This review suggests that surgically menopausal women of a relatively young age who are given bolus injections of E2 could benefit. However, the trials showing the benefit in this subgroup had low methodological quality scores and the findings could be due to chance (type 1 error).

Some of the unanswered questions from this review should be answered within the next few years. Results from the ERT arm of the WHISCA trial should be reported within the next year. Currently, two large prospective randomized placebo-controlled studies are in progress in the USA (PREPARE and KRONOS). The PREPARE trial (Preventing Postmenopausal Memory Loss and Alzheimers with Replacement Estrogens) no longer has an active treatment arm but is still following up participants. The KRONOS trial is assessing the effects of combined treatment on cognitive endpoints as secondary outcomes in women randomized within three years of the menopause over four years and is expected to report results in 2010. Results from the ERT arm of the WHISCA trial are currently being analysed. The WISDOM-COG trial (Women's International Study of long Duration estrogen after Menopause - Cognition) was prematurely terminated in 2002 after WHI results were released. By 2010, some of the unanswered questions about the effects of hormones on cognition may have been clarified. In the meantime, older postmenopausal women should not initiate hormone therapy to prevent cognitive decline.

AUTHORS' CONCLUSIONS

Implications for practice

There is good evidence that both ERT and HRT do not prevent cognitive decline in older postmenopausal women. It is not known whether specific types of ERT or HRT have effects in subgroups of women, although there was evidence that combined hormone therapy (CEE + MPA) in similarly aged women was associated with a decrement in a number of verbal memory tests and a small improvement in a test of figural memory. There is insufficient evidence to determine whether subgroups of women using specific types of hormone therapy could benefit from treatment. Based on the available evidence, ERT or HRT cannot be recommended for cognitive improvement or maintenance in healthy postmenopausal women.

Implications for research

We need to await the results of the ongoing large RCTs to answer the questions raised in the discussion, which included the potential modifying effect on estrogen effects by factors such as age, surgical or natural menopause and menopausal symptoms in maintaining the cognitive functions of healthy postmenopausal women.

ACKNOWLEDGEMENTS

Marc Budge of OPTIMA.

Lindsay Stead, consumer editor.

REFERENCES

References to studies included in this review

Almeida 2006 {published data only}

Almeida OP, Lautenschlager NT, Vasikaran S, Leedman P, Gelavis A, Flicker L. A 20-week randomized controlled trial of estradiol replacement therapy for women aged 70 years and older: effect on mood, cognition and quality of life. *Neurobiology of Aging* 2006;**27**:141-9.

Binder 2001 {published data only}

Binder EF, Schechtman KB, Birge SJ, Williams DB, Kohrt WM. Effects of HRT on cognitive performance in elderly women. *Maturitas* 2001;**38**(2):137-46.

Ditkoff 1991 {unpublished data only}

Ditkoff EC, Crary, WG, Cristo M, Lobo RA. Estrogen improves psychological function in asymptomatic postmenopausal women. *Obstetrics and Gynecology* 1991;**78**:991-5.

Duka 2000 {published data only}

Duka T, Tasker R, McGowan JF. The effects of 3-week estrogen hormone replacement on cognition in elderly healthy females. *Psychopharmacology* 2000;**149**:129-39.

Dunkin 2005 {published data only}

Dunkin J, Rasgon N, Wagner Steh K, David S, Altshuler L, Rapkin A. Reproductive events modify the effects of estrogen replacement therapy on cognition in healthy postmenopausal women. *Psychoneuroendocrinology* 2005;**30**(3):284-96.

Fedor-Freyberg 1977 {published data only}

Fedor-Freyberg P. The influence of estrogens on the well-being and mental performance on the climacteric and postmenopausal women. *Acta Obstetrics Gynecology Scandinavia* 1977;**64**:12-20.

Goebel 1995 {published data only}

Goebel JA, Birge SJ, Price SC, Hanson JM, Fishel DG. Estrogen replacement therapy and postural stability in the elderly. *American Journal of Otolaryngology* 1995;**16**:470-4.

Greenspan 2005 {published data only}

Greenspan SL, Resnick NM, Parker RA. The effect of hormone replacement on physical performance in community-dwelling elderly women. *American Journal of Medicine* 2005;**118**:1232-9.

Hackman 1977 {published data only}

Hackman BW, Galbraith D. Six month study of oestrogen therapy with piperazine oestrone sulphate ('Harmogen') and its effects on memory. *Current Medical Research Opinion* 1977;**4**(3):21-7.

Janowsky 2000 {published data only}

Janowsky JS, Chavez B. Sex steroids modify working memory. *Journal of Cognitive Neuroscience* 2000;**12**:407-14.

Linzmayr 2001 {published data only}

Linzmayr L, Semlitsch HV, Saletu B, Saletu-Zyhlars G, Zoghalmi A, Gruber D, et al. Double-blind, placebo-controlled

psychometric studies of a combined estrogen-progestin regimen versus estrogen alone on performance, mood and personality of menopausal syndrome patients. *Arzneimittel Forschung/Drug Research* 2001;**51**(I):238-45.

Phillips 1992 {published data only}

Phillips SM, Sherwin BB. Effects of estrogen on memory function in surgically menopausal women. *Psychoneuroendocrinology* 1992;**17**:485-95.

Polo-Kantola 1998 {published data only}

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, March):459-66.

Resnick 2006 {published data only}

Resnic SM, Maki PM, Rapp SR, Espeland MA, Brunner R, Coker LH, et al. for the Women's Health Initiative Study of Cognitive Aging Investigators. Effects of combination estrogen plus progestin hormone treatment on cognition and affect. *Journal of Clinical Endocrinology and Metabolism* 2006;**91**(5):1802-10.

Shaywitz 1999 {published data only}

Shaywitz SE, Shaywitz BE, Pugh KR, Fulbright RK, Skudlarski P, Mencl WE, et al. Effects of estrogen on brain activation patterns in postmenopausal women during working memory tasks. *JAMA* 1999;**281**(13):1197-202.

Sherwin 1988 {published data only}

Sherwin BB. Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. *Psychoneuroendocrinology* 1988;**13**:345-57.

Sherwin 1990 {published data only}

Sherwin BB, Phillips S. Estrogen and cognitive functioning in surgically menopausal women. *Annals of the New York Academy of Sciences* 1990;**592**:474-5.

Shumaker 2003 {published data only}

Rapp SR, Espeland MA, Shumaker SA, Henderson VW, Brunner RL, Manson JE, et al. for the WHIMS Investigators. Effect of estrogen plus progestin on global cognitive function in postmenopausal women. *JAMA* 2003;**289**:2663-72.

* Shumaker SA, Legault C, Thal L, Wallace RB, Ockene JK, Hendrix SL, et al. Estrogen plus progestin and the incidence of dementia and mild cognitive impairment in postmenopausal women: the Women's Health Initiative Memory Study: a randomized controlled trial. *JAMA* 2003;**289**(20):2651-62.

Shumaker 2004 {published data only}

Espeland MA, Rapp SR, Shumaker SA, Brunner R, Manson JE, Sherwin BB, et al. for the WHIMS Investigators. Conjugated equine estrogens and global cognitive function in postmenopausal women. *JAMA* 2004;**291**(24):2959-68.

* Shumaker SA, Legault C, Kuller L, Rapp SR, Thal L, Lane DS, et al. Women's Health Initiative Memory Study. Conjugated equine estrogens and incidence of probable dementia and mild cognitive impairment in postmenopausal women: Women's Health Initiative Memory Study. *JAMA* 2004;**291**(24):2947-58.

Vanhulle 1976 {published data only}

Vanhulle G, Demol R. A double-blind study into the influences of estriol on a number of psychological tests in postmenopausal women. In: van Keep PA, Greenblatt RB, Albeaux-Fernet M editor(s). Consensus on Menopausal Research. London: MTP Press, 1976:94-9.

Wolf 1999 {published data only}

Wolf OT, Kudielka BM, Hellhammer DH, Toerber S, McEwen BS, Kirschbaum C. Two weeks of transdermal estradiol treatment in postmenopausal elderly women and its effect on memory and mood. *Psychoneuroendocrinology* 1999;**24**:727-41.

Yaffe 2006 {published data only}

Yaffe K, Vittinghoff E, Ensrud KE, Johnson KC, Diem S, Hanes V, et al. Effects of ultra-low-dose transdermal estradiol on cognition and health-related quality of life. *Archives of Neurology* 2006;**63**:945-50.

References to studies excluded from this review

Drapier 2005 {published data only}

Drapier Faure E, Azoulay C, Abramovici Y. Acceptabilite par les patientes d'un THS combine continu versus cyclique: une etude clinique randomisee multicentrique francaise. [Acceptability of continuous combined versus cyclical HRT: a French multicentric randomized clinical study]. *Gynecologie, obstetrique and fertilité* 2005;**33**(7-8):498-504.

Friebely 2001 {published data only}

Friebely JS, Shifren JL, Schiff I, Regestein QR. Preliminary observations on differing psychological effects of conjugated and esterified estrogen treatments. *Journal of Womens Health and Gender Based Medicine* 2001;**10**(2):181-7.

