This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

AHRQ is the lead Federal agency charged with supporting research designed to improve the quality of health care, reduce its cost, address patient safety and medical errors, and broaden access to essential services. AHRQ sponsors and conducts research that provides evidence-based information on health care outcomes; quality; and cost, use, and access. The information helps health care decisionmakers—patients and clinicians, health system leaders, and policymakers—make more informed decisions and improve the quality of health care services.

Systematic Evidence Review
Number 13

Hormone Replacement Therapy and Cognition

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
2101 East Jefferson Street
Rockviille, MD 20852
www.ahrq.gov

Contract No. 290-97-0018
Task Order No. 2
Technical Support for the U.S. Preventive Services Task Force

Prepared by:

Oregon Health & Science University

Evidence-based Practice Center, Portland, Oregon

Erin Leblanc M.D., M.P.H.
Benjamin Chan, M.S.
Heidi D. Nelson, M.D., M.P.H.

August 2002

Preface

The Agency for Healthcare Research and Quality (AHRQ) sponsors the development of Systematic Evidence Reviews (SERs) through its Evidence-based Practice Program. With guidance from the third U.S. Preventive Services Task Force* (USPSTF) and input from Federal partners and primary care specialty societies, two Evidence-based Practice Centers—one at the Oregon Health Sciences University and the other at Research Triangle Institute-University of North Carolina—systematically review the evidence of the effectiveness of a wide range of clinical preventive services, including screening, counseling, immunizations, and chemoprevention, in the primary care setting. The SERs—comprehensive reviews of the scientific evidence on the effectiveness of particular clinical preventive services—serve as the foundation for the recommendations of the third USPSTF, which provide age- and risk-factor-specific recommendations for the delivery of these services in the primary care setting. Details of the process of identifying and evaluating relevant scientific evidence are described in the "Methods" section of each SER.

The SERs document the evidence regarding the benefits, limitations, and cost-effectiveness of a broad range of clinical preventive services and will help to further awareness, delivery, and coverage of preventive care as an integral part of quality primary health care.

AHRQ also disseminates the SERs on the AHRQ Web site (http://www.ahrq.gov/uspstfix.htm) and disseminates summaries of the evidence (summaries of the SERs) and recommendations of the third USPSTF in print and on the Web. These are available through the AHRQ Web site (http://www.ahrgq.gov/uspstfix.htm), through the National Guideline Clearinghouse (http://www.ncg.gov), and in print through the AHRQ Publications Clearinghouse (1-800-358-9295).

We welcome written comments on this SER. Comments may be sent to: Director, Center for Practice and Technology Assessment, Agency for Healthcare Research and Quality, 6010 Executive Blvd., Suite 300, Rockville, MD 20852.

Carolyn M. Clancy, M.D.
Acting Director
Agency for Healthcare Research and Quality

Robert Graham, M.D.
Director, Center for Practice and Technology Assessment
Agency for Healthcare Research and Quality

---

* The USPSTF is an independent panel of experts in primary care and prevention first convened by the U.S. Public Health Service in 1984. The USPSTF systematically reviews the evidence on the effectiveness of providing clinical preventive services--including screening, counseling, immunization, and chemoprevention--in the primary care setting. AHRQ convened the third USPSTF in November 1998 to update existing Task Force recommendations and to address new topics.

The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services of a particular drug, device, test, treatment, or other clinical service.

STRUCTURED ABSTRACT

Context: Although postmenopausal hormone replacement (HRT) therapy is widely used, its risks and benefits are not well understood. Some observational data suggest that HRT may reduce the risk of cognitive decline and dementia but results have been conflicting.

Objective: To review and evaluate studies of HRT for preventing cognitive decline and dementia in healthy postmenopausal women.

Data Sources: Studies with English-language abstracts identified in MEDLINE (1966-December 2000), HealthSTAR (1975-December 2000), PsychINFO (1984-December 2000); Cochrane Library databases; and articles listed in reference lists of key articles.

Study Selection: Randomized controlled trials and cohort studies were reviewed for the effects of HRT on cognitive decline; cohort and case-control studies were reviewed for dementia risk. No randomized controlled trials regarding dementia risk were identified.

Data Extraction: Twenty-nine studies met inclusion criteria and were rated. Two reviewers rated study quality independently and 100% agreement was reached on Jadad scores and 80% agreement was reached on U.S. Preventive Services Task Force quality scores. A final score was reached through consensus if reviewers disagreed.

Data Synthesis: Studies of cognition were not combined quantitatively because of heterogeneous study design. Women symptomatic from menopause had improvements in verbal memory, vigilance, reasoning, and motor speed, but no enhancement of other cognitive functions. Generally, no benefits were observed in asymptomatic women. A Bayesian meta-analysis of 12 observational studies suggested that HRT was associated with a decreased risk of dementia (summary odds ratio, 0.66; 95% confidence interval, 0.53-0.82). However, possible biases and lack of control for potential confounders limit interpretation of these studies.

Eight studies received a poor quality rating, 3 fair, and 1 good. Studies did not contain enough information to assess adequately the effects of progestin use, various estrogen preparations or doses, or duration of therapy.

Conclusions: In women with menopausal symptoms, HRT may have specific cognitive effects, and future studies should target these effects. The meta-analysis found a decreased risk of dementia in HRT users but most studies had important methodological limitations.

Table Of Contents

Chapter 1. Introduction

Burden of Suffering 1
Prior Recommendations 2
Analytic Frameworks and Key Questions 2

Chapter 2. Methods

Literature Search Strategy 3
Inclusion/Exclusion Criteria 3
Size of Literature Reviewed 4
Data Extraction and Synthesis 5

Chapter 3. Results

Effects on Cognition
- 1. Does HRT improve or stabilize cognition? 9
  - Memory 10
  - Attention 13
  - Concept formation and reasoning 15
  - Motor speed 15
  - Dementia screening measures 16
  - Verbal function 16
  - Influence of symptoms 16
  - Effects of progestins 17
- 2. What is the optimal dose and duration of use? 18
- Summary 19

Effects on Dementia
- 1. Does HRT lower the risk for dementia? 20
- 2. What is the optimal dose and duration of use? 24
- Summary 25

Chapter 4. Discussion

Conclusions 26
Limitations of the Literature 27
Future Research 27

References 29

Figures

1. Benefits of Hormone Replacement Therapy--Analytic Framework 36
2. Adverse Effects of Hormone Replacement Therapy–Analytic Framework 37
3. Results of Meta-analysis of Dementia Studies 38

Tables

1. Summary of Cognitive Test Results 39
2. Influence on Symptoms on Results of Randomized Controlled Trials 41
3. Summary of Meta-analysis Results of HRT and Dementia 42
4. Summary of Evidence 43

Appendices

1. Search Strategy 44
2. Search Results 45
3. Inclusion/Exclusion Criteria 46
4. a. Criteria for Grading Quality of Randomized Controlled Trials: The Jadad Score 47
4. b. Criteria for Grading the Internal Validity of Individual Studies 48
5. Evidence Tables
   - Table 1. HRT and Cognition–Randomized Controlled Trials 50
   - Table 2. HRT and Cognition–Cohort Studies 56
   - Table 3. HRT and Dementia–Case Control Studies 65
   - Table 4. HRT and Dementia–Cohort Studies 75
6. Cognitive Measures Used in Studies 77
7. Abbreviation and Acronym List 80
8. Acknowledgements 82

Chapter 1. Introduction

In this systematic evidence review, we evaluate data on the use of hormone replacement therapy (HRT) to prevent cognitive decline and dementia in healthy postmenopausal women. Specifically, we reviewed the literature reporting effects of HRT on cognitive function in women without dementia. In addition, we conducted a review and meta-analysis of studies of HRT and dementia, focusing on Alzheimer disease. Results of this review have been recently published.[1] This report is part of a larger project on the risks and benefits of HRT prepared for the U.S. Preventive Services Task Force to assist them in making recommendations.

Burden of Suffering

Between 3 and 8 million people in the United States have dementia.[2] The most common type of dementia is Alzheimer disease, which affects 3 to 4 million people and is the fourth leading cause of death in the United States.[3, 4] The incidence of dementia is 1 percent per year in older individuals, although in the most elderly populations this rate may be as high as 2 to 3 percent.[3] One community-based prevalence study in east Boston estimated that almost 50 percent of those aged 85 and over suffered from dementia.[5] Most studies report that after accounting for differences in life expectancy, women have a 1.4 to 3-fold higher risk of Alzheimer disease than men.[3, 6]

The life expectancy of demented patients is greatly reduced. Those with early-onset (before age 60) Alzheimer disease have a median survival of 6.7 to 8.1 years, while survival in those with late-onset disease is 4.8 to 6.0 years.[3, 7] In 1991, the annual cost of Alzheimer disease was estimated to be $67.3 billion.[8] Given the expected growth of the elderly population,

this financial cost, as well as the emotional and physical costs of caring for demented patients, will continue to increase.

Observational data suggest a possible but inconsistent relationship between endogenous estrogen exposure and cognition.[9] Women in the high estrogen phase of the menstrual cycle have been shown to perform better on tests of motor skills compared to when they are in the low estrogen phase of the cycle.[10-12] Bone mineral density, hypothesized to be a marker of cumulative estrogen exposure, has been correlated with risk of cognitive deterioration.[13] Because of the association between endogenous estrogen exposure and cognition, it can be hypothesized that hormone replacement therapy (HRT) after menopause may prevent cognitive decline and the development of dementia.

Prior Recommendations

In 1996, the Second U.S. Preventive Services Task Force thought that the evidence on the effects of HRT on cognitive function was inconclusive.[14] They also reported disparate findings on the association between HRT and Alzheimer disease.

Analytic Frameworks and Key Questions

The analytic frameworks in Figures 1 and 2 show the target populations, interventions, and health outcome measures we examined for the overall question of the benefits and risks of postmenopausal HRT. Arrows 6a and 6b in Figure 1 correspond to issues of HRT and cognition specifically covered in this report. We examined the key questions: (arrow 6a) Does HRT improve or stabilize cognition? What is the optimal dose and duration of use?, (arrow 6b) Does

HRT lower the risk for dementia?, What is the optimal dose and duration of use? These questions guided our literature review and are addressed in the results section below.

We were concerned with HRT as chemoprophylaxis for primary prevention and therefore focused on the use of either estrogen alone or estrogen combined with progestins in healthy, postmenopausal women without dementia. The adverse effects of HRT are discussed in separate reports.

Chapter 2. Methods

Literature Search Strategy

We searched MEDLINE (1966-December 2000), HealthSTAR (1975-December 2000), PsychINFO (1984-December 2000), and Cochrane Library databases using the search strategy in Appendix 1. We combined the Medical Subject Headings *hormone replacement therapy* and *estrogens* with the heading*s dementia, mental processes, cognition disorders*, and *memory disorders*. Additional articles were obtained from reference lists of relevant reviews and from expert reviewers.

Inclusion/Exclusion Criteria

A single reader reviewed all English abstracts identified by the search and determined if a study would be included in this review by applying inclusion and exclusion criteria (Appendix 2). Studies were considered for inclusion if they provided primary data of the effects of HRT on tests of cognition or memory or on any type of dementia diagnosis. We excluded papers from further review if subjects were not postmenopausal women, subjects had pre-existing dementia,

the study did not address links in the analytic framework, or the study was a review, a letter, or an editorial. Only articles published after peer review were included (no abstracts).

In order to identify the most important studies for inclusion in the evidence tables, a "best evidence" approach was used.[15] To address the association between HRT and cognitive testing, we included only randomized, double-blind, placebo-controlled trials and cohort studies. Randomized controlled trials tended to include younger, perimenopausal women who used estrogen for short periods (months), while cohort studies included older populations followed for longer periods. Nonrandomized trials, case-control studies, and cross-sectional studies were excluded from the evidence tables on the grounds that they were too subject to bias. Other studies have found that users are healthier, have healthier lifestyles, and are more highly educated than nonusers.[16-20]

To address the association between HRT and dementia, only prospective cohort and case-control studies were included in the evidence table; there were no randomized controlled trials for this question. Cross-sectional studies were felt to be subject to multiple biases because women with dementia would be less likely to be given HRT. For a case-control study to be included, the study methodology had to provide details about how dementia was determined in the cases and excluded in the controls. If two studies were done on the same population, the more recently updated data were included.

Size of Literature Reviewed

We reviewed all 509 abstracts identified by the search (Appendix 3). We excluded 396 papers from further review because they focused on animals, studied only men, were in a foreign language (unless a key article), or did not address links in the analytic framework; 57 papers

were reviews, letters, or editorials. From the original search, 56 articles with primary data on the relationship between HRT and cognition in postmenopausal women without dementia were then identified. An additional 16 articles with primary data were identified from reference lists of relevant reviews.

Data Extraction and Synthesis

HRT and Cognitive Function

We reviewed 9 randomized controlled trials and 8 cohort studies of the effects of HRT on cognition as measured by formal cognitive testing. Abstracted data from the trials included the type of study and setting; a description of active treatment and placebo groups, including age and menopausal type; type of HRT and duration of use; whether subjects had menopausal symptoms; and the effects of HRT on symptoms in the treatment groups (to look for unblinding of the study). We recorded study design information such as exclusion/inclusion criteria, method of allocation, and compliance and follow-up rates. In both the randomized controlled trials and cohort studies, we recorded the cognitive tests that were used and the results in users and nonusers. Not all studies used the same types of analyses. We recorded analyses based on either 1) comparisons of the change scores (post-pre) of users with those of nonusers, or 2) if the pretest scores were equal or appropriate adjustments were made, comparisons of the post scores of users with those of nonusers. If between-group comparisons were not available, within-group comparisons were documented. The most adjusted values were recorded. Any trends in duration, currency, or dosage were also noted.

Although the original goal was to quantitatively combine the results of the cognitive tests, the studies were too dissimilar. Instead, the cognitive tests were qualitatively combined

according to what they measured (memory, attention, reasoning, mental status, motor speed, verbal function) using a reference guide [21] and expert opinion. When possible, we calculated the increment of effect and normalized the scores using a total of 100 points. In some cases, scores were estimated from figures.[22-24]

Two reviewers independently rated the quality of the randomized controlled trials using Jadad scores and had 100% agreement (Appendix 4a).[25] For the other studies, we applied quality criteria developed by the methods workgroup of the U.S. Preventive Services Task Force and had 80% agreement (Appendix 4b). [26] A final score was reached through consensus if reviewers disagreed. Quality scores are detailed in the evidence tables (Appendix 6).

HRT and Dementia

We reviewed 10 case-control and 2 cohort studies of the relationship between HRT and risk of dementia of any type. Abstracted data included the type of study and setting; a description of the cases and controls or the users and nonusers, including age, menopausal status (surgical or natural), and education; the type and amount of HRT (formulation, duration, and recency); and any confounders that were controlled. We recorded how the history of HRT exposure was obtained because of the potential for recall bias and proxy bias. Demented women would be less likely than controls to remember previous exposures. Although proxy respondents were used in several of the studies, they may not have accurately remembered exposure history or may not have been aware of hormone use because many women consider this a personal decision. The cohort studies, because they document HRT use prior to the development of Alzheimer disease, are less prone to these biases. However, since Alzheimer disease is an

insidious disease with a long latency period,[8] a long followup time is needed to avoid finding a falsely low HRT usage rate in women with early cognitive decline.