Grady 2002 {published data only}

Grady D, Yaffe K, Kristof M, Lin F, Richards C, Barrett-Connor E. Effect of postmenopausal hormone therapy on cognitive function: the Heart and Estrogen/progestin Replacement Study. *American Journal of Medicine* 2002;**113**:543-8.

Greendale 1997 {published data only}

Greendale GA, Melton B, Postmenopausal Estrogen/Progestin Interventions (PEPI) Investigators. Effects of estrogen or estrogen/progestins on physical, cognitive, and affective symptoms. *Journal of the American Geriatrics Society* 1997;**45**:S25.

Hogervorst 1999 {published data only}

Hogervorst E, Boshuisen M, Riedel WJ, Willekes C, Jolles J. The effect of hormone replacement therapy on cognitive function in elderly women. *Psychoneuroendocrinology* 1999;**24**:43-68.

Krug 2003 {published data only}

Krug R, Molle M, Dodt C, Fehm HL, Born J. Acute influences of estrogen and testosterone on divergent and convergent thinking in postmenopausal women. *Neuropsychopharmacology official publication of the American College of Neuropsychopharmacology* 2003;**28**(8):1538-45.

Maki 2004 {published data only}

Maki PM, Resnick SM, Brandt J, Dobs AR, Durso SC, McCrae RR, University of Illinois at Chicago, Chicago. Changes in self-reported personality coincident with declines in cognition: results from parallel hormone therapy trials in elderly men and women. *Neurobiology of Aging* 2004;**25**(S2):106.

McDonald 1952 {published data only}

McDonald Caldwell B, Watson RI. An evaluation of psychologic effects of sex hormone administration in aged women. *Journal of Gerontology* 1952;**7**(2):228-44.

Michael 1970 {published data only}

Michael CM, Kantor HI, Shore H. Further psychometric evaluation of older women-the effect of estrogen administration. *Journal of Gerontology* 1970;**25**(4):337-41.

Natale 2001 {published data only}

Natale V, Albertazzi P, Zini M, Di Micco R. Exploration of cyclical changes in memory and mood in postmenopausal women taking sequential combined oestrogen and progestogen preparations. *British Journal of Obstetrics and Gynaecology* 2001;**108**:286-90.

Nielsen 2006 {published data only}

Nielsen TF, Ravn P, Pitkin J, Christiansen C. Pulsed estrogen therapy improves postmenopausal quality of life: a 2-year placebo-controlled study. *Maturitas* 2006;**53**(2):184-90.

Ohkura 1998 {published data only}

* Ohkura T. Sex steroids and brain function. *Acta Obstetrica et Gynaecologica Japonica* 1998;**50**(8):633-44.

Pan 2003 {published data only}

Pan HA, Wang ST, Pai MC, Chen CH, Wu MH, Huang KE. Cognitive function variations in postmenopausal women treated with continuous, combined hrt or tibolone: a comparison. *Journal of Reproductive Medicine for the Obstetrician and Gynecologist* 2003;**48**(5):375-80.

Pauker 1958 {published data only}

Pauker JD, Kheim T, Mensh I, Kountz WB. Sex hormone replacement in the aged: Psychological and medical evaluation of administration of androgen-estrogen and of androgen-estrogen combined with reserpine. *Journal of Gerontology* 1958;**12**:389-97.

Rasgon 2005 {unpublished data only}

Rasgon N. Estrogen use in protection from cognitive decline. ClinicalTrials.gov 2005a.

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.

Ryan 2001 {published data only}

Ryan N, Rosner A. Quality of life and costs associated with micronized progesterone and medroxyprogesterone acetate in hormone replacement therapy for nonhysterectomized, postmenopausal women. *Clinical Therapeutics* 2001;**23**(7):1099-115.

Stevens 2005 {published data only}

Stevens MC, Clark VP, Prestwood KM. Low-dose estradiol alters brain activity. *Psychiatry Research* 2005;**139**(3):199-217.

Tierney 2005 {published data only}

Tierney MC. Estrogen trial to delay the onset of Alzheimer's disease. <http://www.controlled-trials.com> (accessed 2005).

Viscoli 2005 {published data only}

Viscoli CM, Brass LM, Kernan WN, Sarrel PM, Suissa S, Horwitz RJ. Estrogen therapy and risk of cognitive decline: results from the Women's Estrogen for Stroke Trial (WEST). *American Journal of Obstetrics and Gynecology* 2005;**192**:387-93.

Wisniewski 2002 {published data only}

Wisniewski AB, Nguyen TT, Dobs AS. Evaluation of high-dose estrogen and high-dose estrogen plus methyltestosterone treatment on cognitive task performance in postmenopausal women. *Hormone Research* 2002;**58**(3):150-5.

References to studies awaiting assessment
Dumas 2006 {published data only}

Dumas J, Hancur-Bucci C, Naylor M, Sites C, Newhouse P. Estrogen treatment effects on anticholinergic-induced cognitive dysfunction in normal postmenopausal women. *Neuropsychopharmacology* 2006;**31**:2065-78.

Joffe 2006 {published data only}

Joffe H, Hall JE, Gruber S, Sarmiento IA, Cohen LS, Yurgelun-Todd D, Martin KA. Estrogen therapy selectively enhances prefrontal cognitive processes: a randomized double-blind placebo-controlled study with functional magnetic resonance imaging in perimenopausal and recently postmenopausal women. *Menopause* 2006;**13**(3):411-22.

LeBlanc 2007 {published data only}

LeBlanc ES, Neiss MB, Carello PE, Samuels MH, Janowsky JS. Hot flashes and estrogen therapy do not influence cognition in early menopausal women. *Menopause* 2007;**14**(2):191-202.

Schiff 2005 {published data only}

Schiff R, Bulpitt CJ, Wesnes KA, Rajkumar C. Short-term transdermal estradiol therapy, cognition and depressive symptoms in healthy older women. A randomised placebo controlled pilot cross-over study. *Psychoneuroendocrinology* 2005;**30**(4):309-15.

References to ongoing studies
Kronos 2006 {unpublished data only}

Crenshaw P. Kronos Early Estrogen Prevention Study (KEEPS). Metaregister of Current Controlled Clinical Trials.

Ragnnauth 2000 {published data only}

Ragnnauth A. Dementia prevention study for healthy women. Centerwatch 2000.

Additional references
Barrett-Connor 1991

Barrett-Connor E. Postmenopausal estrogen and prevention bias. *Annals of Internal Medicine* 1991;**115**:455-6.

Bland 2001

Bland RC, Newman SC. Mild dementia or cognitive impairment: the Modified Mini-Mental State examination (3MS) as a screen for dementia. *Canadian Journal of Psychiatry* 2001;**46**:506-10.

Farquhar 2005

Farquhar CM, Marjoribanks J, Lethaby A, Lamberts Q, Suckling JA and the Cochrane Hormone Therapy Study Group. Long term hormone therapy for perimenopausal and postmenopausal women. *Cochrane Database of Systematic Reviews* 2005, Issue 3.

Fratiglioni 2000

Fratiglioni L, Launer LJ, Andersen K, Breteler MM, Copeland JR, Dartigues JF, et al. Incidence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. Incidence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group.. *Neurology* 2000;**54**(11):S10-15.

Henderson 2004

Henderson VW. Hormone therapy and Alzheimer's disease: benefit or harm?. *Expert Opinion on Pharmacotherapy* 2004;**5**(2):389-406.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60.

Hogervorst 2000

Hogervorst E, Williams J, Budge M, Riedel W, Jolles J. The nature of the effect of female gonadal hormone replacement therapy on cognitive function in post-menopausal women: a meta-analysis. *Neuroscience* 2000;**101**(3):485-512.

Hogervorst 2002 b

Hogervorst E, Yaffe K, Richards M, Huppert F. Hormone replacement therapy to maintain cognitive function in women with dementia. *Cochrane Database of Systematic Reviews* 2002, Issue 3.

Huppert 1997

Huppert F, Wilcock G. Ageing, cognition and dementia. *Age and Ageing* 1997;**Dec 26**(Suppl 4):20-3.

Launer 1999

Launer LJ, Andersen K, Dewy ME, Letenneur L, Ott A, Amaducci LA, et al. Rates and risk factors for dementia and Alzheimer's disease. *Neurology* 1999;**1**:78-84.

LeBlanc 2001

LeBlanc ES, Janowsky J, Chan BKS, Nelson HD. Hormone replacement therapy and cognition. *JAMA* 2001;**285**(11):1489-99.

Maki 2005

Maki PM. A systematic review of clinical trials of hormone therapy on cognitive function. *Annals of the New York Academy of Sciences* 2005;**1052**:182-7.

Maki 2006

Maki PM. Hormone therapy and cognitive function: is there a critical period for benefit?. *Neuroscience* 2006;**138**:1027-30.

Matthews 1996

Matthews KA, Kuller LH, Wing RR, Meilahn EN, Plantinga P. Prior to use of estrogen replacement therapy: are users healthier than nonusers?. *American Journal of Epidemiology* 1996;**143**(10):971-8.

McEwen 1997

McEwen BS, Alves SE, Bulloch K, Weiland NG. Ovarian steroids and the brain implications for cognition and aging. *Neurology* 1997;**48**:S14-18.

Nappi 1999

Nappi RE, Sinforiani E, Mauri M, Bono G, Polatti F, Nappi G. Memory functioning at menopause: impact of age in

ovariectomized women. *Gynecologic and Obstetric Investigation* 1999;**47**:29-36.