We also abstracted how the investigators documented dementia in their cases and excluded it in their controls, and whether they used the criteria created by the work group of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer Disease and Related Disorders Association (NINCDS-ADRDA). This is the most widely applied criterion for defining Alzheimer disease clinically and requires a medical history; neurologic, psychiatric, and clinical exam; neuropsychological testing; and laboratory studies. Subjects are categorized as having possible, probable, or definite Alzheimer disease. Definite Alzheimer disease can be diagnosed only by the presence of neurofibrillary tangles and senile plaques at autopsy. A subject with a typical, insidious onset of dementia with progression will be given a diagnosis of probable Alzheimer disease once other systemic or brain diseases such as Parkinson disease and multi-infarct dementia are excluded. A possible Alzheimer disease diagnosis is made if the course of dementia is somewhat uncharacteristic or if the subject has other diseases that may be playing a minor role.[27] The inter-rater agreement for these criteria ranges from poor to good (Kappa statistic 0.36 to 0.65).[28, 29] Using pathological diagnosis of Alzheimer disease as the gold standard, a multi-site reliability and validity study found that the NINCDS-ADRDA criteria had a sensitivity of 0.81 and a specificity of 0.73.[29] Seventy percent of the errors in this study were false-negatives.

The recorded outcome measures were adjusted odds ratios or relative risks with 95 percent confidence intervals. When confidence intervals were not given, they were calculated using the available data from the original paper. In addition, any duration or recency of use data

were recorded. Studies were evaluated for quality using U.S. Preventive Services Task Force criteria as described previously.

We performed a meta-analysis of the 2 cohort[30, 31] and 10 case-control studies[32-41] meeting inclusion criteria. The two cohort studies reported adjusted relative risks. Three of the case-control studies provided raw data to calculate the unadjusted odds ratio but did not provide odds ratios from multivariate models.[32-34] The rest of the case-control studies reported adjusted odds ratios from logistic regression models. Because dementia is a relatively rare event, the odds ratio is a good estimate for the relative risk. For uniformity, we indicated the results from all studies as relative risks (RR).

The logarithm of the relative risk (logRR) was assumed to have a normal distribution. Standard errors for logRR were calculated from reported confidence intervals or p-values. If neither were reported, standard errors were calculated from the raw data. The logRR and the corresponding standard errors were the data points for the meta-analysis.

We tested both fixed-effects and random-effects models. A fixed-effects model is fit on the data and assumes only one source of variability (the variability within studies). It also assumes that the patient populations across studies are sufficiently similar and that the results are suitable to pool together. A random-effects model assumes a second source of variability among studies. Variation among studies implies that each study potentially estimates different effects sizes. Random-effects models are more conservative in the sense that they allow for more variability in treatment effects.[42]

The Bayesian data analytic framework was used for the meta-analysis. Inference on the parameters was done via posterior probability distributions. The data were analyzed using

WinBUGS software,[43] which uses a method of Markov Chain Monte Carlo called Gibbs Sampling to simulate posterior probability distributions.

Noninformative prior probability distributions were used. Inference was made on 5,000 simulated draws (1,000 draws from 5 chains) from the posterior distribution after adequate convergence. Sensitivity analysis was performed using different prior distributions, combining only studies with similar methods (using just the cohort studies, the case-control studies, the studies that looked at Alzheimer disease only, and the studies that used NINCDS-ADRDA criteria), excluding poor quality studies and studies with potential proxy bias, and excluding a study with uncertain confidence intervals.

We also evaluated all studies for selection bias using funnel plots [44] and investigated the sensitivity of the analysis to possible missing studies due to publication bias by trim and fill[45, 46]. Some asymmetry was detected in the funnel plots, suggesting that study selection bias was possible although interpretation of the plots is subjective.

Chapter 3. Results

Effects on Cognition (Figure 1, Arrow 6a)

1. Does HRT improve or stabilize cognition?

The literature search identified 9 randomized controlled trials (Appendix 5, Evidence Table 1), and 8 cohort studies (Appendix 5, Evidence Table 2) that used formal testing to measure the effects of HRT on cognition of women without dementia.

The randomized controlled trials are dissimilar in several ways. Three of the studies used a cross-over design,[22, 47, 48] while the rest used separate experimental and placebo groups. The

mean age of the women in the studies ranged from 45 to 80 years. Three studies included women immediately after a total abdominal hysterectomy/bilateral salpingo-oophroectomy,[22, 49, 50] while 6 other studies included community volunteers with only a small percentage of women with surgical menopause.[23, 47, 48, 51, 52] Some of the studies included women with menopausal symptoms,[22, 23, 49, 52] while one specifically excluded symptomatic women.[50] More than 40 different cognitive tests were used as outcome measures in these studies, and 30 of these tests were used by only one investigator (Appendix 6). Only 7 tests were used by more than 2 studies, and even when tests were repeated by several investigators, administration was not uniform. Only 2 studies, both done by the same research group, used the identical estrogen formulation and dose. Duration of use ranged from 21 days to 6 months. When analyzing the data, some studies performed between-group comparisons while others looked at within-group changes. Because of these differences among studies, results were not combined quantitatively. Instead, we grouped tests according to the cognitive process they measured: memory, attention, concept formation and reasoning, motor speed, mental status, verbal function, and learning ability (Table 1).

Memory. Memory is the first process to be affected in Alzheimer disease; thus, memory tasks are often used to track dementia onset. Although cross-sectional studies suggest that HRT affects memory, especially verbal memory,[53-57] results from randomized controlled trials and cohort studies are conflicting.

In one study[51] of 18 women with menopausal symptoms, 9 women given 1.5 mg of piperazine oestrone for 6 months scored, on average, 5 points higher than baseline on a memory battery (the Guild Memory Test). The 9 women given placebo had a 1.8 point decrease in

scores. The overall increase in the estrogen group's score can be attributed to 3 women whose scores increased by 13 to 20 points while on estrogen. The other women in the estrogen arm had either minor improvements or even decreases in scores.

Six studies, 2 randomized controlled trials and 4 cohort studies, used 4 tests of immediate verbal recall. [49, 58-62] In these tests, subjects are shown verbal stimuli, such as lists of words, word pairs, or paragraphs, and asked to recall them immediately after seeing them.[21] One randomized controlled trial [49] and one randomized cross-over study,[58] both done by the same investigators, found that 10 women randomized to intramuscular estradiol for 3 months performed better on 2 tests of immediate verbal recall. The women receiving estrogen immediately recalled 3.5 to 5 more words of a paragraph compared to when they were not receiving estrogen. There was no improvement in the placebo groups.[58, 49] On another test, at the end of the study the placebo group recalled 5 fewer word pairs, but there was no change in the number of pairs recalled by women in the estrogen group.[49] However, the women in both of these studies were immediately post-operative for total abdominal hysterectomy/bilateral salpingo-oophorectomy and likely had menopausal symptoms, though these were measured in only one study. When the same research group used a cohort design to study elderly women (average age 73 to 74 years) who were unlikely to be symptomatic, there was no longer a significant difference in scores on these tests between long-term users and nonusers.[59]

The third test of immediate verbal recall, the selective reminding test, was used in the 3 cohort studies. In this test, subjects are asked to recall a list of words and with each successive trial are reminded of the words they omitted.[21] In one cohort study, 81 users recalled 1.58 more words and 646 nonusers recalled 0.41 fewer words after 2.5 years of follow-up.[60] However, in the Rancho Bernardo cohort, 394 ever-users of HRT did not perform better on this same test than

the 406 nonusers.[61] The younger women in the Rancho Bernardo cohort performed significantly better than older women, suggesting that this test was sensitive enough to detect differences between groups. In addition, the users in the Rancho Bernardo cohort had a longer duration of use than those in the previous study (16.5 years versus 2.5 years), and were followed for longer (15 years versus 2.5 years).[61] A third study[59] compared the immediate selective reminding test scores of 10 long-term users to 27 nonusers. Although the users had higher baseline scores, there was no difference in the amount the scores changed over 18 months.

"Delayed verbal recall" refers to recall of the same verbal stimuli after 15 to 60 minutes. Estrogen exposure was associated with improvement in at least one test of delayed verbal recall in 2 out of 4 studies. Two randomized controlled trials that found improvements in immediate paragraph recall did not find that these benefits extended to delayed recall of a paragraph.[22, 49] One of these trials found that while the placebo group recalled 2.3 fewer word pairs after the delay period, the average score of the estrogen group did not change.[49] However, this result was not confirmed by a recent cohort study by the same research group.[59]

Two of the 3 cohort studies that looked at the delayed portion of the selective reminding test found an effect of estrogen.[59-61] The studies found that after a 15-minute delay, users recalled 0.64 more words from a word list while nonusers recalled 0.41 fewer words. However, the scores on the other tests of delayed verbal recall that were used in these cohort studies did not differ between exposure groups.[59, 61]

On visual memory tests, subjects are shown configural stimuli (such as shapes and figures) and usually respond through visuomotor response (such as drawing). Six studies used 5 tests of visual recall. Eighteen women who became users of postmenopausal estrogen during the course of followup in the Baltimore Longitudinal Study of Aging had 0.28 fewer errors on the

Benton Visual Retention test after 6.5 years of followup while never users had 1.67 more errors.[24] However, 2 trials did not find that women given estrogen for 3 months did better than those given placebo on this same test.[47, 52] No study found that women exposed to estrogen did better on 3 other tests of visual recall.[49, 59, 61]

Attention. Although performance on tests of attention is not as useful as performance on memory tests in predicting Alzheimer disease, 10 studies[23, 47-50, 52, 58, 59, 61, 63] have looked at the effects of HRT on various aspects of attention. The most recent randomized controlled trial [48] used nonstandardized tests to measure "working memory," which refers to the "ability to hold information in mind, to initialize information, and to use information to guide behavior without the aid of external cues." [21] Although this study did not find a difference in scores according to estrogen exposure, the ease of the tests may have precluded finding a difference (ceiling effect). The study did find that the women given estrogen had increased activation of several areas of the brain on functional magnetic resonance imaging when performing the working memory tasks.[48] Two other studies used tests to try to evaluate working memory; neither study found a difference in scores according to estrogen exposure.[47, 64] It is unlikely that the null results are secondary to a ceiling effect, since the results in at least one of these studies were measured in milliseconds.[47]

Complex attention tasks require both attention and concentration. Performance on 2 standardized tests of complex attention, the digit symbol and trail making tests, was not affected by estrogen in any of the 5 studies that used them as cognitive measures.[47, 50, 52, 61, 63] However, one of these trials found that nuns given estrogen had borderline improvement on another nonstandardized test of attention, the spot-pattern test ($p=0.08$).[52] Another study found

that estrogen-treated women did better on a Swedish test of attention.[23] Both of these latter two studies found that women treated with estrogen had more improvement in symptoms than the untreated subjects. In fact, the latter study states that "there was a remarkable improvement in the ability to sleep in all oestrogen-treated patients,"[23] and so the subjects on estrogen may have done better because they were less fatigued.

Women using estrogen showed improvement compared to the placebo group on 2 of the 13 tests of mental tracking, which includes such tasks as arithmetic, repeating numbers backwards, and repeating months backwards.[23, 47, 52, 61] Although one randomized controlled trial found that estrogen exposure was related to performance on digit span (a test requiring subjects to repeat a list of 2 to 8 numbers backwards),[22] the women randomized to estrogen in 3 other randomized controlled trials did not perform better on this test.[47, 50, 52] Two of these studies[50, 52] used women who were post-operative from total abdominal hysterectomy/bilateral salpingo-oophorectomy; however, the women had to have fewer than 4 hot flashes in a 2-week period to qualify for one of them.[50] The 24 women randomized to oral conjugated equine estrogen (0.625 and 1.25 mg) for 3 months did not perform better on the test of digit span. [50] However, all of the pre-treatment scores being in the normal range may have precluded finding a difference (ceiling effect).

A cross-over study of 62 women also did not find that women given estrogen for 3 months did better on the digit span test.[47] Since this study was the largest of the randomized trials and reported a power of 90 percent, it is likely that the women in the other study[50] may have done better on the digit span test secondary to improvement in menopausal symptoms. Indeed, when this research group used a cohort design to look at asymptomatic women, there

was no difference between users' and nonusers' performances on another test of mental tracking, the visual memory span.[59]

The results on yet another test of mental tracking, the Stroop color word test, are also conflicting. A randomized controlled trial of 21 symptomatic women in Germany found that while the placebo group had no change in scores, the estrogen users were better able to read the word of a color despite it being printed in a different color, or to state the color of the word even though the word stood for another color. [23] However, these results were not confirmed by the later, larger study.[47]

Of the 3 studies that measured vigilance, 2 found that estrogen improved women's abilities to sustain attention. In both of these studies, women were symptomatic with fatigue, sleep problems, hot flashes, and depression.[23, 52] In contrast, a larger randomized cross-over trial did not find that women given estrogen performed better on 2 sensitive tests of vigilance.[47]

Concept formation and reasoning. Concept formation and reasoning refers to the quality or process of thinking. It was tested in 3 studies with conflicting results. While a randomized cross-over study and a cohort study found that subjects given estrogen improved in their abstract reasoning scores compared to when they took placebo,[22, 62] a New York based cohort study did not find that scores for ever-users changed over 2.5 years compared to never-users.[60]

Motor speed. Although motor speed is not very predictive of Alzheimer disease because it is not affected until later stages of the disease, losses in motor speed will affect performance on any task that requires a motor response. For example, motor slowing can look like a reasoning deficit if the reasoning task requires any sort of motor response or manipulation. Motor speed, as

measured by simple reaction time, was improved by postmenopausal estrogen in one study[23] but not another.[47] A randomized controlled trial that found estrogen improved reaction time by over 100 milliseconds included symptomatic women.[23] A larger, more recent trial did not find a difference in reaction time between exposure groups, even though measurements were also in milliseconds.[47] Women given estrogen had improvement in clerical speed and accuracy in another study.[22]

Dementia screening measures. Three early studies did not find that women given estrogen had improvement in mental status as measured by the mini-mental status exam or Cognitive Capacity Screening Exam.[61, 63, 65] However, this is not unexpected, given that the simplicity of the mental status exam for cognitively intact women might preclude finding subtle differences (ceiling effect). Two recent cohort studies using more sensitive multidimensional tests of cognition found that over time users performed better than nonusers.[62, 66]

Verbal function. Only one[62] of 4 studies found that users performed better than nonusers on tests of verbal functions and language skills.[59-62] Users of unopposed estrogen were more fluent in naming categories.[62]

Influence of symptoms. In the randomized controlled trials, changes in cognitive measures were most likely to occur in symptomatic women (Table 2). In all 4 trials of women with somatic complaints, HRT improved at least one cognitive function. The cognitive processes that were most consistently improved were verbal recall[22, 49, 59, 60] and vigilance,[23,52] although

complex attention,[23, 52] mental tracking,[22, 52] concept formation and reasoning,[22] and motor speed[22, 23] were also affected.

Symptomatic women may perform better on cognitive testing when given estrogen because they sleep better, are less fatigued, have fewer hot flashes, or have improved mood. Many of the effects were in tests of attention, which are particularly sensitive to anything that elevates mood. Alternatively, the subtle effects of estrogen on cognition may be apparent only in subjects who are not performing at maximum cognitive ability because of fatigue and loss of sleep secondary to menopausal symptoms.