Sherwin 2003

Sherwin BB. Estrogen and cognitive functioning in women. *Endocrine Reviews* 2003;**24**:133-51.

Silva 2001

Silva I, Mor G, Naftolin F. Estrogen and the aging brain. *Maturitas* 2001;**Feb 28;38**(1):100-1.

Yaffe 1998 a

Yaffe K, Sawaya G, Lieberburg I, Grady D. Estrogen therapy in postmenopausal women. *JAMA* 1998;**279**:688-95.

Yaffe 1998 b

Yaffe K, Grady D, Pressman A, Cummings S. Serum estrogen levels, cognitive performance and risk of cognitive decline in older community women. *JAGS* 1998;**46**:816-21.

Yaffe 2000

Yaffe K, Lui L-Y, Grady D, Cauley J, Kramer J, Cummings SR. Cognitive decline in women with relation to non-protein bound estradiol concentrations. *Lancet* 2000;**356**(9231):708-12.

References to other published versions of this review

Hogervorst 2002a

Hogervorst E, Yaffe K, Richards M, Huppert F. Hormone replacement therapy for cognitive function in postmenopausal women. *Cochrane Database of Systematic Reviews* 2002, Issue 2.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Almeida 2006

Methods	Design: randomized, double-blind, parallel groups, 20 weeks
Participants	Country: Australia, n = 86 healthy women over 70, mean age = 73.8, natural or surgical menopause not stated
Interventions	1. E2 oral (0.5 mg weeks 1 to 2; 1 mg weeks 3 to 4; 2 mg weeks 5 to 16; 1 mg weeks 17 to 18; 0.5 mg weeks 19 to 20) 2. Placebo
Outcomes	-General: CAMCOG -Verbal memory/language: CVLT (IR+DR), Verbal Fluency, Faces (IR+DR) -Visuospatial: Block Design
Notes	HRT > placebo 1/5 tests (Faces - immediate recall, not delayed recall). No effect: CAMCOG, Verbal Fluency, CVLT, Block Design

Binder 2001

Methods	Design: randomized double-blind parallel groups 9 months
Participants	Country: USA n = 52 elderly healthy women, mean age 82 (75 to 91 years), mixed natural and surgical menopause ? Exclusion criteria: dementia, depression, HRT contraindications
Interventions	1. CEE oral 0.625 mg/day + (in women without hysterectomy) MPA 5 mg for 13 days every third month 2. Placebo
Outcomes	- Verbal Memory: Paired Associate Learning, Verbal Fluency - Speed/concentration: Trial Making Test (TMT) A / B, Cancellation Random Letters and Figures test
Notes	No effect HRT 0/5 tests

Ditkoff 1991

Methods	Design: randomized double-blind parallel groups 3 months
Participants	Country: USA n = 36 healthy women, mean age: 53 (45 to 60) Hysterectomized, asymptomatic patients Exclusion criteria: active disease or psychoactive medication use
Interventions	1. CEE oral 0.625 or 1.25 mg/day 2. Placebo
Outcomes	- Short term memory (STM) and Concentration (Digit Span) - Speed (Digit Symbol Modalities test or DSST)
Notes	No effect HRT 0/2 tests

Duka 2000

Methods	Design: randomized double-blind parallel groups 3 weeks
Participants	Country: UK n = 37 healthy women, mean age: 65 (55 to 75). Natural menopausal ? Status unclear. Exclusion criteria: abnormal weight, mental, neurological, or chronic disease, substance abuse
Interventions	1. 17 beta E2 transdermal patch (0.1 mg /day) 2. Placebo
Outcomes	- Visual Memory: Picture Recall (BETAM), Visual Paired Associates Learning (CANTAB)

Hormone replacement therapy for cognitive function in postmenopausal women (Review)

Duka 2000 (Continued)

- Visuospatial speed: mental rotation/ reaction time (RT)
- Executive function: CANTAB ID/ED test, Stroop , Random Number Generation Test, Tower of London

Notes	HRT> placebo memory (BETAM and PAL) and visual rotation speed, 3/ 7 tests (baseline differences seem larger than effect size)
-------	---

Dunkin 2005

Methods	Design: randomized, double-blind, parallel groups, 10 weeks
Participants	Country: USA n = 17, healthy women, mean age 57, natural or surgical menopause not stated
Interventions	1. E2 transdermal 0.1 mg/day 2. Placebo
Outcomes	-Verbal memory/language : CVLT, WMS-R PAL and logical memory, Fluency -Non-verbal memory: WMS-R VRT, Rey-Osterreith -Attention, Speed processing, Executive functions: Stroop, Grooved Pegboard Test WCST, TMT-B, Digit Span, DSST, WAIS-R picture arrangement
Notes	No effect of HRT 0/13 tests (but trend level significance in attention and speed of information processing tests combined)

Fedor-Freyberg 1977

Methods	Design: randomized double-blind parallel groups 3 months
Participants	Country: Sweden n = 21 women with many climacteric symptoms, mean age: 57 (47 to 70). Exclusion criteria: mental disease
Interventions	1. E2 oral 2 mg/day 2. Placebo
Outcomes	- Speed/concentration: Choice RT, KVT (visual search) , Stroop, USTM test (attention, STM & reasoning), Visual Search test
Notes	HRT > placebo 5/5 tests

Goebel 1995

Methods	Design: randomized double-blind parallel groups 2, 4 and 8 months
Participants	Country: USA

Goebel 1995 (Continued)

n = 87 elderly women, mean age: 75 (70 to 88) Exclusion criteria: age < 69, smoking, psychotropic medication, depression, cancer, liver or vascular disease

Interventions	1. CEE oral 0.625 mg (+ MPA 5 mg for 12 days every third month) 2. Placebo
Outcomes	- Speed: TMT-B
Notes	No effect HRT 0/1 test

Greenspan 2005

Methods	Design: randomized double-blind parallel groups, 3 years
Participants	Country: USA, n = 373 healthy women, mean age 71.3 years, 243 with a uterus and 130 without a uterus
Interventions	1. CEE oral (0.625 mg/day) or CEE oral (0.625 mg/day) + medroxyprogesterone acetate (2.5 mg/day) 2. Placebo
Outcomes	- General: MMSE
Notes	No effect of HRT 0/1 test Half of the participants also took alendronate and all participants had co-intervention (calcium and vitamin D supplement and a multivitamin)

Hackman 1977

Methods	Design: randomized double-blind parallel groups, 6 months
Participants	Country: UK n = 18 healthy women, aged between 29 and 68 years. Mixed surgical and natural menopausal, many climacteric symptoms.
Interventions	1. piperazine E1 oral , 1.5 mg/2x day 2. Placebo 2x day for 3 weeks each month
Outcomes	- Verbal memory: Guild memory test
Notes	No effect HRT 0/1 test

Janowsky 2000

Methods	Design: randomized double-blind parallel groups
---------	---

Janowsky 2000 (Continued)

1 month

Participants	Country: USA n = 13 healthy women, mean age: 69 (61 to 74) Natural menopausal ? (all had normal gynaecological exams)
Interventions	1. CEE 0.625 mg/day 2. Placebo
Outcomes	- Visual memory: Working memory task (with abstract figures)
Notes	No effect HRT 0/1 test

Linzmayr 2001

Methods	Design: randomized double-blind parallel groups (3-arm trial) 2 months
Participants	Country: Austria n = 49, mean age: 58 (46 to 67 years). Mixed ? natural and surgical. Inclusion: Sleep disorder (DSM), menopausal > 24 months
Interventions	1. E2 valerate (2 mg) + dienogest (P, 3 mg)/day 2. E2 2 mg/day 3. Placebo
Outcomes	- Verbal memory (Grunberger, also includes associative, numeric and total verbal memory) - Visual memory (Benton Visual Retention Test) - Intelligence (Mehrfachwahl-Wortschatz Test or MWT)
Notes	HRT E2+P> E2/ Placebo on associative verbal memory E2 > E2+P on numerical memory (baseline difference) E2+P >E2 on visual memory (only after open label, 4 mnth HRT) No effect Intelligence or other verbal memory

Phillips 1992

Methods	Design: randomized double-blind parallel groups 2 months
Participants	Country: Canada n = 19 healthy women, mean age: 48 (SD 5) All surgical menopausal. Exclusion criteria: contra-indications for HRT, mental disease, ever use HRT
Interventions	1. E2 (I.M.) 10 mg/month 2. Placebo (I.M. sesame oil 1 ml)
Outcomes	- Verbal memory: Paragraph recall, Paired Associate learning, Digit Span - Visuospatial Memory: Visual Reproduction Test (VRT of the Wechsler Memory Scale or WMS)

Phillips 1992 (Continued)

Notes	HRT > placebo 2/4 tests: verbal memory (not digit span) no effect VRT
-------	--

Polo-Kantola 1998

Methods	Design: randomized double-blind cross-over, 3 months drug-1 month- 3 month drug wash-out
Participants	Country: Finland n=62 very healthy hysterectomized women, mean age: 56 (47 to 65). Exclusion criteria: neurological, cardiovascular, endocrinological, and mental disease, malignancies, smoking >10 cigarettes, psychoactive medication
Interventions	1. E2 transdermal 2.5 mg/day (< 56 years of age) or 0.05 mg/day (> 55 years of age) 2. Placebo
Outcomes	- Memory: Benton Visual Retention Test (BVRT) and Digit Span - Speed: Stroop, RT, Letter cancellation, Paced Auditory Serial Addition Test (PASAT), DSST, subtraction test
Notes	no effect HRT 0/8 tests

Resnick 2006

Methods	Design: randomised, double-blind, parallel groups, 3 years
Participants	Country: USA n = 1417 (subgroup WHIMS), mean age 73.7, healthy women, natural or surgical menopause not stated
Interventions	1. CEE oral (0.625 mg/day) + medroxyprogesterone (2.5 mg/day) 2. Placebo
Outcomes	-General/Verbal knowledge: PMA vocabulary -Memory/language: BVRT, CVLT (IR+DR), Verbal Fluency (2 types) -Attention, working memory: Digit Span forward and backwards -Spatial ability: Card Rotations -Fine motor skills: Finger tapping
Notes	HRT >placebo 1/9 tests (BVRT) Placebo>HRT 1/9 tests (CVLT)

Shaywitz 1999

Methods	Design: randomized double-blind cross-over, 21 days drug-14 days wash-out- 21 days drug
Participants	Country: USA

Hormone replacement therapy for cognitive function in postmenopausal women (Review)

Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shaywitz 1999 (Continued)

n = 46 healthy postmenopausal, mean age: 51 (SD 5, 33 to 61). Exclusion criteria: MRI abnormalities, IQ < 85, <5 months postmenopausal, HRT for last 3 months

Interventions	1. CEE oral 1.25 mg/day 2. Placebo
Outcomes	- Memory: word and token recall
Notes	No effect HRT 0/2 tests

Sherwin 1988

Methods	Design: randomized double-blind cross-over 3 months drug-1 month placebo- 3 month drug
Participants	Country: Canada n = 50, healthy women, mean age: 45.4. All surgically menopausal. Exclusion criteria: psychological problems.
Interventions	1. E2 (I.M. 10 mg/ mnth), n=10 2. E2 (I.M. 8.5 mg)+ T (I.M. 150 mg) /mnth, n=10 3. T (I.M. 200 mg/mnth), n=10 4. Placebo, n=10
Outcomes	- Memory: Digit Span, Paragraph recall (not WMS!) - Speed: Clerical speed and accuracy - Abstract reasoning
Notes	ERT > placebo 4/4 tests TRT=ERT

Sherwin 1990

Methods	Design: randomized double-blind parallel groups, 2 months
Participants	Country: Canada n=12 healthy women, mean age: 47. All surgical menopausal. Exclusion criteria: psychological or physical disease
Interventions	1. E2 (I.M.) 10 mg/month 2. Placebo
Outcomes	- Verbal memory: Paragraph recall, Paired Associate Learning, Digit span (not sure, mentions 4 tests), - Visuospatial memory: VRT (WMS)
Notes	HRT > placebo 2/4 tests, not on VRT

Shumaker 2003

Methods	Design: randomised, double-blind, parallel groups, 7 years (mean 4.2 years)
Participants	Country: USA n = 4344, healthy women (WHIMS participants) age: 65 + (half below 70 years of age)
Interventions	1. CEE oral 0.625 mg/day + 2.5 mg MPA 2. Placebo
Outcomes	- General: 3 MSE - Diagnosis of mild cognitive impairment (defined by a standardised protocol)
Notes	Placebo>HRT 1/1 test. More women in HT group had substantial decline

Shumaker 2004

Methods	Design: randomised, double-blind, parallel groups, 7 years (mean 5.4 years)
Participants	Country: USA n = 2808, healthy women (WHIMS participants), mean age: 72 (65 to 79 years), hysterectomy
Interventions	1. CEE oral 0.625 mg/day 2. Placebo
Outcomes	- General: 3MSE - Diagnosis of mild cognitive impairment defined by a standardised protocol
Notes	Placebo>HRT 1/1 test

Vanhulle 1976

Methods	Design: randomized double-blind parallel groups, 3 months
Participants	Country: Belgium n = 26 healthy postmenopausal nuns, mean age: 57.6. Mixed surgical and natural menopausal Exclusion criteria: HRT or tranquillizer use
Interventions	1. E3 oral 4 mg/daily 2. Placebo
Outcomes	- Memory: BVRT, Digit span, Labyrinth of Rey - Speed/concentration: DSST, calculations, vigilance test, Bourdon- Wiersma (B-W) test
Notes	HRT> placebo 2/7 tests of alertness and attention (vigilance and B-W test), not on memory and concentration

Wolf 1999

Methods	Design: randomized double-blind parallel groups, 2 weeks
Participants	Country: Germany n = 38, healthy women, mean age: 69 (SD 6). Mixed surgical and natural menopausal. Exclusion criteria: psychiatric, endocrine, cardiovascular or other chronic disease, use of psychoactive medication and steroids.
Interventions	1. E2 transdermal, 0.1 mg/ day 2. Placebo
Outcomes	- Verbal memory: Paired Word recall, Verbal Fluency - Visuospatial memory: Route finding - Mental rotation test (different from Duka) - Speed/concentration: Stroop test
Notes	No effect HRT 0/3 test

Yaffe 2006

Methods	Design: randomized triple-blind parallel groups 2 years
Participants	Country: USA n = 417 women, mean age 67 years (range 60 to 80 years), intact uterus, at least 5 years beyond menopause, normal bone density for age Exclusion: unexplained bleeding, endometrial hyperplasia or endometrium >5mm, abnormal mammo-gram, history of metabolic bone disease, cancer, coronary disease stroke or TIA, venous thromboem-bolism, uncontrolled hypertension, uncontrolled thyroid disease, live disease, fasting triglyceride level >180mg/dL, ever taken fluoride calcitonin or bisphosphonates, or taken E or P within 3 months of ran-domization
Interventions	1. 3.25 cm2 estradiol patch (approx 0.014 mg estradiol/day) 2. Placebo patch All participants had concomitant oral supplements of 400mg calcium (twice daily) and 400IU of vitamin D once daily
Outcomes	All secondary outcomes: - General: 3MSE - Episodic memory: Logical Memory, Immediate and Delayed, Brief Visuospatial Memory Test Immedi-ate and Delayed, Word List Memory, Word List Recall - Attention, visual scanning and executive function: Trails B Test - Language function: Modified Boston Naming Test - Semantic memory and language: Verbal Fluency Test
Notes	No effect

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Drapier 2005	This RCT was a dose finding study with two HRT groups and no placebo control.
Friebely 2001	This RCT was a dose finding study with two HRT groups and no placebo control.
Grady 2002	Participants had cardiovascular disease.
Greendale 1997	Measurement of cognitive outcomes was not by validated scales but were self-reports of the participants.
Hogervorst 1999	Not RCT.
Krug 2003	This RCT compared the effects of estrogen and testosterone; there was no control group.
Maki 2004	Primary outcomes were measures of self-reported personality changes.
McDonald 1952	Confounded (androgens and progesterone also given).
Michael 1970	No measure of cognition.
Natale 2001	Not RCT.
Nielsen 2006	Outcomes were measured by the Women's Health Questionnaire, self reports by the participants.
Ohkura 1998	Not RCT.
Pan 2003	This RCT compared HRT with tibolone; there was no control group.
Pauker 1958	Confounded; no objective measure of cognition.
Rasgon 2005	Single blind trial. The inclusion criteria for the review specify that included studies must be double blind.
Rauramo 1975	Not RCT.
Ryan 2001	No measure of cognition.
Stevens 2005	Outcome was the measurement of hemodynamic activity using MRI; this is a surrogate outcome.
Tierney 2005	Participants were 'memory impaired'.
Viscoli 2005	Participants had cerebrovascular disease.
Wisniewski 2002	A comparison of estrogen with estrogen plus testosterone: there was no control group.

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

Kronos 2006

Trial name or title	Kronos Early Estrogen Prevention Study (KEEPS)
Methods	
Participants	Healthy women within 3 years of the menopause

Kronos 2006 (Continued)

Interventions	(1) CEE 0.45mg/day (2) Transdermal estradiol 50 MCG/day (3) Micronized progesterone 200mg/day X 12d/mth (4) Placebo
Outcomes	Primary: Rate of change of carotid intimal medial thickness by ultrasound Secondary: (1) Plasma lipid profiles (2) Blood clotting factors (3) Serum inflammatory markers (4) Hormone levels (5) Cognitive and affective scores on standard psychometric tests (6) Quality of life
Starting date	September 2005
Contact information	Patricia Crenshaw (info@kronosinstitute.org)
Notes	Multicenter trial with expected completion in December 2010.