One study that looked at asymptomatic women did not find that women given estrogen had improved performance on tests of immediate verbal recall or attention.[50] The largest study, a cross-over study of 62 women using transdermal estrogen for 3 months, also did not find any improvement in women exposed to estrogen on tests of immediate verbal recall, visual memory, attention (including working memory), or motor speed.[47] This study used sensitive tests (outcomes measured in milliseconds) and had a power of 90 percent to detect a difference between users and nonusers.

Effects of progestins. All of the randomized controlled trials used unopposed estrogen. The 4 cohort studies that looked at the type of HRT found that most (greater than 70 percent) of the women used unopposed oral conjugated equine estrogen.[59-61, 65] None of these studies looked at subgroups that used progestins. One small, nonrandomized trial of 19 symptomatic women used estradiol combined with progestins. They concluded that progestins did not attenuate the cognitive benefits of estrogen because users of the combined regimen had more improvement on a test of delayed verbal recall than nonusers. However, unlike the results in several previous

randomized controlled trials of symptomatic women, immediate verbal recall was not enhanced by HRT exposure.[67]

One recent cohort study looked separately at users of unopposed and combined HRT regimens. Women currently using unopposed estrogen had more improvement in global cognition, abstract reasoning, and category fluency, compared with never-users. However, current users of estrogen combined with progestin had a decline in global cognition and mental tracking scores.[62]

2. What is the optimal dose and duration of use?

Most of the women in the cohort studies used oral estrogen, but dosages were not given. Therefore, information on dosing comes from randomized controlled trials that used a variety of preparations and doses. Only 3 of the randomized controlled trials used oral conjugated equine estrogen in doses commonly prescribed in the U.S. (either 0.625 or 1.25 milligrams), and none found that women randomized to estrogen performed better on several tests of cognition.[48, 64, 50] A study using transdermal estrogen also did not find a difference in cognitive test scores between the estrogen and placebo groups 47. Although the 2 studies that used intramuscular estradiol found that estrogen favorably affected women's performances on cognitive testing, these studies were done by the same author and enrolled symptomatic women.22, 49 Some of the early studies that found beneficial effects on cognition used larger doses of oral estrogen than are currently prescribed.[51, 52] Although it is tempting to conclude that the different estrogen formulations or dosages may have contributed to the studies' disparate findings, too many other factors vary among studies to draw any conclusions about which formulation or dose may be most protective.

The trials lasted only several months, and so information about duration of use comes from the cohort studies. The one cohort study that looked at duration found that users of greater than 20 years scored 1 point higher on category fluency, a test of verbal functions and language skills.[61] However, these long-term users did not perform better on any of the other 8 measures of cognition. Although one cohort study looking at recency of use actually found that past users had more benefit than current users, [63] another study found that past users had scores intermediate between current and never-users.[62]

Summary

- It is difficult to compare studies about HRT and cognitive function and report overall conclusions because the studies enlist different patient populations and report different cognitive test outcomes.
- Only studies that used women with menopausal symptoms found that estrogen improved cognitive performance. The most consistent findings in these studies appeared to be on verbal memory and vigilance tests, although effects on complex attention, mental tracking, concept formation and reasoning, and motor speed were also seen.
- Postmenopausal estrogen does not seem to enhance asymptomatic women's performance on cognitive testing.
- There is insufficient evidence about the effects of the addition of progestins to estrogen. One non-randomized trial of a small number of women found that progestins may attenuate some of estrogen's effects on immediate verbal recall in symptomatic women.

- The randomized controlled trials are too dissimilar in design to conclude that any formulations or dosage may be more beneficial for cognitive function in symptomatic women.
- Only one cohort study looked at duration of use, and it did not find that long-term users were performing consistently better over time than never-users.

Effects on Dementia (Figure 2, Arrow 6b)

1. Does HRT lower the risk for dementia?

Ten case-control studies (Appendix 5, Evidence Table 3) and 2 cohort studies (Appendix 5, Evidence Table 4) on the association between HRT use and risk of dementia were identified from the literature review and met inclusion criteria.

The early case-control studies did not find an association between HRT and dementia (Odds Ratios [ORs] of 0.78 to 2.38) and all received poor quality scores.[32, 33, 35, 36] HRT use was only one of many risk factors evaluated in these studies. These early studies all used proxy interviews to determine exposure to HRT in both cases and controls, and none used blinded interviewers. Two of the studies evaluated the agreement in reported HRT use between controls and their surrogates and found good agreement (Kappa values were 0.63 and 0.64).[33, 35] The studies were relatively small, with the number of users ranging from 8 to 21, or about 8 to 18 percent of the study population. All of the studies controlled for age, but only one study controlled for education.[33]

All but one of the case-control studies published since 1990 found a significantly decreased risk of dementia among HRT users (ORs of 0.33 to 1.1).[34, 37-41] In some of these studies, however, the method of determining estrogen exposure may have resulted in a falsely

low rate of HRT use in cases. For example, several of the studies used proxy interviews for cases but self-interviews for controls.[34, 38, 41] Because proxy informants would be expected to have less knowledge about HRT use, the rate of estrogen exposure in cases may have been underestimated. These studies were rated poor.

The largest case-control study, which was nested in the Leisure World Cohort and also rated poor, used data from death certificates to determine dementia outcomes, and defined HRT exposure with a self-administered questionnaire completed an average 5 years before death. Because dementia is insidious in onset and women with cognitive decline were not excluded at baseline, cases might have been less likely than controls to remember previous estrogen use at the time of the original questionnaire. This might have falsely lowered the percentage of users among demented subjects and led to the finding of decreased risk among HRT users (OR 0.65; 95% CI 0.29, 0.88).[39]

Three studies used more objective measures such as medical or pharmacy records to determine estrogen use, but their results are conflicting. One poor-quality study stated that "medical records were the primary source material,"[38] but did not specify how the material was abstracted or how much of the information on HRT use was actually derived from proxy interviews. They found that HRT users had a 45 percent reduction in risk of dementia (CI 0.26, 1.16).

A good-quality study enlisted subjects from the Group Health Cooperative of Puget Sound, a health maintenance organization (HMO) in Seattle, Washington.[37] Cases were identified from the Alzheimer Disease Patient Registry, which uses NINCDS-ADRDA criteria to diagnose Alzheimer disease. The HMO's computerized pharmacy records were then used to identify subjects who had filled a prescription for any form of postmenopausal estrogen since

1977. Proxy interviews were used for information prior to 1977. Almost half of both groups used HRT, which is a higher percentage of users than in the general U.S. population and suggests a highly selected study population. Also, because HRT use was defined through prescription data, women who never took or discontinued the medication within days after first filling the prescription would be classified as users. Such misclassification could have biased the results to the null. After controlling for age and history of hysterectomy before and after age 55, this study found no decreased risk among ever-users of HRT (OR 1.1; 95% CI 0.6, 1.8).

The Rochester Epidemiology Project records-linkage system was used to identify cases and controls in a fair-quality study.[40] Dementia diagnosis and HRT use were determined through blinded record abstraction. After controlling for age, education, and length of time in the linkage system, this study found that users of any form of estrogen for greater than 6 months after the menopause but before the onset of Alzheimer disease was associated with a 58 percent reduction (OR 0.42; 95% CI 0.18, 0.96).

The strongest evidence for an association between HRT use and Alzheimer disease comes from 2 fair-quality cohort studies (Appendix 5, Evidence Table 4). The Manhattan Study of Aging cohort was formed from Medicare listings and senior housing centers.[30] The average age of subjects was 74.2 years. One hundred fifty-six ever-users of postmenopausal estrogen and 968 nonusers were followed for 1 to 5 years for the development of Alzheimer disease as defined by the NINCDS-ADRDA criteria. After controlling for education, age, and ethnicity, the study found that users were significantly less likely to develop Alzheimer disease (Relative Risk (RR) 0.5; 95% CI 0.25, 0.9). Users who developed Alzheimer disease also had a later age of onset. One problem with this study, however, was that subjects developed Alzheimer disease within 5 years of the initial interview. Given that a diagnosis of Alzheimer disease lags symptom onset

by 3.5 to 5.5 years,[8] cases may have been less likely to remember HRT usage. Also, concerns about compliance in women with mild cognitive problems could have made it less likely that they were prescribed HRT.

The Baltimore Longitudinal Study of Aging followed 230 HRT users and 242 nonusers aged 28 to 94 (average 61.5) for up to 16 years, which makes it less likely that the Alzheimer disease subjects had subtle memory problems at the beginning of the study.[31] They evaluated the subjects every 2 years for the development of Alzheimer disease as diagnosed by the NINCDS-ADRDA criteria. After controlling for age and education, they found that relative risk of dementia in users was 0.457, with a 95% confidence interval of 0.209-0.997.

The results of these 10 case-control and 2 cohort studies were combined by meta-analysis (Figure 3). The test of homogeneity indicated that the studies were homogeneous ($p>0.10$). The fixed and random effects models did not differ, so we present only the random effects model. When the studies were combined quantitatively, the summary odds ratio was 0.66 (CI 0.53, 0.82). When case-control or cohort studies were analyzed separately, the estimates did not change (Table 3). Excluding poor-quality studies or studies with potential proxy bias did not changes these estimates. Also, restricting the analysis to studies that looked only at Alzheimer disease or used only NINCDS criteria did not change the estimate, although the confidence intervals were wider given the smaller number of studies. Sensitivity analysis using different prior distributions and using various values for confidence intervals also did not significantly change the risk estimates.

Although the summary odds ratio indicates a decreased risk of dementia in women exposed to postmenopausal estrogen, confounders may explain this relationship. Women who use HRT are more educated,[16, 17, 19] and formal education has been found to be protective

against dementia.[3, 7] However, several of the studies controlled for education and still found a decreased risk. 34, 40, 41HRT users are also healthier and have healthier lifestyles, and physical health status has been associated with cognitive changes with advancing age.[2] Users are also younger than nonusers, and the most important risk factor for dementia is advancing age. All of the studies controlled for age.

Only one study using pharmacy records looked at the effect of progestins. Although the number of progestin users was small, adding progestins to the logistic regression model did not change the risk estimates, indicating that it was not a significant confounder.[37] Another study did not find that excluding women who reported a surgical menopause (who are usually taking unopposed estrogen) affected the results.[31]

Only one study looked at dementia other than Alzheimer disease.[38] This study included women with dementia secondary to ischemic vascular disease. Women who used HRT had a 50 percent decreased risk of developing ischemic vascular disease, although this result was not statistically significant (OR 0.50; CI 0.26, 1.20).

2. What is the optimal dose and duration of use?

The older case-control studies do not specify the formulation of HRT used.[32, 33, 35, 36] The more recent case-control studies define HRT exposure as the use of any form (oral, IM, topical, suppository) of estrogen after menopause.[34, 37, 39, 40] The cohort studies included only women who used oral or transdermal forms of estrogen.[30, 31] In the studies that looked at this information, 66 to 95 percent of the women used oral conjugated equine estrogen.[34, 37, 40] While one study found that only oral estrogen was associated with a decreased risk of Alzheimer

disease,[37] another study found that oral, injectable, and/or cream forms of estrogen were all associated with a decreased risk.[39]

Although one study found that taking at least 1.25 mg of oral conjugated estrogen was associated with greater dementia risk reduction,[39] another study did not find that cumulative dose was associated with dementia risk.[40]

The results for duration of use were also mixed. In the Manhattan Study of Aging, users with more than one year of use had a relative risk of 0.13 (CI 0.02, 0.92) compared to a relative risk of 0.47 (CI 0.20, 1.10) in users for less than one year.[30] In another study only users for more than 6 months had a decreased risk.[40] While one case-control study[68] and one cohort study[30] found that increasing duration of use was associated with a decreased risk of dementia, another case-control study and cohort study did not confirm this finding.[31, 37].

Studies that looked at currency of use also found different results. Two of the case-control studies that looked only at current users found a decreased risk,[34, 38] while a third did not.[33] Another study found that the odds ratio for the risk of dementia in current users was 0.6 (CI 0.3, 1.2) compared to 1.7 (CI 0.9, 3.2) in former users.[37] However, women with dementia may be less likely to be prescribed HRT because of compliance issues or because of complex medication regimens (prescribing bias).

Summary

- Based on data from 12 case-control and cohort studies, there appears to be a 34 percent decreased risk of dementia among HRT users (OR 0.66, CI 0.53, 0.82).
- The studies upon which this risk estimate is based have several limitations.

  -In case-control studies, proxy respondents may not be aware of previous HRT use.

-In case-control studies, demented women may have been less likely to receive HRT because of compliance issues and multiple medications.

-In cohort studies, women with early memory problems may not have remembered previous HRT use.

-Women who use HRT are healthier ("healthy user bias").

- There are insufficient data about the addition of progestins.
- There are insufficient data about other forms of dementia.
- There are insufficient and conflicting data about a dose-response or duration effect.
- Although some studies have found that current users had a decreased risk, this could be secondary to prescribing bias.

Chapter 4. Discussion

Conclusions

Table 4 represents a summary of evidence for the questions addressed in this review. Although the study populations and outcome measures differ, the 9 RCTs and 8 cohort studies offer some provisional conclusions about the effects of postmenopausal HRT on cognition. No deleterious effects on cognition have been reported. HRT does not appear to enhance performance on formal cognitive testing for asymptomatic women. However, some studies have found that symptomatic women have improved cognitive performance with HRT, especially in tests of verbal memory and vigilance. There is insufficient evidence about whether progestins attenuate these cognitive effects. Because of the heterogeneous study designs, no conclusions

can be drawn about whether specific estrogen formulations or dosages might be more beneficial. Duration of use did not appear to be related to cognitive performance.

Ten case-control and 2 cohort studies suggest that HRT users have a 34 percent decreased risk of Alzheimer disease (CI 0.53, 0.82), however, the studies included in the meta-analysis have important methodological limitations. It is unclear whether estrogen is also associated with a decreased risk of other forms of dementia. No conclusions can be drawn about the effects of adding progestins to the regimen, or whether specific dosages or formulations of estrogen are more protective.

Limitations of the Literature

Studies of HRT use and cognition and dementia have many limitations. Risk estimates may have been falsely low in some of the case-control studies that used proxy respondents, who may not be aware of previous HRT exposure. Also, demented women may have been less likely to receive HRT because of compliance issues or because they were already receiving multiple medications. The relative risk estimates in the cohort studies might have been artificially decreased if women with early, subtle memory changes were less likely to remember previous HRT use. Finally, HRT users may be less likely to develop Alzheimer disease not because of postmenopausal estrogen exposure, but because they are healthier and more educated.

Future Research

Since women have menopausal symptoms for only a limited time, future research on the effects of HRT on cognitive performance should focus on older, asymptomatic women instead of perimenopausal women. Because HRT users are different from nonusers in many lifestyle and

health behaviors, studies should control for health status, health behavior, and education. The ideal studies would be large double-blind placebo-controlled trials with intervention arms containing estrogen with and without progestins. The trial should last for at least several years so that the effects of long-term HRT can be studied, and that subtle changes between treatment groups could be detected. Progestins should be included as part of the intervention arm because of the possibility that it may attenuate some of estrogen's cognitive effects. It should also include the various formulations and dosages of estrogen that are commonly used. Future studies also need to control for the psychological effects of estrogen to ensure that any cognitive effects are not secondary to changes in depressive symptoms. Studies should control for education.

Of crucial importance in designing a study of cognition is deciding which outcome measures to employ. The tests should not have ceiling values and need to be sensitive to very small differences because the effects of estrogen on cognition may be subtle. These tests should examine particular cognitive domains because the evidence indicates that estrogen may have neural and cognitive specificity. In particular, future studies should focus on the cognitive domains most consistently affected in previous studies, such as verbal memory, vigilance, complex attention, mental tracking, concept formation and reasoning, and motor speed. Future studies should include measures of the ability to care for oneself, live independently, and complete activities of daily living

Estrogen's cognitive and neural specificity should also be considered when interpreting the results of future research studies, including the two ongoing primary prevention trials of HRT and cognition, the Women's Health Initiative Study of Cognitive Aging (WHISCA) [69] and the Women's International Study of Long Duration Oestrogen after Menopause in the United

Kingdom.[70] The results of these trials are still several years away but may answer many of the questions raised by observational data.

References

1. LeBlanc E, Janowsky J, Chan B, Nelson H. Hormone replacement therapy and cognition: Systematic review and meta-analysis. *JAMA*. 2001;285:1489-99.
2. Teri L, McCurry S, Logsdon R. Memory, thinking, and aging. What we know about what we know. *West J Med*. 1997;4:269-75.
3. Keefover R. The clinical epidemiology of Alzheimer's disease. *Neurol Clin*. 1996;2:337-51.
4. Burke JR, Morgenlander JC. Update on Alzheimer's disease. Promising advances in detection and treatment. *Postgrad Med*. 1999;106:85-6, 89-90, 93-4 passim.
5. Evans DA, Funkenstein HH, Albert MS, et al. Prevalence of Alzheimer's disease in a community population of older persons. Higher than previously reported. *JAMA*. 1989;262:2551-6.
6. Jorm AF, Korten AE, Henderson AS. The prevalence of dementia: a quantitative integration of the literature . *Acta Psychiatr Scand*. 1987;76:465-79.
7. Breteler MM, Claus JJ, van Duijn CM, Launer LJ, Hofman A. Epidemiology of Alzheimer's disease. *Epidemiol Rev*. 1992;14:59-82.
8. Ernst RL, Hay JW. The US economic and social costs of Alzheimer's disease revisited. *Am J Public Health*. 1994;84:1261-4.
9. Smith CA, McCleary CA, Murdock GA, et al. Lifelong estrogen exposure and cognitive performance in elderly women. *Brain Cogn*. 1999;39:203-18.
10. Chiarello C, McMahon MA, Schaefer K. Visual cerebral lateralization over phases of the menstral cycle: A preliminary investigation. *Brain Cogn*. 1989;11:18-36.
11. Hampson E, Kimura D. Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. *Behav Neurosci*. 1988;102:456-459.
12. Hampson E. Estrogen-related variation in human spatial and articulatory-motor skills. *Psychoneuroendocronology*. 1990;15:97-111.
13. Yaffe K, Browner W, Cauley J, Launer L, Harris T. Association between bone mineral density and cognitive decline in older women. *J Am Geriatr Soc*. 1999;47:1176-82.
14. US Preventive Services Task Force. *Guide to Clinical Preventive Services. Report of the U.S. Preventive Services Task Force,* 2nd ed. Baltimore: Williams & Wilkins; 1996.

15. Slavin RE. Best practice synthesis: An alternative to meta-analytic and traditional reviews. *Educ Res*. 1986;15:5-11.
16. Keating N, Cleary P, Rossi A, Zaslavsky A, Ayanian J. Use of hormone replacement therapy by postmenopausal women in the United States. *Ann Intern Med*. 1999;130.
17. Matthews KA, Kuller LH, Wing RR, Meilahn EN, Plantinga P. Prior to use of estrogen replacement therapy, are users healthier than nonusers? *Am J Epidemiol*. 1996;143:971-8.
18. Folsom AR, Mink PJ, Sellers TA, Hong CP, Zheng W, Potter JD. Hormonal replacement therapy and morbidity and mortality in a prospective study of postmenopausal women. *Am J Public Health*. 1995;85:1128-32.
19. Derby CA, Hume AL, Barbour MM, McPhillips JB, Lasater TM, Carleton RA. Correlates of postmenopausal estrogen use and trends through the 1980s in two southeastern New England communities. *Am J Epidemiol*. 1993;137:1125-35.
20. Barrett-Connor E. Postmenopausal estrogen and prevention bias. *Ann Intern Med*. 1991;115:455-6.
21. Lezak M. *Neuropsychological Assessement*. 3rd ed. New York: Oxford University Press; 1995.
22. Sherwin BB. Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. *Psychoneuroendocrinology*. 1988;13:345-57.
23. Fedor-Freybergh P. The influence of oestrogens on the wellbeing and mental performance in climacteric and postmenopausal women. *Acta Obstet Gynecol Scand Suppl*. 1977;64:1-91.
24. Resnick SM, Metter EJ, Zonderman AB. Estrogen replacement therapy and longitudinal decline in visual memory. A possible protective effect? *Neurology*. 1997;49:1491-7.
25. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17:1-12.
26. Harris R, Helfand M, Woolf S, et al. Current methods of the Third US Preventive Services Task Force: A review of the process. *Am J Prev Med*. 2001;20:21-35.
27. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34:939-44.

28. Lopez OL, Swihart AA, Becker JT, et al. Reliability of NINCDS-ADRDA clinical criteria for the diagnosis of Alzheimer's disease. *Neurology*. 1990;40:1517-22.
29. Blacker D, Albert MS, Bassett SS, Go RC, Harrell LE, Folstein MF. Reliability and validity of NINCDS-ADRDA criteria for Alzheimer's disease. The National Institute of Mental Health Genetics Initiative. *Arch Neurol*. 1994;51:1198-204.
30. Tang MX, Jacobs D, Stern Y, et al. Effect of oestrogen during menopause on risk and age at onset of Alzheimer's disease. *Lancet*. 1996;348:429-32.
31. Kawas C, Resnick S, Morrison A, et al. A prospective study of estrogen replacement therapy and the risk of developing Alzheimer's disease: the Baltimore Longitudinal Study of Aging [published erratum appears in Neurology 1998 Aug;51(2):654]. *Neurology*. 1997;48:1517-21.
32. Amaducci LA, Fratiglioni L, Rocca WA, et al. Risk factors for clinically diagnosed Alzheimer's disease: a case-control study of an Italian population. *Neurology*. 1986;36:922-31.
33. Heyman A, Wilkinson WE, Stafford JA, Helms MJ, Sigmon AH, Weinberg T. Alzheimer's disease: a study of epidemiological aspects. *Ann Neurol*. 1984;15:335-41.
34. Henderson VW, Paganini-Hill A, Emanuel CK, Dunn ME, Buckwalter JG. Estrogen replacement therapy in older women. Comparisons between Alzheimer's disease cases and nondemented control subjects. *Arch Neurol*. 1994;51:896-900.
35. Graves AB, White E, Koepsell TD, et al. A case-control study of Alzheimer's disease. *Ann Neurol*. 1990;28:766-74.
36. Broe GA, Henderson AS, Creasey H, et al. A case-control study of Alzheimer's disease in Australia. *Neurology*. 1990;40:1698-707.
37. Brenner DE, Kukull WA, Stergachis A, et al. Postmenopausal estrogen replacement therapy and the risk of Alzheimer's disease: a population-based case-control study. *Am J Epidemiol*. 1994;140:262-7.
38. Mortel KF, Meyer JS. Lack of postmenopausal estrogen replacement therapy and the risk of dementia. *J Neuropsychiatry Clin Neurosci*. 1995;7:334-7.
39. Paganini-Hill A, Henderson VW. Estrogen replacement therapy and risk of Alzheimer disease. *Arch Intern Med*. 1996;156:2213-7.

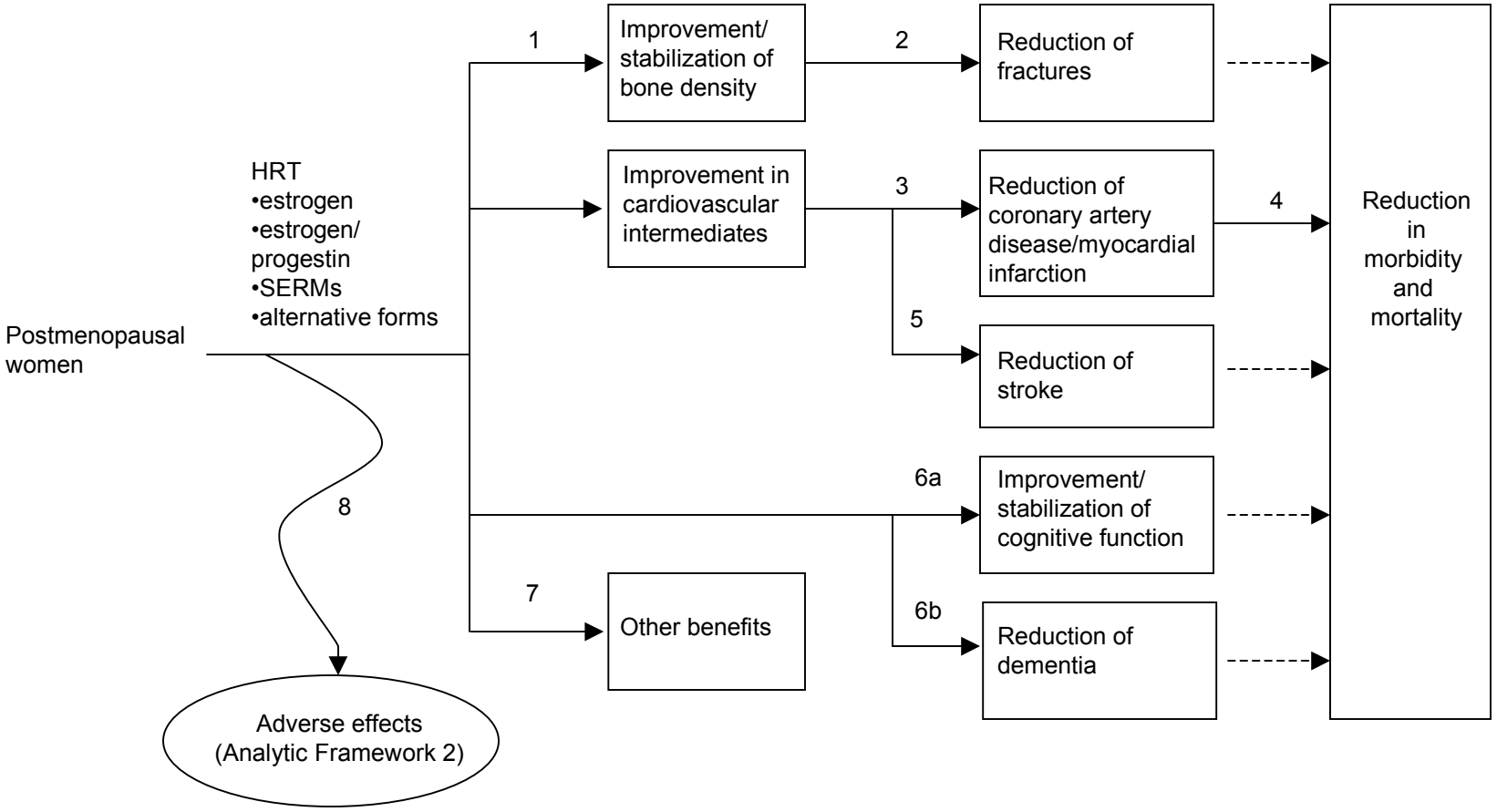
40. Waring SC, Rocca WA, Petersen RC, O'Brien PC, Tangalos EG, Kokmen E. Postmenopausal estrogen replacement therapy and risk of AD: a population-based study. *Neurology*. 1999;52:965-70.
41. Harwood DG, Barker WW, Loewenstein DA, et al. A cross-ethnic analysis of risk factors for AD in white Hispanics and white non-Hispanics. *Neurology*. 1999;52:551-6.
42. Normand SL. Meta-analysis: formulating, evaluating, combining, and reporting. *Stat Med*. 1999;18:321-59.
43. Spiegelhalter D, Thomas A, Best N. WinBUGS Version 1.2 User Manual. 1.2 ed. Cambridge: MRC Biostatistics Unit; 1999.
44. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J*. 1997;315:629-34.
45. Duval S, Tweedie R. A Nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *J Am Stat Assoc*. 2000;95:89-98.
46. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for Meta-analysis in Medical Research*. Chichester: John Wiley and Sons; 2000.
47. 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. *Obstet Gynecol*. 1998;91:459-66.
48. Shaywitz SE, Shaywitz BA, Pugh KR, et al. Effect of estrogen on brain activation patterns in postmenopausal women during working memory tasks. *JAMA*. 1999;281:1197-202.
49. Phillips SM, Sherwin BB. Effects of estrogen on memory function in surgically menopausal women. *Psychoneuroendocrinology*. 1992;17:485-95.
50. Ditkoff EC, Crary WG, Cristo M, Lobo RA. Estrogen improves psychological function in asymptomatic postmenopausal women. *Obstet Gynecol*. 1991;78:991-5.
51. Hackman BW, Galbraith D. Replacement therapy and piperazine oestrone sulphate ('Harmogen') and its effect on memory. *Curr Med Res & Opin*. 1976;4:303-6.
52. Vanhulle G, Demol P. A double-blind study into the influence of estriol on a number of psychological tests in post-menopausal women. In: van Keep PA, et al., ed. *Consensus on MenopauseResearch*. Lancaster, England: MTP; 1976.

53. Szklo M, Cerhan J, Diez-Roux AV, et al. Estrogen replacement therapy and cognitive functioning in the Atherosclerosis Risk in Communities (ARIC) Study. *Am J Epidemiol*. 1996;144:1048-57.
54. Resnick SM, Maki PM, Golski S, Kraut MA, Zonderman AB. Effects of estrogen replacement therapy on PET cerebral blood flow and neuropsychological performance. *Horm Behav*. 1998;34:171-82.
55. Kimura D. Estrogen replacement therapy may protect against intellectual decline in postmenopausal women. *Horm Behav*. 1995;29:312-21.
56. Robinson D, Friedman L, Marcus R, Tinklenberg J, Yesavage J. Estrogen replacement therapy and memory in older women. *J Am Geriatr Soc*. 1994;42:919-22.
57. Kampen DL, Sherwin BB. Estrogen use and verbal memory in healthy postmenopausal women. *Obstet Gynecol*. 1994;83:979-83.
58. Sherwin BB. Cognitive assessment for postmenopausal women and general assessment of their mental health. *Psychopharmacol Bull*. 1998;34:323-6.
59. Carlson LE, Sherwin BB. Relationships among cortisol (CRT), dehydroepiandrosterone-sulfate (DHEAS), and memory in a longitudinal study of healthy elderly men and women. *Neurobiol Aging*. 1999;20:315-24.
60. Jacobs DM, Tang MX, Stern Y, et al. Cognitive function in nondemented older women who took estrogen after menopause. *Neurology*. 1998;50:368-73.
61. Barrett-Connor E, Kritz-Silverstein D. Estrogen replacement therapy and cognitive function in older women. *JAMA*. 1993;269:2637-41.
62. Rice MM, Graves AB, McCurry SM, et al. Postmenopausal estrogen and estrogen-progestin use and 2-year rate of cognitive change in a cohort of older Japanese American women: The Kame Project. *Arch Intern Med*. 2000;160:1641-9.
63. Matthews K, Cauley J, Yaffe K, Zmuda JM. Estrogen replacement therapy and cognitive decline in older community women. *J Am Geriatr Soc*. 1999;47:518-23.
64. Janowsky JS, Chavez B. Sex steroids modify working memory. *J Cogn Neurosci*; 2000:407-14.
65. Funk JL, Mortel KF, Meyer JS. Effects of estrogen replacement therapy on cerebral perfusion and cognition among postmenopausal women. *Dementia*. 1991;2:268-272.