Raghnauth 2000

Trial name or title	Alzheimer's disease prevention trial
Methods	
Participants	Healthy women >65 years who have a family history of memory problems and Alzheimer's disease.
Interventions	(1) Estrogen - unknown dose (2) Estrogen + progestogen - unknown dose (3) Placebo
Outcomes	Prevention of dementia
Starting date	unknown
Contact information	A. Raghnauth, Columbia-Presbyterian Medical Center, G.H. Sergievsky Center. sdm9001@nyp.org
Notes	Sponsored by the National Institute on Aging (NIA) in the USA

DATA AND ANALYSES

Comparison 1. The effect of ERT/HRT versus placebo on global cognition scores

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 CAMCOG after 20 weeks of E2 versus placebo	1	115	Mean Difference (IV, Fixed, 95% CI)	0.60 [-0.62, 1.82]
2 3MSE score after CEE + MPA (variable follow up)	1		WMD (Fixed, 95% CI)	Subtotals only
2.1 One year follow up	1		WMD (Fixed, 95% CI)	-0.22 [-0.45, 0.01]
2.2 2 years follow up	1		WMD (Fixed, 95% CI)	-0.13 [-0.37, 0.11]
2.3 3 years follow up	1		WMD (Fixed, 95% CI)	-0.36 [-0.61, -0.11]
2.4 4 years follow up	1		WMD (Fixed, 95% CI)	-0.52 [-0.81, -0.23]
2.5 5 years follow up	1		WMD (Fixed, 95% CI)	-0.16 [-0.58, 0.26]
3 3MSE score after CEE alone (variable follow up)	2		WMD (Fixed, 95% CI)	Subtotals only
3.1 One year follow up	2		WMD (Fixed, 95% CI)	-0.44 [-0.73, -0.16]
3.2 2 years follow up	2		WMD (Fixed, 95% CI)	-0.25 [-0.54, 0.03]
3.3 3 years follow up	1		WMD (Fixed, 95% CI)	-0.21 [-0.56, 0.14]
3.4 4 years follow up	1		WMD (Fixed, 95% CI)	-0.17 [-0.55, 0.21]
3.5 5 years follow up	1		WMD (Fixed, 95% CI)	-0.26 [-0.72, 0.20]
3.6 6 years follow up	1		WMD (Fixed, 95% CI)	-0.45 [-0.99, 0.09]
4 Folstein Mini-Mental State Examination with CEE alone or CEE + MPA after 3 years	1	373	Mean Difference (IV, Fixed, 95% CI)	-0.1 [-0.34, 0.14]

Analysis 1.1. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 1 CAMCOG after 20 weeks of E2 versus placebo.

Analysis 1.2. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 2 3MSE score after CEE + MPA (variable follow up).

Analysis 1.3. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 3 3MSE score after CEE alone (variable follow up).

Analysis 1.4. Comparison 1 The effect of ERT/HRT versus placebo on global cognition scores, Outcome 4 Folstein Mini-Mental State Examination with CEE alone or CEE + MPA after 3 years.

Comparison 2. The effect of ERT/HRT versus placebo on prevention of mild cognitive impairment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Mild cognitive impairment (based on WHIMS protocol)	2	7474	Odds Ratio (M-H, Fixed, 95% CI)	1.20 [0.93, 1.55]
1.1 Unopposed E therapy at 5 years follow up	1	2942	Odds Ratio (M-H, Fixed, 95% CI)	1.34 [0.95, 1.90]
1.2 Combined E and P Rx at 4 years follow up	1	4532	Odds Ratio (M-H, Fixed, 95% CI)	1.05 [0.72, 1.54]

Analysis 2.1. Comparison 2 The effect of ERT/HRT versus placebo on prevention of mild cognitive impairment, Outcome 1 Mild cognitive impairment (based on WHIMS protocol).

Comparison 3. The effect of ERT/HRT versus placebo on verbal memory and language tests

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Paired Associates (intermediate recall) after HRT or placebo	3	57	Std. Mean Difference (IV, Random, 95% CI)	0.73 [-0.02, 1.48]
1.1 After 2 months of E2 i.m. or placebo	2	31	Std. Mean Difference (IV, Random, 95% CI)	1.09 [0.12, 2.07]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.2 After 10 weeks E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.25 [-0.52, 1.03]
2 Paired associate learning (different scoring methods immediate recall)	3	83	Std. Mean Difference (IV, Random, 95% CI)	0.66 [0.02, 1.31]
2.1 After 9 months with CEE + MPA or placebo	1	52	Std. Mean Difference (IV, Random, 95% CI)	0.30 [-0.27, 0.88]
2.2 After 2 to 3 months with E2 i.m. or placebo	2	31	Std. Mean Difference (IV, Random, 95% CI)	1.02 [0.19, 1.85]
3 Paired Associate (delayed recall) after HRT or placebo	3	97	Std. Mean Difference (IV, Random, 95% CI)	0.12 [-0.29, 0.53]
3.1 After 9 months with CEE+ MPA	1	52	Std. Mean Difference (IV, Random, 95% CI)	0.0 [-0.57, 0.57]
3.2 After 2 months of E2 i.m. or placebo	1	19	Std. Mean Difference (IV, Random, 95% CI)	0.59 [-0.33, 1.52]
3.3 After 10 weeks E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.0 [-0.77, 0.77]
4 Other Paired Associate tests	2	70	Std. Mean Difference (IV, Random, 95% CI)	-0.09 [-0.56, 0.39]
4.1 After 2 months of E2 oral + P or placebo	1	32	Std. Mean Difference (IV, Random, 95% CI)	0.17 [-0.52, 0.86]
4.2 After 2 weeks of E2 transdermal or placebo	1	38	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.95, 0.34]
5 Paragraph recall (immediate) after 2 to 3 months of E2 i.m. or placebo	2	31	Mean Difference (IV, Fixed, 95% CI)	4.64 [-2.10, 11.37]
6 Paragraph recall (delayed) after 2 months with E2 i.m. or placebo	1	19	Mean Difference (IV, Fixed, 95% CI)	5.3 [-5.39, 15.99]
7 Verbal Fluency (semantic memory)	4	582	Std. Mean Difference (IV, Random, 95% CI)	0.02 [-0.14, 0.19]
7.1 After 9 months of CEE + MPA or placebo	1	52	Std. Mean Difference (IV, Random, 95% CI)	-0.22 [-0.80, 0.35]
7.2 After 2 weeks of E2 transdermal or placebo	1	38	Std. Mean Difference (IV, Random, 95% CI)	0.31 [-0.34, 0.95]
7.3 After 20 weeks of E2 oral or placebo	1	115	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.46, 0.27]
7.5 After 2 years low E transdermal versus placebo	1	377	Std. Mean Difference (IV, Random, 95% CI)	0.06 [-0.14, 0.26]
8 Word list recall tests	4	991	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.10, 0.15]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.1 CVLT memory after 20 weeks with E2 oral or 10 weeks with E2 transdermal or placebo	2	141	Std. Mean Difference (IV, Random, 95% CI)	-0.12 [-0.45, 0.21]
8.2 CVLT (short delay free recall) after 10 weeks E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.30 [-0.48, 1.07]
8.3 CVLT (long delay free recall) after 10 weeks E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.75, 0.80]
8.4 CVLT (recognition) after 10 weeks E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.0 [-0.77, 0.77]
8.5 Guild memory test after 6 months of piperazine E1 or placebo	1	18	Std. Mean Difference (IV, Random, 95% CI)	0.72 [-0.24, 1.68]
8.6 Word List Memory Test after 2 years low E transdermal versus placebo	1	377	Std. Mean Difference (IV, Random, 95% CI)	0.08 [-0.12, 0.29]
8.7 Word List Recall after 2 yrs low E transdermal versus placebo	1	377	Std. Mean Difference (IV, Random, 95% CI)	-0.02 [-0.22, 0.18]
9 Wechsler Memory Scale - Revised (logical memory - immediate recall)	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
9.1 10 weeks of E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-2.61 [-10.35, 5.13]
9.2 After 1 year of low E transdermal or placebo	1	388	Mean Difference (IV, Fixed, 95% CI)	-0.13 [-0.84, 0.58]
9.3 After 2 years of low E transdermal or placebo	1	377	Mean Difference (IV, Fixed, 95% CI)	-0.52 [-1.29, 0.25]
10 Wechsler Memory Scale - Revised (logical memory - delayed recall)	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
10.1 After 10 weeks of E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-0.32 [-9.25, 8.61]
10.2 After 1 year of low E transdermal or placebo	1	388	Mean Difference (IV, Fixed, 95% CI)	-0.07 [-0.77, 0.63]
10.3 After 2 years of low E transdermal or placebo	1	377	Mean Difference (IV, Fixed, 95% CI)	-0.43 [-1.19, 0.33]
11 Wechsler Memory Scale - Revised (logical memory - % retained) after 10 weeks E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	7.34 [-5.06, 19.74]
12 Boston Naming Test after low E transdermal versus placebo	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
12.1 After 1 year	1	388	Mean Difference (IV, Fixed, 95% CI)	0.07 [-0.07, 0.21]
12.2 After 2 years	1	377	Mean Difference (IV, Fixed, 95% CI)	0.0 [-0.15, 0.15]
13 PMA Vocabulary after 1+ years of HRT versus placebo	1		WMD (Fixed, 95% CI)	-0.1 [-0.43, 0.23]
14 Verbal Fluency - letter after 1+ years of HRT versus placebo	1		WMD (Fixed, 95% CI)	-0.28 [-0.85, 0.29]
15 Verbal Fluency - categories after 1+ years of HRT versus placebo	1		WMD (Fixed, 95% CI)	-0.09 [-0.46, 0.28]
16 Word list recall tests after 1+ years of HRT versus placebo (inverse variance)	1		WMD (Fixed, 95% CI)	Subtotals only
16.1 CVLT total	1		WMD (Fixed, 95% CI)	-0.52 [-0.91, -0.13]
16.2 CVLT Interference	1		WMD (Fixed, 95% CI)	0.05 [-0.11, 0.21]
16.3 CVLT short delay	1		WMD (Fixed, 95% CI)	-0.24 [-0.44, -0.04]
16.4 CVLT long delay	1		WMD (Fixed, 95% CI)	-0.23 [-0.43, -0.03]

Analysis 3.1. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 1 Paired Associates (intermediate recall) after HRT or placebo.