66. Yaffe K, Haan M, Byers A, Tangen C, Kuller L. Estrogen use, APOE, and cognitive decline: evidence of gene-environment interaction. *Neurology*. 2000;54:1949-54.
67. Hogervorst E, Boshuisen M, Riedel W, Willeken C, Jolles J. 1998 Curt P. Richter Award. The effect of hormone replacement therapy on cognitive function in elderly women. *Psychoneuroendocrinology*. 1999;24:43-68.
68. Paganini-Hill A. Oestrogen replacement therapy and Alzheimer's disease. *Br J Obstet Gynaecol*. 1996;103:80-6.
69. Shumaker SA, Reboussin BA, Espeland MA, et al. The Women's Health Initiative Memory Study (WHIMS): a trial of the effect of estrogen therapy in preventing and slowing the progression of dementia. *Control Clin Trials*. 1998;19:604-21.
70. Wren BG. Megatrials of hormonal replacement therapy. *Drugs Aging*. 1998;12:343-8.
71. Yaffe K, Sawaya G, Lieberburg I, Grady D. Estrogen therapy in postmenopausal women: effects on cognitive function and dementia. JAMA. 1998;279:688-95.

Figure 1. Benefits of Hormone Replacement Therapy

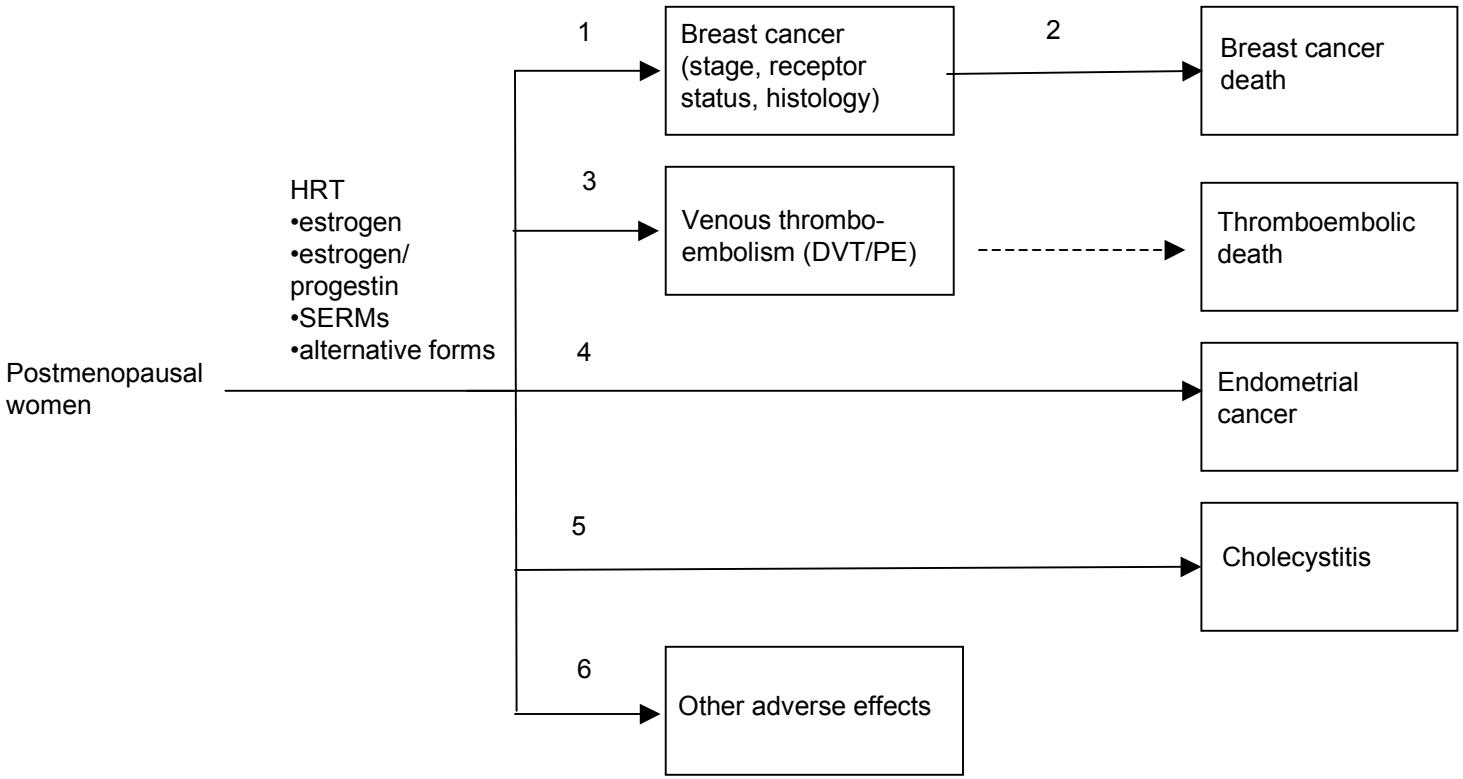
Analytic Framework 1



SERMs=Selective estrogen receptor modulators

Figure 2. Adverse Effects of Hormone Replacement Therapy

Analytic Framework 2



SERMs=Selective estrogen receptor modulators
DVT/PE=Deep-vein thrombosis/pulmonary embolus

Figure 3. Results of Meta-Analysis of Dementia Studies

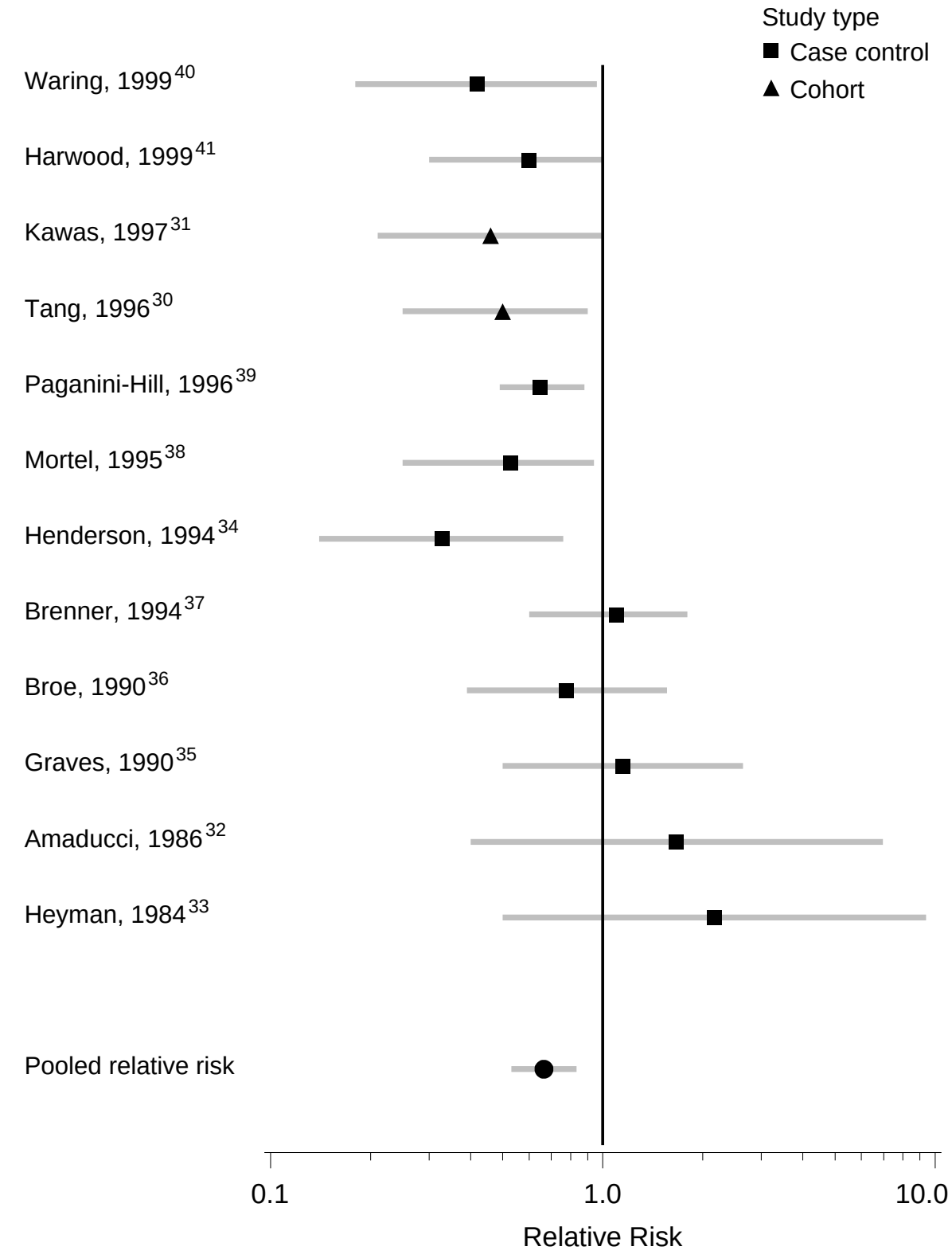


Table 1. Summary of Cognitive Test Results

Cognitive function	Reference	Positive tests/ total tests*	Subject profile- Studies with positive tests	Education adequately controlled for in studies with positive tests**	Explanation of results and magnitude of effect§
Memory					
Memory battery	51¥	1/1	Symptomatic	No	Increase of 8.0 with use
mmediate verbal recall	22#†,49,59, 60,61¥,62	4/9	Symptomatic	Yes in 2/3 studies	Paragraph recall: increase of 2.2, 5.9¶, and 11.5¶ with use; selective reminding: increase of 2.4 and 2.8¶ with use; associate learning: increase of 1.7 and 14.0¶ with use
Delayed verbal recall	49,59,60,61¥	3/8	Asymptomatic/ Symptomatic	Yes in 2/3 studies	Paragraph recall: change of -5.4 and 1.52¶ with use; selective reminding: increase of 16.6¶ and 21.6¶ with use; associate learning: increase of 2.6 and 19.3¶ with use
Visual memory	24,47,49,52, 59, 61¥	1/9	Not stated	Yes	Fewer errors made by users in 1 study; 8 measures in 5 other studies were negative
Attention					
Working memory	47¥,48,64	0/5	N/A	N/A	Increase of 0.2‡, 0.7‡, and 3.2 with use
Complex attention	23#,47,50, 52¥, 61,62,63	2/9	Symptomatic	No	Positive findings were on 2 tests not repeated by other studies; 1 was only of borderline significance (*P* =.08); 4 studies found no effects on digit symbol; 2 studies found no effect on trail making
Mental tracking	22#,23#, 47¥,,49,50¥, 52¥,59, 61,62	2/14	Symptomatic	Yes in 1/2 studies	1 of 5 studies had improvement on digit span‡; change of -1.67, 2.25 and 11.25¶ with use
Vigilance	47¥,23#,52¥	3/5	Symptomatic	No	5 different tests were used; in 1 study, visual search improved by 0.4 to 4 min¶ and sorting improved by 3 to 4 min¶ with use; other positive result was only of borderline significance (*P* =.07)
Concept Formation & Reasoning	22#,60¥,62,	2/3	Asymptomatic/ Symptomatic	Yes	Abstract reasoning; increase of 3.4¶ and 11.0¶ with use
Motor Speed	47¥,22#,23#	2/3	Symptomatic	Yes in 1/2 studies	Clerical speed and accuracy: increase of 9.5¶ with use; reaction time: 160-millisecond improvement with use
Mental Status	61¥,62,63¥, 65¥,66	2/5	Asymptomatic	N/A	Dementia screening exams‡: increase of 0.89 & 0.90¶ with use
Verbal Functions & Language	59,61¥,62,65¥	1/4	Asymptomatic	N/A	Category fluency and retrieval: increase of 3.4¶ & 6.0 with use

* Positive test indicates that women using estrogen scored significantly higher (at 0.10 significance level) than nonusers.
Total tests refers to the total number of test sessions on that cognitive measure. The same test may have been used by more than one study and some studies may have used more than one type of test to measure that cognitive function.
** Either baseline education was measured and was equal between groups or was adjusted for in the analysis.
§ Scores normalized using a total of 100 points.
# Scores derived from figure and represent estimates of effect magnitude.
† Denominator not reported so estimated by authors.
¶ Significant (P<.05).
¥ No data given in study.
‡ Ceiling effect may have contributed to null findings.

Table 2. Influence of Symptoms on Results of Randomized Controlled Trials

Subject profile	Reference Number	Results
Symptomatic Subjects had various complaints, including hot flashes, sleep problems, depression, fatigue, chest pressure	22 49 23 51 52	All of the studies found that women given estrogen had improvement in at least one cognitive test
Not symptomatic Subjects had fewer than 4 hot flashes during 2 week trial	50	No improvement in cognition in women using estrogen
Not known if subjects were symptomatic	47 48	Neither study found an improvement in cognition in women using estrogen

Table 3. Summary of Meta-analysis Results of HRT and Dementia

Studies included	Number of Studies	References	Fixed Effects		Random Effects		Test of Homogeneity X2 (p value)
			Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	
All studies	12	30-41	0.66	0.55-0.78	0.66	0.53-0.82	11.92 (p>0.10)
Case-control only	10	32-41	0.69	0.58-0.83	0.71	0.56-0.91	11.04 (P>0.10)
Cohort only	2	30, 31	0.50	0.30-0.77	0.50	0.30-0.80	0.87 (P>0.10)
Excluding poor quality studies	4	30, 31, 37, 40			0.64	0.32-1.06	
Excluding studies with proxy bias	9	30 - 33, 35-37, 39, 40			0.72	0.55-0.96	
Alzheimer disease only	11	30-37, 39-41			0.68	0.51-0.89	
Alzheimer disease by NINCDS-R criteria	7	30, 31, 34-37, 41			0.67	0.46-0.92	
Alzheimer disease by NINCDS-R criteria-case - control only	5	34-37, 41			0.77	0.46-1.16	
All studies except Heyman[33]	11	30-32, 34-41	0.65	0.55-0.77	0.65	0.52-0.80	
Using confidence intervals for Heyman[33] from Yaffe[71]	12	30-41	0.67	0.57-0.79	0.68	0.53-0.84	
Using t distribution	12	30-41			0.67	0.53-0.81	
Using Yaffe[71] as prior distribution	12	30-41			0.67	0.51-0.87	

NINCDS-R = National Institute of Neurological and Communicative Disorders and Stroke

Table 4. Summary of Evidence

Key Questions	Evidence codes*	Quality of Evidence
Arrow 6a		
1. Does HRT improve or stabilize cognition?	I, II-2	Fair-poor: studies enlist different patient populations and measure many different outcomes; results for symptomatic women are different than asymptomatic women. Duration of studies is too short to be meaningful. Difficult to draw any conclusions because outcome measures are so diverse.
2. What is the optimal dose and duration of use?		Fair-poor: too many other factors varied among studies to conclude which type or dose is protective. RCTs lasted only a few months and could not evaluate long-term use.
Arrow 6b		
1. Does HRT lower the risk for dementia?	II-2	Fair-poor: although the meta-analysis supports a protective effect, methodologic limitations and biases exist in individual studies (e.g., healthy user effect, use of proxy interviews, historical data obtained from demented subjects).
2. What is the optimal dose and duration of use?	II-2	Fair-poor: several studies did not specify type of preparation or dose; results for duration of use were based on different measures and are mixed.

***Study Design Categories**
(from Guide to Clinical Prevention Services, 1996)[14]

I: Randomized, controlled trials
II-1: Controlled trials without randomization
II-2: Cohort or case-control analytic studies
II-3: Multiple time series, dramatic uncontrolled experiments
III: Opinions of respected authorities, descriptive epidemiology

Appendix 1. Search Strategy

Hormone replacement therapy and cognition

1 exp hormone replacement therapy
estrogen replacement therapy

2 hormone replacement.tw. (text word taken from title and abstract of article)

3 estrogen replacement.tw.

4 exp estrogens/ad,tu (ad = administration & dosage; tu = therapeutic use)

equilenin	estrogens, catechol
equilin	estrogens, conjugated
estradiol	estrogens, non-steroidal
estriol	estrone

5 exp estrogens, synthetic/ad,tu

estrogens, non-steroidal	epimestrol
chlorotrianisene	ethinyl estradiol
coumestrol	mestranol
dienestrol	quinestrol
diethylstilbestrol	hexestrol
zearalenone	zeranol

6 1 **or** 2 **or** 3 **or** 4 **or** 5

7 exp mental processes

cognition	learning
mental fatigue	mind-body relations (metaphysics)
perception	thinking
volition	

8 cognition disorders

9 exp dementia

AIDS dementia complex	dementia, vascular
Alzheimer disease	Creutzfeldt-Jakob syndrome

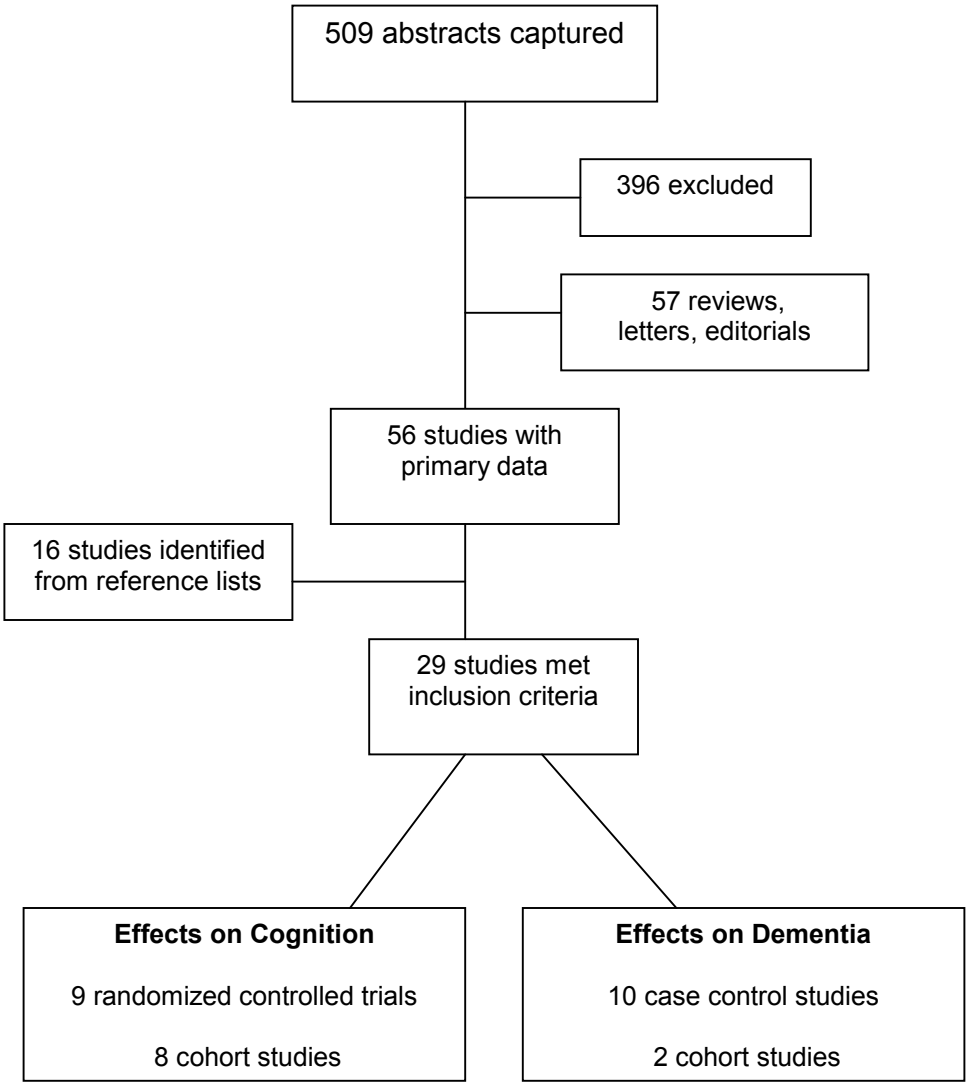
10 exp memory

deja vu	memory, short-term
retention (Psychology)	recall

11 memory disorders

12 7 **or** 8 **or** 9 **or** 10 **or** 11
13 6 **and** 12
14 **limit** 13 to human
15 **limit** 14 to english language
16 *looked at english abstracts of foreign articles*

Appendix 2. HRT and Cognition--Search Results



Appendix 3. Inclusion/Exclusion Criteria

Title and Abstract Review--Exclusion Criteria

1. Non-human
2. Foreign language (unless key article)
3. Looked only at men
4. Did not address links in analytic framework
5. Reviews/letters/editorials that did not seem to offer new perspective or helpful reference list

Literature Review--Inclusion Criteria

1. Human
2. Postmenopausal women
3. Non-demented subjects
4. Any type of study (cohort, cross-sectional, case-control, randomized clinical trial) with primary data on the relationship between HRT and cognition (key question 6a) or HRT and dementia (any type of dementia) (key question 6b)
5. Review all meta-analyses

Evidence Tables—Inclusion Criteria
"Best Evidence Approach"

1. For the association between HRT and cognition, studies that meet the following criteria:
   a. Randomized double-blind placebo-controlled and cohort studies
   b. Objective measurement of cognition (not just subjective cognition)

2. For the association between HRT and dementia, studies that meet the following criteria:
   a. Cohort or case control study
   b. State dementia criteria (state how determined cases had dementia and controls did not)

3. If two reports of the same population, the more recent updated data will be included.

4. Peer reviewed published articles (no abstracts)

Appendix 4a. Criteria for Grading Quality of Randomized Controlled Trials: The Jadad Score[25]

Study received one point for each “yes” or zero point for each “no” for each of the following questions:

1. Was the study described as randomized such as using the words randomly, random, and randomization?
    a. An additional point was given if method of randomization was described and it was appropriate (for example, table of random numbers, computer generated)
    b. A point was deducted if the method of randomization was inappropriate (for example, patients allocated alternately by birth date or hospital number)
2. Was the study described as double blind?
    a. An additional point was given if method of blinding was described and it was appropriate (for example, identical placebo)
    b. An additional point was deducted if method of blinding was inappropriate (for example, comparing placebo tablet with injection)
3. Was there a description of withdrawals and dropouts?

Maximum number of points is 5.

Appendix 4b. Criteria for Grading the Internal Validity of Individual Studies

Design-Specific Criteria and Quality Category Definitions[26]

Presented below are a set of minimal criteria for each study design and then a general definition of three categories- "good," "fair," and "poor" – based on those criteria. These specifications are not meant to be rigid rules but rather are intended to be general guidelines, and individual exceptions, when explicitly explained and justified, can be made. In general, a "good" study is one that meets all criteria well. A "fair" study is one that does not meet (or it is not clear that it meets) at least one criterion but has no known "fatal flaw." "Poor" studies have at least one fatal flaw.

Case Control Studies

Criteria:

- Accurate ascertainment of cases
- Nonbiased selection of cases/controls with exclusion criteria applied equally to both
- Response rate
- Diagnostic testing procedures applied equally to each group
- Measurement of exposure accurate and applied equally to each group
- Appropriate attention to potential confounding variable

Definition of ratings based on criteria above:

Good: Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equal to or greater than 80 percent; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables.

Fair: Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rate less than 80 percent or attention to some but not all important confounding variables.

Poor: Major selection or diagnostic work-up biases, response rates less than 50 percent, or inattention to confounding variables.

Randomized Controlled Trials and Cohort Studies

Criteria:

- Initial assembly of comparable groups:
  -for RCTs: adequate randomization, including first concealment and whether potential confounders were distributed equally among groups
  -for cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies, or intention to treat analysis for RCTs.

(continued)

Appendix 4b. Criteria for Grading the Internal Validity of Individual Studies (continued)

Definition of ratings based on above criteria:

Good: Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes are considered; and appropriate attention to confounders in analysis. In addition, for RCTs, intention to treat analysis is used.

Fair: Studies will be graded "fair" if any or all of the following problems occur, without the fatal flaws noted in the "poor" category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred in follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention to treat analysis is done for RCTS.

Poor: Studies will be graded "poor" if any of the following fatal flaws exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention to treat analysis is lacking.

Author, year	Type of study	How recruited	Setting	Number of subjects	Number originally randomized	Mean age (years) (Range)	Confirm menopause with FSH/ estradiol?	Percentage with surgical menopause	Symptomatic?
Janowsky 2000[64]	Trial	Paid volunteers	Portland, OR	6 & 7	Not stated	69	Yes	Not stated	Not stated
Shaywitz, 1999[48]	Crossover	Paid volunteers from community	New Haven, CT	46	47	50.8 (33-61)	Yes	Not stated	Not stated
Polo-Kantola, 1998[47]	Crossover	Volunteers-newspaper ads	Finland	62	70	56.3 (47-65)	Yes	24% BSO	Not stated
Phillips, 1992[49]	Trial	Recruitment after TAH/BSO	Canada	10 & 9	31	overall 48.2 +/- 4.7	TAH/BSO	100% TAH/BSO	Yes-Placebo groups with more hot flashes; No difference in mood,depression, anxiety, hostility
Ditkoff, 1991[50]	Trial	Not stated	East Los Angeles-Hispanic	24 & 12	Not stated	Overall 53 (45-60)	Yes	100% TAH	No- <4 hot flash episodes for 2 weeks
Sherwin, 1988[22]	Crossover	Recruitment after TAH/BSO (cases) or TAH (controls)	Canada	10&10& 10&10	59	45.4	Yes	Cases-100% TAH/BSO	Not specifically stated but most were likely symptomatic after TAH/BSO

Author, year	Type of study	How recruited	Setting	Number of subjects	Number originally randomized	Mean age (years) (Range)	Confirm menopause with FSH/ estradiol?	Percentage with surgical menopause	Symptomatic?
Fedor-Freybergh 1977[23]	Trial	Recruitment at outpatient clinic	Gynecology clinic in Stockholm	11 & 10	25	56.5 (47-70)	Yes	Not stated	Yes. Did not ask about hot flashes but majority had sleep problems, depression, fatigue.
Hackman, 1976[51]	Trial	Not stated	England	9 & 9	Not stated	29-68	No	44% TAH/BSO < 6 months ago	Yes
Vanhulle, 1976[52]	Trial	Volunteers	Nuns in Belgium	11 & 15	12 & 17	56.6 (cases) & 58.7 (controls)	No	Not stated	Yes. Estrogen group had fewer menopausal symptoms. Had less hot flashes, chest pressure, and fatigue.

Author, year	HRT form	Duration of use	Follow-up rate	Percentage compliant to HRT	Possible confounders?	Method of measuring outcome
Janowsky 2000[64]	.625 mg conjugated estrogen	I month	100%	Not stated	None	Interview
Shaywitz, 1999[48]	Oral-CEE 1.25 mg/day	21 day course with 14 day break	100%	94%	Not measured	1.Interview 2.Functional MRI
Polo-Kantola, 1998[47]	Transdermal- If age <56-Estrogel 0.6 mg/g (2.5 g/d), If age >55-Evorel patch 50 ug/24 hr	3 month course with 1 month wash-out	100%	92%	No difference in depression scores with estrogen therapy	Interview
Phillips, 1992[49]	10 mg of estradiol valerate IM each month	2 months	100%	61%	No difference on anxiety, depression, hostility scores	Interview
Ditkoff, 1991[50]	Oral CEE .625(12) or oral CEE 1.25(12) daily for 25 days/month	3 months	?100%	?100%	Users had improved depression (not controlled)	Interview
Sherwin, 1988[22]	1. Estradiol valerate 10.0 mg IM 2. Testosterone 150 mg IM 3. Estradiol dienanthate 7.5 mg + Estradiol benzoate 1.0 mg + Testosterone 150 mg IM	2- 3 month treatment periods with 2 month wash out	100%	85%	No difference in: baseline scores, education, occupation, personality inventory	Interview

Author, year	HRT form	Duration of use	Follow-up rate	Percentage compliant to HRT	Possible confounders?	Method of measuring outcome
Fedor-Freybergh 1977[23]	2 mg estradiol-17B-valerianate (Progynon) daily	3 months	100%	84% (exclude noncompliant)	Estrogen users slightly older and had decreased depression, anxiety, and fatigue and improved sleep	Interview
Hackman, 1976[51]	1.5 mg piperazine oestrone sulphate twice daily	6 months	100%	Not stated	None	Interview
Vanhulle, 1976[52]	4 mg estriol daily	3 months	100%	90%	Age	Interview