Analysis 3.2. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 2 Paired associate learning (different scoring methods immediate recall).

Analysis 3.3. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 3 Paired Associate (delayed recall) after HRT or placebo.

Study or subgroup	Treatment		Control		Std. Mean Difference			Weight	Std. Mean Difference		
	N	Mean(SD)	N	Mean(SD)	Random, 95% CI						
Test for subgroup differences: Chi²=1.27, df=1 (P=0.53), I²=0%											
					Favours control	-4	-2	0	2	4	Favours treatment

Analysis 3.4. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 4 Other Paired Associate tests.

Study or subgroup	Treatment		Control		Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
	N	Mean(SD)	N	Mean(SD)			
3.4.1 After 2 months of E2 oral + P or placebo							
Linzmayr 2001	16	2 (5)	16	1 (6.4)		46.21%	0.17[-0.52,0.86]
Subtotal ***	16		16			46.21%	0.17[-0.52,0.86]
Heterogeneity: Not applicable							
Test for overall effect: Z=0.48(P=0.63)							
3.4.2 After 2 weeks of E2 transdermal or placebo							
Wolf 1999	21	0.3 (3)	17	1.4 (3.5)		53.79%	-0.31[-0.95,0.34]
Subtotal ***	21		17			53.79%	-0.31[-0.95,0.34]
Heterogeneity: Not applicable							
Test for overall effect: Z=0.94(P=0.35)							
Total ***	37		33			100%	-0.09[-0.56,0.39]
Heterogeneity: Tau ² =0; Chi ² =0.97, df=1(P=0.32); I ² =0%							
Test for overall effect: Z=0.36(P=0.72)							
Test for subgroup differences: Chi ² =0.97, df=1 (P=0.32), I ² =0%							
					-4 -2 0 2 4		
					Favours control	Favours treatment	

Analysis 3.5. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 5 Paragraph recall (immediate) after 2 to 3 months of E2 i.m. or placebo.

Study or subgroup	Treatment		Control		Mean Difference Fixed, 95% CI	Weight	Mean Difference Fixed, 95% CI
	N	Mean(SD)	N	Mean(SD)			
Phillips 1992	10	6.8 (10.2)	9	1.5 (13.2)		39.66%	5.3[-5.39,15.99]
Sherwin 1990	6	4.7 (6.5)	6	0.5 (8.7)		60.34%	4.2[-4.47,12.87]
Total ***	16		15			100%	4.64[-2.1,11.37]
Heterogeneity: $\tau^2=0$; $\chi^2=0.02$, $df=1$ ($P=0.88$); $I^2=0\%$ Test for overall effect: $Z=1.35$ ($P=0.18$)							
					-10 -5 0 5 10		Favours control Favours treatment

Analysis 3.6. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 6 Paragraph recall (delayed) after 2 months with E2 i.m. or placebo.

Analysis 3.7. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 7 Verbal Fluency (semantic memory).

Analysis 3.8. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 8 Word list recall tests.

Analysis 3.9. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 9 Wechsler Memory Scale - Revised (logical memory - immediate recall).

Analysis 3.10. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 10 Wechsler Memory Scale - Revised (logical memory - delayed recall).

Analysis 3.11. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 11 Wechsler Memory Scale - Revised (logical memory - % retained) after 10 weeks E2 transdermal or placebo.

Analysis 3.12. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 12 Boston Naming Test after low E transdermal versus placebo.

Analysis 3.13. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 13 PMA Vocabulary after 1+ years of HRT versus placebo.

Analysis 3.14. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 14 Verbal Fluency - letter after 1+ years of HRT versus placebo.

Analysis 3.15. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 15 Verbal Fluency - categories after 1+ years of HRT versus placebo.

Analysis 3.16. Comparison 3 The effect of ERT/HRT versus placebo on verbal memory and language tests, Outcome 16 Word list recall tests after 1+ years of HRT versus placebo (inverse variance).

Comparison 4. The effect of ERT/HRT versus placebo on visual memory

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Visuo(spatial) memory tests	8	1342	Std. Mean Difference (IV, Random, 95% CI)	0.04 [-0.09, 0.16]
1.1 VRT immediate recall after 2 months of E2 or placebo	1	19	Std. Mean Difference (IV, Random, 95% CI)	0.26 [-0.64, 1.17]
1.2 Paired objects recall (BETAM) after 3 weeks of E2 transdermal or placebo	1	37	Std. Mean Difference (IV, Random, 95% CI)	0.37 [-0.28, 1.02]
1.3 Visuospatial map recall after 2 weeks of E2 transdermal or placebo	1	38	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.74, 0.54]
1.4 Benton Visual Retention Test (# correct) after 2 months of E2 oral or placebo	1	32	Std. Mean Difference (IV, Random, 95% CI)	-0.34 [-1.04, 0.35]
1.5 Visual immediate recall (drop in errors) after 1 month of CEE or placebo	1	13	Std. Mean Difference (IV, Random, 95% CI)	0.07 [-1.02, 1.16]
1.6 Visual memory for faces (total score) after 20 weeks E2 oral or placebo	1	115	Std. Mean Difference (IV, Random, 95% CI)	0.43 [0.06, 0.79]
1.7 Visual memory for faces (delayed recall) after 20 weeks E2 oral or placebo	1	115	Std. Mean Difference (IV, Random, 95% CI)	0.06 [-0.30, 0.43]
1.8 Visuospatial Organisation (Block Design) after 20 weeks of E2 oral or placebo	1	115	Std. Mean Difference (IV, Random, 95% CI)	0.32 [-0.05, 0.69]
1.9 Rey-Osterreith (delayed recall) after 10 weeks of E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-1.28, 0.29]
1.10 Wechsler Memory Scale - Revised (visual reproduction immediate) after 10 weeks of E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	-0.43 [-1.21, 0.35]
1.11 Wechsler Memory Scale - Revised (visual reproduction delayed) after 10 weeks of E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.74, 0.80]
1.12 Wechsler Memory Scale - Revised (visual reproduction recognition) after 10 weeks of E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	-0.55 [-1.34, 0.23]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.13 Brief Visuospatial Memory Test (immediate) after 2 years low E transdermal versus placebo	1	377	Std. Mean Difference (IV, Random, 95% CI)	-0.02 [-0.22, 0.18]
1.14 Brief Visuospatial Memory Test (delayed) after 2 years low E transdermal versus placebo	1	377	Std. Mean Difference (IV, Random, 95% CI)	-0.03 [-0.24, 0.17]
2 Visuospatial memory tests (inverse variance)	1		WMD (Fixed, 95% CI)	-0.27 [-0.49, -0.05]
2.1 Benton Visual Retention Test after 1+ years of HRT versus placebo	1		WMD (Fixed, 95% CI)	-0.27 [-0.49, -0.05]

Analysis 4.1. Comparison 4 The effect of ERT/HRT versus placebo on visual memory, Outcome 1 Visuo(spatial) memory tests.

Analysis 4.2. Comparison 4 The effect of ERT/HRT versus placebo on visual memory, Outcome 2 Visuospatial memory tests (inverse variance).

Comparison 5. The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Test of simple speed of information processing	1	52	Mean Difference (IV, Fixed, 95% CI)	-10.5 [-24.05, 3.05]
1.1 TMT-A after 9 months of CEE + MPA or placebo	1	52	Mean Difference (IV, Fixed, 95% CI)	-10.5 [-24.05, 3.05]
2 Tests of other information processing	3	179	Std. Mean Difference (IV, Random, 95% CI)	-0.01 [-0.31, 0.28]
2.2 Stroop (digits) after 3 weeks of E2 transdermal or placebo	1	37	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.75, 0.55]
2.3 Stroop (colour) after 2 - 10 weeks of E2 transdermal or placebo	2	64	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.61, 0.38]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.5 Stroop (word) after 10 weeks of E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.13 [-0.65, 0.90]
2.8 Pegboard (dominant hand) after 10 weeks E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.85, 0.69]
2.9 Pegboard (non-dominant hand) after 10 weeks of E2 transdermal or placebo	1	26	Std. Mean Difference (IV, Random, 95% CI)	0.30 [-0.48, 1.07]
3 Attention and working memory tests after 1+ year HRT versus placebo (inverse variance)	1		WMD (Fixed, 95% CI)	Subtotals only
3.1 Digits forward	1		WMD (Fixed, 95% CI)	0.06 [-0.06, 0.18]
4 Fine Motor Speed after 1+ yrs of HRT versus placebo (inverse variance)	1		WMD (Fixed, 95% CI)	Subtotals only
4.1 Finger tapping dominant	1		WMD (Fixed, 95% CI)	0.0 [-0.41, 0.41]
4.2 Finger tapping non dominant	1		WMD (Fixed, 95% CI)	0.02 [-0.31, 0.35]
5 Tests of general information processing speed and recognition	3	91	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-1.40, 1.31]
5.1 Wechsler Adult Intelligence Scale - Revised (Digit Symbol) after 10 weeks of E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-3.11 [-13.27, 7.05]
5.2 Wechsler Adult Intelligence Scale - Revised (Digit Span) after 9-10 weeks of E2 or placebo	3	65	Mean Difference (IV, Fixed, 95% CI)	0.01 [-1.35, 1.37]

Analysis 5.1. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 1 Test of simple speed of information processing.

Study or subgroup	Treatment		Control		Mean Difference Fixed, 95% CI	Weight	Mean Difference Fixed, 95% CI
	N	Mean(SD)	N	Mean(SD)			
5.1.1 TMT-A after 9 months of CEE + MPA or placebo							
Binder 2001	34	-7.5 (25)	18	3 (23)		100%	-10.5[-24.05,3.05]
Subtotal ***	34		18			100%	-10.5[-24.05,3.05]
Heterogeneity: Not applicable							
Test for overall effect: Z=1.52(P=0.13)							
Total ***	34		18			100%	-10.5[-24.05,3.05]
Heterogeneity: Not applicable							
Test for overall effect: Z=1.52(P=0.13)							
					Favours treatment -100 -50 0 50 100	Favours control	

Analysis 5.2. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 2 Tests of other information processing.

Analysis 5.3. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 3 Attention and working memory tests after 1+ year HRT versus placebo (inverse variance).

Study or subgroup	Treatment N	Control N	WMD (SE)	WMD IV, Fixed, 95% CI	Weight	WMD IV, Fixed, 95% CI
Test for overall effect: $Z=1(P=0.32)$						
Favours control -1 -0.5 0 0.5 1 Favours treatment						

Analysis 5.4. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 4 Fine Motor Speed after 1+ yrs of HRT versus placebo (inverse variance).

Study or subgroup	Treatment N	Control N	WMD (SE)	WMD IV, Fixed, 95% CI	Weight	WMD IV, Fixed, 95% CI
5.4.1 Finger tapping dominant						
Resnick 2006	1	1	0 (0.21)		100%	0[-0.41,0.41]
Subtotal (95% CI)					100%	0[-0.41,0.41]
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
5.4.2 Finger tapping non dominant						
Resnick 2006	1	1	0 (0.17)		100%	0.02[-0.31,0.35]
Subtotal (95% CI)					100%	0.02[-0.31,0.35]
Heterogeneity: Not applicable						
Test for overall effect: $Z=0.12(P=0.91)$						
Test for subgroup differences: $\text{Chi}^2=0.01, \text{df}=1 (P=0.94), I^2=0\%$						
Favours control -1 -0.5 0 0.5 1 Favours treatment						

Analysis 5.5. Comparison 5 The effect of ERT/HRT versus placebo on speed tests, attention and manual dexterity, Outcome 5 Tests of general information processing speed and recognition.

Study or subgroup	Treatment		Control		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	Fixed, 95% CI		Fixed, 95% CI
5.5.1 Wechsler Adult Intelligence Scale - Revised (Digit Symbol) after 10 weeks of E2 transdermal or placebo							
Dunkin 2005	12	-0.2 (14.9)	14	2.9 (10.8)		1.77%	-3.11[-13.27,7.05]
Subtotal ***	12		14			1.77%	-3.11[-13.27,7.05]
Heterogeneity: Not applicable							
Test for overall effect: Z=0.6(P=0.55)							
5.5.2 Wechsler Adult Intelligence Scale - Revised (Digit Span) after 9-10 weeks of E2 or placebo							
Dunkin 2005	12	0 (4)	14	-0.2 (6.2)		11.72%	0.25[-3.7,4.2]
Phillips 1992	10	0.1 (1.8)	9	0.3 (1.6)		74.61%	-0.2[-1.76,1.36]
Sherwin 1988	10	0.5 (4.5)	10	-0.6 (4.5)		11.9%	1.1[-2.82,5.02]
Subtotal ***	32		33			98.23%	0.01[-1.35,1.37]
Heterogeneity: Tau²=0; Chi²=0.38, df=2(P=0.83); I²=0%							
Test for overall effect: Z=0.02(P=0.99)							
Total ***	44		47			100%	-0.04[-1.4,1.31]
Heterogeneity: Tau²=0; Chi²=0.74, df=3(P=0.86); I²=0%							
Test for overall effect: Z=0.06(P=0.95)							
Test for subgroup differences: Chi²=0.36, df=1 (P=0.55), I²=0%							
Favours control -10 -5 0 5 10 Favours treatment							

Comparison 7. The effect of ERT/HRT versus placebo on tests of mental rotations

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Mental Rotations time (120 degrees only) after 3 weeks of E2 transdermal or placebo	1	37	Mean Difference (IV, Fixed, 95% CI)	-84.6 [-442.28, 273.08]
2 Mental Rotations number after 2 weeks of E2 transdermal or placebo	1	38	Mean Difference (IV, Fixed, 95% CI)	0.73 [-1.79, 3.25]
3 Card rotations after 1+ year of HRT versus placebo (inverse variance)	1		WMD (Fixed, 95% CI)	-0.39 [-1.76, 0.98]

Analysis 7.1. Comparison 7 The effect of ERT/HRT versus placebo on tests of mental rotations, Outcome 1 Mental Rotations time (120 degrees only) after 3 weeks of E2 transdermal or placebo.

Analysis 7.2. Comparison 7 The effect of ERT/HRT versus placebo on tests of mental rotations, Outcome 2 Mental Rotations number after 2 weeks of E2 transdermal or placebo.

Analysis 7.3. Comparison 7 The effect of ERT/HRT versus placebo on tests of mental rotations, Outcome 3 Card rotations after 1+ year of HRT versus placebo (inverse variance).

Comparison 8. The effect of ERT/HRT versus placebo on accuracy

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clerical speed and accuracy (# correct) after 3 months of E2 i.m. or placebo	1	20	Mean Difference (IV, Fixed, 95% CI)	6.0 [4.72, 7.28]
2 Letter Cancellation (# correct) after 9 months of CEE or placebo	1	52	Mean Difference (IV, Fixed, 95% CI)	0.0 [-0.06, 0.06]
3 Random number generation after 3 weeks of E2 transdermal or placebo	1	37	Mean Difference (IV, Fixed, 95% CI)	0.8 [-5.54, 7.14]

Analysis 8.1. Comparison 8 The effect of ERT/HRT versus placebo on accuracy, Outcome 1 Clerical speed and accuracy (# correct) after 3 months of E2 i.m. or placebo.

Analysis 8.2. Comparison 8 The effect of ERT/HRT versus placebo on accuracy, Outcome 2 Letter Cancellation (# correct) after 9 months of CEE or placebo.

Analysis 8.3. Comparison 8 The effect of ERT/HRT versus placebo on accuracy, Outcome 3 Random number generation after 3 weeks of E2 transdermal or placebo.

Comparison 9. The effect of ERT/HRT versus placebo on executive function

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Wisconsin Card Sorting Test (categories) after 10 weeks E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-0.37 [-1.46, 0.72]
2 Wisconsin Card Sorting Test (perseverative responses) after 10 weeks E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	3.14 [-10.39, 16.67]
3 Wisconsin Card Sorting Test (% conceptual level responses) after 10 weeks E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-4.03 [-21.33, 13.27]
4 Weschler Adult Intelligence Scale - Revised (picture arrangement) after 10 weeks E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-0.94 [-6.16, 4.28]
5 Tower of London after 3 weeks of E2 transdermal or placebo	1	37	Mean Difference (IV, Fixed, 95% CI)	0.8 [-5.54, 7.14]
6 Abstract reasoning (DAT) after 3 months of IM E2 or placebo	1	20	Mean Difference (IV, Fixed, 95% CI)	6.80 [5.52, 8.08]
7 Controlled Word Association Test (COWAT) after 10 weeks E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-2.99 [-13.31, 7.33]
8 Stroop test (interference time) after 10 weeks of E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-6.69 [-31.85, 18.47]
9 Stroop (interference errors) after 10 weeks of E2 transdermal or placebo	1	26	Mean Difference (IV, Fixed, 95% CI)	-0.47 [-1.43, 0.49]
10 Trails B with either CEE + MPA or low E transdermal versus placebo	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
10.1 After 9 months	1	52	Mean Difference (IV, Fixed, 95% CI)	-16.1 [-34.42, 2.22]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
10.2 After 1 year	1	388	Mean Difference (IV, Fixed, 95% CI)	-6.02 [-16.12, 4.08]
10.3 After 2 years	1	377	Mean Difference (IV, Fixed, 95% CI)	2.29 [-7.70, 12.28]
11 Digits Backward after 1+ year HRT versus placebo	1		WMD (Fixed, 95% CI)	0.08 [-0.04, 0.20]

Analysis 9.1. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 1 Wisconsin Card Sorting Test (categories) after 10 weeks E2 transdermal or placebo.

Analysis 9.2. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 2 Wisconsin Card Sorting Test (perseverative responses) after 10 weeks E2 transdermal or placebo.