Author, year	Outcome	Results	Analysis	Further explanation of main differences	Additional Information on Study Methodology	Jadad Score
Janowsky 2000[64]	SOP working memory table	Not significant				2
Shaywitz, 1999[48]	Verbal working memory tasks Nonverbal working memory tasks Brain activation	Not significant Not significant Significant	Intention to treat	Increased activation of certain brain areas during verbal storage	6 not menopausal by FSH/ estradiol, did not use standard tests for measuring cognition; May not have been able to discriminate with tests because all scored high "ceiling effect"	5
Polo-Kantola, 1998[47]	Simple reaction time Multistep reaction time Subtraction test Statement verification test Auditory serial addition Digit span Digit symbol Benton visual retention Letter cancellation	Not significant for each test	Only those that complete study		Mean serum estradiol levels lower with patch (190 vs 431 pmol/u)	4
Phillips, 1992[49]	Digit span (WMS) Paragraph recall-immediate Paragraph recall-delayed Associate learning (WMS)-Immediate associate learning (WMS)-Delayed Visual reproduction (WMS)	Not significant p<0.05 Not significant p<0.05 p<0.05 Not significant	Only those that complete study	Immediate paragraph recall--users had improvement in score;no change in placebo Associate learning--users stayed the same; placebo had decline	Compared pre and post scores. Did not compare scores of the estrogen and placebo groups	4
Ditkoff, 1991[50]	Digit span (WAIS) Digit symbol (WAIS)	Not significant Not significant	Not clear		Hispanic (American born); Compared pre and post scores. Did not compare scores of the estrogen and placebo groups	4
Sherwin, 1988[22]	Digit span Clerical speed & accuracy Paragraph recall test Abstract reasoning	p<0.01 p<0.01 p<0.01 p<0.01	Only those that complete study and compliant	Scores of all treatment groups higher than placebo; Scores of treatment groups dropped during the	Only 10 women in estradiol treatment alone; Likely many women with symptoms	3

Author, year	Outcome	Results	Analysis	Further explanation of main differences	Additional Information on Study Methodology	Jadad Score
Fedor-Freybergh 1977[23]	Subjective cognition Reaction time Visual search Color word test -Stroop Sorting task (KTV) Attention test (USTM)	"More improvement" in estrogen group p<0.001 (for HRT vs placebo);p<.01 (for change HRT) p<0.01 (simple); p<0.001 (with memory load) p<0.01 (simple);p<0.001 (with interference) p<0.01 (time); p<0.001 (errors)		Estrogen users improved but placebo with no change	Compared difference in estrogen and placebo groups; Also looked at change in pre and post test score for estrogen and placebo groups--p value for estrogen group and difference between estrogen and placebo same unless specified	3
Hackman, 1976[51]	Guild Memory Test	p<0.02	Not reported	Estrogen users improved but placebo with no change	Did not compare estrogen and placebo groups; No correlation between subjective improvement in memory and GMT Score; 10/18 identified because of menopausal signs/symptoms; Sherwin reanalyzed and did not find significant result; 3 estrogen users had large improvement and 1 had large	2
Vanhulle, 1976[52]	Subjective cognition, BVRT, Series of numbers (WAIS)(Digit span), Substitution (WAIS) (Digit symbol), Arithmetic (GIT), Manual Labyrinth of Rey, Reaction time, Vigilance, Tempo of work (spot pattern test), Attention (spot pattern test)	"no significant differences" for each test	Only those that complete	Compared difference in means of both groups (pretest-posttest). Estrogen users had greater improvement in	Compared difference in means of both groups (pretest-posttest). When compared post-tests, only attention was different (p<0.03 for unadjusted analysis). Change in overall health score--no change in yes but did have change in no's	3

Author year	Setting/ Project	HRT user (n)	Non-user (n)	Eligibility	Mean Age (years) (User/ Nonuser)	Percentage with surgical menopause (User/ Nonuser)	Education (years) (User/ Nonuser)	Other differences	How determine use/ nonuse	Definition user	Definition nonuser
Rice, 2000[62]	Kame project	Current-196 Past-186	455	Japanese-American living in King County, Washington Aged ≥65 years with 2 years of followup	69.61 71.47 72.1	45.41 55.41 74.5	13.1/12.8/ 12.1		Interview	Current or past use	Never use
Carlson 1999[59]	McGill University	14	41	No major acute or chronic medical or psychiatric illness; no psychotropic medications or glucocorticoids	71.2	72.4	14.4/11.9 ($p<0.05$)	Users had higher socio-economic status	Questionnaire	Current use	Not current use

Author year	Setting/ Project	HRT user (n)	Non-user (n)	Eligibility	Mean Age (years) (User/ Nonuser)	Percentage with surgical menopause (User/ Nonuser)	Education (years) (User/ Nonuser)	Other differences	How determine use/ nonuse	Definition user	Definition nonuser
Matthews 1999[63]	Study of Osteoporotic Fractures (SOF)	Current-1325 Past-2612	5714	>64 yrs, Not Black, Able to walk without help, No history of bilateral hip replacement	Users younger	Users more likely to report surgical menopause	Users more educated	Users more likely to use sedatives/anxiolytics, less likely to smoke	Interview	Ever use of oral estrogen at initial assessment	Never use of oral estrogen at initial assessment
Jacobs, 1998[60]	Community based study of aging/ dementia in NY	81	646	Free of dementia, stroke, CVA; complete data	73.8/74.3	Not stated	11.0/9.1 ($p<0.05$)	Users were more likely to be white; medical conditions not different; same level of depression	Questionnaire	Ever use of HRT	Never use of HRT
Yaffe, 2000[66]	Cardiovascular Health Study	Current 297 Past 336	2083	Age >65, community dwelling, able to give informed consent	70.6/71.9/72.5	Not stated	14.5/13.6/13.4	Users with younger age, current users drank more alcohol per week	Interview	Current or past use	Never use

Author year	Setting/ Project	HRT user (n)	Non-user (n)	Eligibility	Mean Age (years) (User/ Nonuser)	Percentage with surgical menopause (User/ Nonuser)	Education (years) (User/ Nonuser)	Other differences	How determine use/ nonuse	Definition user	Definition nonuser
Resnick, 1997[24]	Baltimore Longitudinal Study of Aging	18	18	>39 yr; No dementia; Normal BVRT1 at start; Short interval between HRT use and BVRT; No past use of HRT; No use of vaginal cream	59.9/ 60.2	33.3/16.7	15.6/15.8		Interview	Never user at first test and current user at time of follow-up test	Never user at both tests
Barrett-Connor, 1993[61]	Rancho Bernardo Cohort	394	406	>64 years, live in Rancho Bernardo, CA	76.9	Not stated	2/3 completed college or more-not state if users differ	Users were less depressed	Interview; pill & prescription review	Ever use of oral estrogen at initial assessment	Never use of oral estrogen at initial assessment
Funk, 1991[65]	Veterans Administration longitudinal study of aging	30	77	Age 40-69, Caucasion	67	Not stated	Not stated	Users were more likely to smoke and drink more than 2 alcholic drinks per day	Medical record review	Current use	Not current use

Author year	HRT form	Duration of HRT use	Length of follow-up	Follow-up rate	Con-founders controlled	Con-founders not controlled	How outcome determined	Results (For *current or ever users* - not past users)	Trends
Rice, 2000[62]	Use of any form of unopposed estrogen or combined with progestin	4 - 15 years	2 years	89.9%	Age, education, in care, marital status, birthplace, language, formal education, WSAID, health check-ups, thyroid disease, cancer history, blood pressure, BMI, smoking, physical activity, alcohol, fiber, hysterectomy status, de-pression, age of		Cognitive Abilities Screening Instrument (CASI) Abstract reasoning subsection Category fluency subsection Attention subsection Recent verbal memory subsection Mental tracking subsection	p<0.05 p<0.001 p<0.05 Not significant Not significant Not significant	
Carlson 1999[59]	50% unopposed oral 0.625 mg CEE; 30% CEE with 2.5 mg MPA; 20% CEE 0.30 mg	19.5 years	18 months	67%	Socioeconomic status and years of education	Mood, medical problems	Immediate paragraph recall Delayed paragraph recall Immediate paired associates Delayed paired associates Immediate Selective Reminding Delayed Selective Reminding Immediate visual paired associates Delayed visual paired associates Visual reproduction Figural memory Digit span Visual memory span Category retrieval	Not significant Not significant Not significant Not significant Not significant p<0.01 Not significant Not significant Not significant Not significant Not significant Not significant Not significant	None studied

Author year	HRT form	Duration of HRT use	Length of follow-up	Follow-up rate	Con-founders controlled	Con-founders not controlled	How outcome determined	Results (For *current or ever users* - not past users)	Trends
Matthews 1999[63]	Oral only	Current-14.3 yr Past-5.2 yr	4-6 years	77%	Age, education, activity limitations, initial performance	Mood, medical problems	Modified MMSE Trails B Digit Symbol	Not significant Not significant Not significant	Past users had more benefit than current users
Jacobs, 1998[60]	Any but most used unopposed oral CEE	4.55 yr	2.5 years	72%	Education, age, ethnicity	Mood, medical problems	Immediate Selective Reminding Delayed Selective Reminding Similaries Subtest Boston Naming Test	p<=0.01 p<=0.001 Not significant Not significant	None studied
Yaffe, 2000[66]	Unopposed estrogen only	Not stated	7 years	84%	Age, education, race, stroke history	Mood, medical problems	Modified mini mental status exam	p=0.0023	Only current users had less decline

Author year	HRT form	Duration of HRT use	Length of follow-up	Follow-up rate	Con-founders controlled	Con-founders not controlled	How outcome determined	Results (For *current or ever users* - not past users)	Trends
Resnick, 1997[24]	Oral or transdermal		Not stated	Not stated	Age, baseline BVRT score, interval between assessments	Education, mood, symptoms, medical problems	BVRT	p=0.05	None studied
Barrett-Connor, 1993[61]	Any but most used unopposed oral CEE (80% used premarin)	Current-19.1 yr Past-7.7 yr	15 years	80%	Education, age, ethnicity	Mood, medical problems	Immediate Selective Reminding MMSE Trails B Category naming/Fluency Visual reproduction tests Months backwards 5 minute recall Serial sevens World backwards	Not significant Not significant Not significant Not significant Not significant Not significant Not significant Not significant Not significant	Long term users (>20 years) scored 1 point higher on Category Fluency than never users (p<0.01)
Funk, 1991[65]	"Almost entirely unopposed" estrogen	Not stated	Maximum of 6 years	Not stated	Length of time since menopause	Education, mood, medical problems	Cognitive Capacity Screening Examination (CCSE)	Not significant	

Author year	Differences	Comments	Quality Score
Rice, 2000[62]	Significant improvement only seen in users of current unopposed estrogen. Current users of estrogen-progestin actually had decline in scores		Fair
Carlson 1999[59]	Users had improvement in scores on delayed selective reminding but non-users had decreased scores	Analysis based on only 10 users and 27 nonusers	Fair

Author year	Differences	Comments	Quality Score
Matthews 1999[63]	Only past users exhibited smaller decline in MMSE (p=0.03) and Trails B (p=0.02); Current users did not differ from nonusers	When only looked at those who were consistent current, past, or never users, results not changed	Fair
Jacobs, 1998[60]	Selective Reminding Test-Users with improved scores while nonusers scores declined; No difference in scores over time on other tests	Level of depression did not differ by HRT use history	Fair
Yaffe, 2000[66]	Current users had a 1.5 point decline in 3 months while never users had a 2.7 point decline	When model was adjusted for confounders, there was no longer a significant difference	Fair

Author year	Differences	Comments	Quality Score
Resnick, 1997[24]	Users with stable number of errors over time compared to increased number of errors in nonusers		Poor; Did not state follow-up rate
Barrett-Connor, 1993[61]	No difference in age related decrease in cognitive function in current or past users		Fair
Funk, 1991[65]	May have been no difference because of ceiling effects on the CCSE--both groups almost to maximum scores; Cerebral blood flow was in normal range for both groups across the length of the study and no difference in perfusion was seen between groups	Study included women with a history of transient ischemic attacks (TIA) or reversible ischemic neurologic deficits (RIND) but only the results from the women without this history are included in this table; there were benefits on cognition and cerebral blood flow in estrogen users who have history of RIND or TIA	Poor; No education information or adjustment

Appendix 5 Evidence Table 3: HRT and Dementia -- Case-Control Studies 1 of 1

Author, year	Setting	Number of Cases /Controls	Type of dementia	Criteria for dementia	How cases were found	Definition of Controls-How exclude dementia
Waring, 1999[40]	Rochester, Minnesota-Population based	222/222	Alzheimer disease	Diagnostic criteria "equivalent" to NINCDS-ADRDA	Rochester Epi Project Records Linkage System & retrospective review of medical records by one neurologist	Extensive medical evaluation in index yr of case but no sign of dementia per neurologist medical record review
Harwood, 1999[41]	Alzheimer disease center-Miami 30% Hispanic	White -229/139 Hispanic-133/53	Alzheimer disease	NINCDS-ADRDA	Were evaluated at Alzheimer disease center	Age >=65, Normal MMSE, Normal 4 trial recall of 3 words in MMSE
Paganini-Hill, 1996[39]	Nested in Leisure World Cohort of 8877 women-Retirement Community California-High Socioeconomic status	248/1198	Alzheimer disease	Dementia diagnosis listed on death certificate; exclude multi-infarct dementia or dementia from another likely cause	Death certificate list Alzheimer disease, "senile dementia," "dementia," "senility"	No mention of dementia on death certificate
Mortel, 1995[38]	Baylor College of Medicine and Houston VA	93(Alzheimer disease)/65(IVD)/148	1.Alzheimer disease 2. IVD	1.NINCDS-ADRDA 2. State of California Alzheimer Disease Diagnostic and Treatment Centers Criteria	Referral by local physicians and support groups Subjects in studies of aging & dementia	Neurological examination and neuropsychologic assessment revealed neurologically, cognitively normal
Henderson, 1994[34]	Altzheimer disease research center California	143/92	Alzheimer disease	NINCDS-ADRDA (70 confirmed with autopsy)	Volunteers recruited from community outreach who meet criteria for Alzheimer disease by history, exam, lab	Neurological examination and detailed neuropsychologic assessment – nondementia neurologically unimpaired

Author, year	Setting	Number of Cases /Controls	Type of dementia	Criteria for dementia	How cases were found	Definition of Controls-How exclude dementia
Brenner, 1994[37]	HMO, Washington	107/120	Alzheimer disease	NINCDS-ADRDA	Alzheimer disease patient registry in HMO	MMSE score of at least 28 and no evidence of dementia on psychometric evaluation, chart review, judgement of study nurse
Graves, 1990[35]	Two Alzheimer disease referral centers in Washington State	60/60 women (130/130 total)	Alzheimer disease	NINCDS-ADRDA	90% from Alzheimer disease center and 10% from Veteran's Administration	No memory loss (not stated how determine this)
Broe, 1990[36]	Two dementia clinics in Sydney, Australia	106/106 women (170/170 total)	Alzheimer disease	NINCDS-ADRDA	Referral by local physicians who had been requested to refer all new dementia cases	MMSE score of at least 26; Neurology of Aging examination
Amaducci, 1986[32]	Neurology departments in Italy	60 female cases /60 hospital and 50 female community controls (116/116/97 total)	Alzheimer disease	Blessed dementia scale; 2 signs or symptoms of cognitive decline; no depression; no evidence for dementias other than Alzheimer disease by history, exam, testing	Admission to neurology departments of seven centers	Blessed dementia scale excluded dementia
Heyman, 1984[33]	Duke Medical Center, North Carolina	28 female cases/ 56 female controls (40/80 total)	Alzheimer disease	"Rigorous criteria"	Participants of another comprehensive study of Alzheimer disease	MMSE>20