Analysis 9.3. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 3 Wisconsin Card Sorting Test (% conceptual level responses) after 10 weeks E2 transdermal or placebo.

Analysis 9.4. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 4 Weschler Adult Intelligence Scale - Revised (picture arrangement) after 10 weeks E2 transdermal or placebo.

Analysis 9.5. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 5 Tower of London after 3 weeks of E2 transdermal or placebo.

Analysis 9.6. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 6 Abstract reasoning (DAT) after 3 months of IM E2 or placebo.

Analysis 9.7. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 7 Controlled Word Association Test (COWAT) after 10 weeks E2 transdermal or placebo.

Analysis 9.8. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 8 Stroop test (interference time) after 10 weeks of E2 transdermal or placebo.

Analysis 9.9. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 9 Stroop (interference errors) after 10 weeks of E2 transdermal or placebo.

Analysis 9.10. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 10 Trails B with either CEE + MPA or low E transdermal versus placebo.

Analysis 9.11. Comparison 9 The effect of ERT/HRT versus placebo on executive function, Outcome 11 Digits Backward after 1+ year HRT versus placebo.

ADDITIONAL TABLES

Table 1. List of abbreviations

Abbreviation	Full description
AD	Alzheimer's disease
Blessed or BIMC	Blessed Information Memory and Concentration test
BSRT	Buschke's Selective Reminding Test
BSO	Bilateral Salpingo-Oophorectomy (removal of the ovaries)
BVRT	Benton Visual Retention Test
CAMCOG	Cambridge Cognitive Examination for Mental Disorders of the Elderly
CEE	Conjugated Equine Estrogens or Premarin
CVLT	Californian Verbal Learning Test
DAT	Dementia of the Alzheimer's Type
Dep	depressed
DSM	Diagnostic Statistical Manual
DSST	Digit Symbol Substitution Test

Table 1. List of abbreviations *(Continued)*

E1	Estrone ¹
E2	Estradiol
E3	Estriol
ERT	Estrogen Replacement Therapy
HRT	Hormone Replacement Therapy (estrogen plus progestagen)
I.M.	Intramuscular
MANOVA	Multivariate Analyses of Variance
3MSE	Mini-Mental Status Examination
MPA	Medroxy Progesterone Acetate
NINCDS/ADRDA	National Institute of Neurological and Communicative Disorders and the Alzheimer's Disease and Related Disorders Association
RCT	Randomized Controlled Trial
SD	Standard Deviation
SEM	Standard Error of the Mean
SQRT	Square Root
TMT	Trail Making Test
VaD	Vascular Dementia
Var	Variance
VRT WMS	Visual Retention Test of the WMS (Visuospatial Memory)
WHIMS	the Women's Health Initiative Study
WHISCA	the Women's Health Initiative Study of Cognitive Aging
WISDOM	Women's International Study of long Duration estrogen after the Menopause
WMD	Weighted Mean Difference
WMS	Wechsler Memory Scale

Table 2. Quality of included studies

Study ID	Randomisation method	All concealment	Attrition	Power calculation	Intention to treat	Prognostic balance
Almeida 2006	Random list of numbers generated by computer	Adequate	17.5% in ERT group and 32.8% in placebo group dropped	Not reported	Yes - also completers	Balanced for age, education, cognitive test scores, prior use of HRT, quality of life, depression

Table 2. Quality of included studies (Continued)

	er - randomisation in blocks of 8		out either because of side effects, withdrawal of consent or health problems		analysis performed	and history of psychiatric treatment
Binder 2001	Computer generated random permutation procedure in 2:1 ratio	Adequate	22% (10/45) in ERT/HRT group and 4.5% of women in the placebo group dropped out. A further 4 women were excluded (2% in HRT group; 14% in placebo group) because of depression or mild cognitive impairment	Not reported	No	Balanced for age, education, ethnic origin, medications, smoking status, comorbidities and previous HRT use
Duka 2000	Supplied by the pharmaceutical company that supplied the treatment - no details given	Not reported	Nil	Not reported	Yes	Balanced for age and education
Dunkin 2005	Not described	Not reported	33% of ERT group and 36% of placebo group dropped out because of time constraints, concern over randomisation or being lost to follow up	Not reported	No	Balanced for age, IQ, education, ratio of surgical or natural menopause, prior use of ERT, BMI and parity
Greenspan 2005	Randomized lists prepared by the study statistician	Adequate	9.6% of HRT/ERT group and 9.7% of placebo group with similar reasons for each group	Yes, 80% power to detect a difference >2.91 times the within group SD	Yes, also per protocol analysis	Balanced for age, BMI, physical performance, functional assessment and MMSE scores
Hackman 1977	"Coding system"	Not reported	10% (1/10) of ERT group and 10% (1/10) of control group dropped out before the end of the trial and were not included in the analyses	Not reported	No	Not reported
Janowsky 2000	Not described	Not reported	Not reported	Not reported	Not reported	Balanced for age and education
Linzmayr 2001	Not described	Not reported	Not reported	Not reported	Not reported	Not reported
Phillips 1992	Not described	Not reported	19% (6/31) of study group dropped out initially and an-	Not reported	No	Balanced for age, education, memory, mood and menopausal symptoms

Table 2. Quality of included studies (Continued)

			other 19% (6/31) dropped out as a result of severe menopausal symptoms after surgery			
Resnick 2006	Permutated block algorithm according to age group and clinical centre site	Adequate	At the conclusion of the trial (mean follow up 1.35 years), 6.7% of HRT group and 6.9% of placebo group had either died or had dropped out	Yes, 90% power to detect an observed 50% treatment associated slowing of declines in cognition and memory	No, data not carried forward for dropouts	Balanced for age, 3MSE scores, education, ethnicity, BMI, smoking status, prior CVD, hypertension, alcohol use, diabetes, vasomotor symptoms and prior HRT use
Sherwin 1988	Not described	Not reported	17% (9/53) of original study group (another 6 women were recruited and randomised)	Not reported	No	Balanced for age, education, occupation, personality and marital adjustment
Sherwin 1990	Not described	Not reported	Not reported	Not reported	Not reported	Not reported
Shumaker 2003	Permutated block algorithm stratified according to age group and clinical centre site	Adequate	2.5% of HRT group and 1.9% of placebo group either refused further testing, had incomplete data or died	Yes, 80% power to detect a 40% reduction in the incidence of dementia	Yes, for mild cognitive improvement outcome	Balanced for age, education, smoking status, history of stroke or diabetes, prior hormone use and 3MSE scores
Shumaker 2004	Permutated block algorithm stratified by age and clinical centre	Adequate	7.2% of ERT group and 6.9% of placebo group either refused further testing, had incomplete data or died	Yes, 80% power to detect 40% reduction in dementia	Yes, for mild cognitive impairment	Balanced for age, education, ethnicity, smoking status, history of CVD or diabetes, prior hormone therapy and 3MSE scores
Wolf 1999	Not described	Not reported	5% (2/40) were excluded from analysis because of non compliance	Not reported	No	Balanced for age, BMI and education
Yaffe 2006	Randomly permitted blocks of 4, stratified by clinical centre	Adequate	8.2% of ERT group and 11.5% in placebo group did not complete the trial (for similar reasons)	Yes, 80% power to detect between group differences of 0.29SD of change scores	Not true ITT - dropouts not included in analysis. ITT without regard to adherence	Balanced for age, education, ethnicity, smoking status, BMI, depression, vasomotor symptoms and estradiol levels

WHAT'S NEW

Date	Event	Description
5 June 2008	Amended	Converted to new review format.

HISTORY

Protocol first published: Issue 1, 2001

Review first published: Issue 3, 2002

Date	Event	Description
10 September 2007	New citation required and conclusions have changed	<p>The following changes have been made to the review since it was published in 2002:</p> <p>(1) Search updated to March 2006</p> <p>(2) Results updated for the following six included trials:</p> <ul style="list-style-type: none"> - Almeida 2006 - Dunkin 2005 - Resnick 2006 - Shumaker 2003 - Shumaker 2004 - Yaffe 2006 <p>Four other trials that met the inclusion criteria are in the 'Studies Awaiting Assessment' section of the review because the reviewers are awaiting suitable data.</p>

CONTRIBUTIONS OF AUTHORS

-AL took on the update of the review in March 2006: all correspondence, all aspects of the review.

-EH became co-author in March 2006; previously main reviewer, correspondent, all aspects of review

-KY search for trials, extraction of data, interpretation data analyses, updating review

-MR search for trials, extraction of data, interpretation data analyses, updating review

-FAH drafting review versions, selection of trials, interpretation data analyses; she withdrew as co-author in March 2006 when Anne Lethaby took over as main author.

-AY extracted data for the 2006 update.

-Contact editor: Leon Flicker.

-Consumer editor: Lindsay Stead.

-This review has been peer reviewed by two peer reviewers in 2007.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- OPTIMA, Department of Pharmacology, University of Oxford, UK.

External sources

- Bristol-Myers Squibb, USA.

INDEX TERMS**Medical Subject Headings (MeSH)**

*Estrogen Replacement Therapy; Cognition [*drug effects] [physiology]; Cognition Disorders [prevention & control]; Hormone Replacement Therapy [methods]; Postmenopause [*physiology]; Randomized Controlled Trials as Topic

MeSH check words

Female; Humans