Author, year	How controls were found	Mean Age (years) (Case/Control)	%surgical menopause (Case/Control)	Average Education (years) (Case/Control)	Other differences
Waring, 1999[40]	Linkage system-residents during index year and matched by age (+/- 3 yr) and length of time in linkage system	Not stated-Case matched to control +/- 3 yr	10 / 9	12 / 12	Cases with less breast cancer; No difference in age at menarche, age at menopause
Harwood, 1999[41]	85% recruited for free memory screening; 9% evaluated at Alzheimer disease center;	Cases-White 79.9/ Hispanic 76.0 Controls-White 75.7/ Hispanic 71.5	Not stated	Cases-White 12.1/ Hispanic 9.9 Controls-White 13.8/ Hispanic 10.8	Cases with higher alcohol use and more hypertension but not statistically significant
Paganini-Hill, 1996[39]	Death certificates-Matched on year of death and year of birth	87.7/87.3	Not stated	Not stated	Significant trend of decreasing risk with increasing weight
Mortel, 1995[38]	Friends and relatives	73.7(Alzheimer disease)/74.4(IVD) /72.3 (Controls)	Not stated	Not stated	No difference in postmenopausal interval, age of onset of dementia, and duration of cognitive impairment for the two demented groups
Henderson, 1994[34]	Volunteers recruited from community outreach in whom Alzheimer disease excluded by exam, assessment	76.0/76.3	39/44	12.2/13.9	No difference in number of medications

Author, year	How controls were found	Mean Age (years) (Case/Control)	% surgical menopause (Case/Control)	Average Education (years) (Case/Control)	Other differences
Brenner, 1994[37]	Stratified random sample of HMO matched within 2 yr	78.7/76.6	47/16	Percentage with > 12yrs: 34.6/60.8	--
Graves, 1990[35]	Friends and relatives-matched for sex and age within 10 years	66.2/63.6 (men & women)	Not stated	Not stated	Cases more likely to have first-degree relative with history of dementia
Broe, 1990[36]	Clinic controls-matched for sex and age within 2 years	78.6/78.7 (men & women) 78.6 men 78.7 women	Not stated	Not stated	Cases more likely to have first-degree relative with history of dementia
Amaducci, 1986[32]	Hospital (116) and friend/neighbor (97) controls-matched for age (within 3 years), sex, and region of residence	31 aged 51-60; 25 aged 61-70; 19 aged 71-80 for cases	13.7 / 10.0 had oophorectomy	No significant association found between education/literacy and case/control status	Cases more likely to have first-or second-degree relative with history of dementia
Heyman, 1984[33]	Population controls--random digit dialing; matched for sex, race, and 5 year age interval	60.8 (51-71) for cases	Not stated	49% of cases and 22% of controls had education beyond high school	Cases had greater history of thyroid disease and history of severe heart disease

Author, year	Response rate	How HRT use determined	Definition of HRT use	Definition of Nonuser	HRT form/dose	Duration of HRT use
Waring, 1999[40]	Not applicable	Record abstraction-blinded to case/control status or hypothesis	Any form (oral, IM, topical, suppository) of estrogen used for > 6 months after menopause but before onset of Alzheimer disease	Never use	90% used oral +/ - topical; most CEE	Not stated
Harwood, 1999[41]	Not stated	Cases- proxy interview Controls-self interview	Ever use	Never use	Not stated	Median 2 years
Paganini-Hill, 1996[39]	Not applicable	Questionnaire prior to death-85% complete >=5 years before death	Ever use	Never use	Any	Not stated
Mortel, 1995[38]	Not stated	Medical records, questionnaires and interviews-surrogate used for patient with dementia	Current user	Not current user	Not stated	Not stated
Henderson, 1994[34]	Not stated	Cases- proxy interview Controls-self interview	Current user	Not current user	Any (>81% oral CEE)	Not stated

Author, year	Response rate	How HRT use determined	Definition of HRT use	Definition of Nonuser	HRT form/dose	Duration of HRT use
Brenner, 1994[37]	Not stated 2 cases excluded for missing ?? data	1. From 1977-Computerized pharmacy records 2. Prior 1977-Proxy interview	Ever use, Use in year prior to diagnosis	Never use	Any; 66% used oral CEE	Looked at number of prescriptions
Graves, 1990[35]	Screened 800 medical records and 188 met criteria; 143 entered study, 130 provided control	Proxy telephone interview (88% were spouse of >10 years)	"Estrogen replacement" use prior to symptoms	No "estrogen replacement"	Not stated	Not stated
Broe, 1990[36]	Screened 333 to obtain 174 Alzheimer disease cases & 170 participants cases; Screened 270 to obtain 170 controls	Proxy interview (>85% were spouse or 1st degree relative)	"Hormonal treatment"	No "hormonal treatment"	Not stated	Daily for at least 6 months
Amaducci, 1986[32]	Of 152 eligible admissions, obtained 116 cases; did not state how many screened to get controls	Proxy interview (>90% spouse or offspring);	"Use of estrogens in menopause"	No "use of estrogens in menopause"	Not stated	Not stated
Heyman, 1984[33]	Not stated	Proxy interview	Current use of "estrogen replacement"	Not current user	Not stated	Not stated

Author, year	Confounders controlled	Confounders not controlled	Number cases & controls that used HRT	Adjusted OR (95% CI)	Significance	Trends
Waring, 1999[40]	Age, education, length of time in linkage system	Mood, symptoms, medical problems, ethnicity	9 & 20 ?? for matched data – which is how got OR for ≥6 mo. vs. never	0.42 (0.18-0.96)	p=0.04	No significant decreased risk with use less 6 months; trend fro decreasing risk with increasing duration(p=0.04); no cumulative dose effect
Harwood, 1999[41]	Education, age	Mood, symptoms, medical problems	White-28 & 44 Hispanic-14 & 35	White-0.6 (0.3-1.0) Hispanic-0.4 (0.2-1.0)	p=0.05 for both	None stated
Paganini-Hill, 1996[39]	Age, weight, blood pressure medication, weight, menopause type, age, last menstrual period, age at menarche	Mood, education, symptoms, ethnicity	96 & 568	0.65 (0.49-0.88)	p=0.005	Oral only: 0.7(0.5-0.98); Significant dose trend; Significant duration trend (only users for >5yrs. Had significant decreased risk of AD)
Mortel, 1995[38]	For analysis of Alzheimer disease and IVD: none For analysis of all dementia: age	Mood, education, symptoms, ethnicity	11 (Altzheimer disease) & 7 (IVD) & 29 (Controls)	Alzheimer disease: 0.55(0.26-1.16) IVD: .0.50(0.20-1.2) All: 0.53(0.26-0.94)	Not significant	None stated
Henderson, 1994[34]	Education, age	Mood, symptoms, ethnicity	10 & 17	0.33 (0.14-0.76)¶	p=0.01	

¶ Odds ratios and confidence intervals are unadjusted and calculated from data in tables.

Author, year	Confounders controlled	Confounders not controlled	Number cases & controls that used HRT	Adjusted OR (95% CI)	Significance	Trends
Brenner, 1994[37]	Age, history of hysterectomy before & after age 55, (education & ethnicity not found to be confounders so not included in the model)	Mood, symptoms	52 & 58	1.1 (0.6-1.8)	Not significant	Oral: 0.7 (0.4-1.5); Current 0.6 (0.3-1.2); No dose trend in number of prescriptions
Graves, 1990[35]	Age	Education, medical problems, symptoms, mood, ethnicity?	11 & 10	1.15 (0.50-2.64)	Not significant	
Broe, 1990[36]	Age	Education, medical problems, symptoms, mood, ethnicity?	14 & 18	0.78 (0.39-1.56)	p=0.48	
Amaducci, 1986[32]	Age, area of residence	Education, medical problems, symptoms, mood, ethnicity?	6 & 4	1.67 (0.39-6.97)[§]	p=0.73	
Heyman, 1984[33]	Age, race, education residence	Medical problems, symptoms, mood	4 & 4	2.17 (0.5-9.41)[†]	p>0.05	

[§] Matched odds ratios and confidence intervals calculated from data in paper using SAS.

[†] Odds ratios and confidence intervals are unadjusted and calculated from data in tables. Adjusted OR given in study is 2.38. Confidence intervals obtained by Yaffe et al were 0.7 to 7.8.

Author, year	Comments	Quality score
Waring, 1999[40]		Fair
Harwood, 1999[41]	Study was of both men and women--info in table on both	Poor
Paganini-Hill, 1996[39]	Did not exclude those with dementia at baseline; Earlier 1994 report had similar results) From previous reports of cohort – primarily Caucasian & highly educated	Poor
Mortel, 1995[38]		Poor
Henderson, 1994[34]	OR is unadjusted, calculated using data in table. Authors state that univariate analysis same as multivariate analysis.	Poor

Author, year	Comments	Quality score
Brenner, 1994[37]		Good
Graves, 1990[35]	Study of both men and women; Number of cases and controls may not be correct because they gave percentage and not clear if this is the percentage of the total or of just women (used women); Not blinded interviewers; HRT only one of many risk factors studied; Kappa for agreement in reported HRT use between controls in the validation subsample and their surrogates was 0.64.	Poor
Broe, 1990[36]	Study of both men and women; Not stated if blinded interviewers; HRT only one of many risk factors studied; Kappa for agreement in reported HRT use between controls in the validation subsample and their surrogates was not specifically stated.	Poor
Amaducci, 1986[32]	Odds ratios and confidence intervals are for population controls- When hospital controls are used the OR is 0.71 with a p value of 0.77; Study of both men and women; Not state if blinded interviewers; Estrogen only one of many risk factors studied; Only 52% of proxy respondents could answer question about estrogen use during menopause; Agreement in reported HRT use between controls in the validation subsample and their surrogates was not specifically stated but was greater than 60%.	Poor
Heyman, 1984[33]	Study of both men and women; Number of cases and controls may not be correct because they gave % and not clear if this % of total or of just women (used women); Not blinded interviewers; HRT only one of many risk factors studied; Kappa for agreement in reported HRT use between controls and their surrogates was 0.63.	Poor

Author, year	Cohort	Recruit-ment Strategy	HRT (n)	Non-user (n)	Mean Age (years)	Percentage with surgical menopause (User/ Nonuser)	Education (years) (User/ Nonuser)	Other differences	Eligibility	Definition of user	How determine use	HRT form	Duration of use
Kawas, 1997[31]	Baltimore Long-itudinal Study of Aging	Not stated	230	242	61.5-range 28-94/ No difference (data not given)	Not stated	No difference (data not given) 63% college graduates	No difference in age, menopause, type of menopause	Information on HRT, up to 16 yr follow-up	Ever use of oral or trans-dermal	Interview	Oral (212)/ Patch (18)	Not stated
Tang, 1996[30]	Manhattan Study of Aging	Recruitment from Medicare and senior housing	156	968	73.0, 74.4 - users younger (p=0.01)	50/26 (p=0.0001)	10.2 / 9.0 (p=0.005)	Users: fewer Blacks, earlier menopause	No cognitive impair-ment at baseline/ Informa-tion on HRT	Ever use after meno-pause	Interview	"Majority"-Con-jugated Equine Estrogen	2 mo-49 yrs (average 6.8 yr)

Author, year	Neuro-psycho-logical testing at baseline	Outcome	Criteria	How determine outcome	Length of follow-up	Follow-up rate	Con-founders	Confounders not controlled	Number of users and nonusers with AD	Adjust-ed RR (95% CI)	Trends	Comments	Quality Score
Kawas, 1997[31]	Yes	Alzheimer disease	NINCDS-ADRDA	Multi-disciplinary evaluations every 2 year including neuropsy-chological assess-ments	Up to 16 years	Not stated	Education Age	Mood, Symptoms, Medical problems, Ethnicity	9 & 25	0.457 (0.209-0.997)	No duration effect		Fair; Did not state follow-up rate
Tang, 1996[30]	Yes	Alzheimer disease	NINCDS-ADRDA	Medical records and imaging studies and data from initial and follow-up study exam-inations	1-5 years	84%	Education Age, Ethnicity Part-icipation group	Mood, Symptoms, Medical problems	9 & 158	0.5 (0.25-0.9)	RR 0.13 for users > 1 yr; users with later age of onset	Women who did not remember HRT classified as nonusers	Fair; Poss-ibility of misclass-ification bias

Appendix 6. Cognitive Measures Used in Studies

Cognitive Measure	Test	Sub-category	Differences in performance between users and nonusers: Reference Number																	
			48	47	49	50	22	23	51	52	59	63	60	24	61	65	64	66	62	
Attention	Attention	Complex attention						p<0.01												
	Digit Symbol	Complex attention		NS		NS				NS		NS								
	Spot Pattern Test	Complex attention							p=0.08											
	Trail Making Test-Part B	Complex attention										NS			NS					
	Arithmetic-Groninger Intelligence test	Mental tracking								NS										
	Auditory Serial Addition Test	Mental tracking		NS																
	Digit Span	Mental tracking		NS	NS	NS	p<0.01			NS	NS									
	Months Backwards	Mental tracking													NS					
	Serial Sevens	Mental tracking													NS					
	Stroop Color Word Test	Mental tracking		NS				p<0.01												
	World Backwards	Mental tracking													NS					
	Letter Cancellation	Vigilance		NS																
	Multi-step Reaction Time	Vigilance		NS																
	Sorting (KVT)	Vigilance						p<0.01												
	Vigilance Test	Vigilance							p=0.07											
	Visual Search	Vigilance						p<0.01												
	Working Memory	Working memory															NS			
	Nonverbal Working Mem. Test	Working memory	NS																	
Attention cont.	Subtraction Test	Working memory		NS																

Appendix 6. Cognitive Measures Used in Studies

Cognitive Measure	Test	Sub-category	Differences in performance between users and nonusers: Reference Number 48	47	49	50	22	23	51	52	59	63	60	24	61	65	64	66	62
	Verbal Working Memory Test	Working memory	NS																
	Working Memory Task	Working memory		NS													NS		
	Verification Test	Working memory		NS															
Concept Formation & Reasoning																			
	Similarities (WAIS-R)	Concept formation											NS						
	Abstract Reasoning	Reasoning				p<0.01													
Learning Ability																			
	Manual Labyrinth of Rey	Learning ability								NS									
Memory	Guild memory Test	Memory battery							p<0.02										
Mental Status																			
	CCSE	Mental status														NS			
	MMSE	Mental status										NS			NS				
	Modified 3MS	Mental status																p<0.05	
	CASI	Mental status																	p=0.023
Motor Speed	Simple Reaction Time	Motor speed		NS				p<0.01											
	Clerical Speed and Accuracy	Speed or perception				p<0.01													

Appendix 6. Cognitive Measures Used in Studies

Cognitive Measure	Test	Sub-category	Differences in performance between users and nonusers: Reference Number 48	47	49	50	22	23	51	52	59	63	60	24	61	65	64	66	62
Skills																			
	Boston Naming Test	Verbal expression											NS						
	Category Naming/Retrieval	Verbal fluency									NS				NS				
Verbal Memory																			
	5 Minute Recall	Delayed verbal recall													NS				
	Associate Learning-Delayed	Delayed verbal recall		p<0.05							NS								
	Paragraph Recall-Delayed	Delayed verbal recall			NS						NS								
	Selective Reminding-Delayed	Delayed verbal recall								p<0.01			p<0.001		NS				
	Associate Learning-Intermediate	Delayed verbal recall		p<0.05							NS								
	Paragraph Recall - Intermediate	Delayed verbal recall		p<0.05		p<0.01					NS								
	Selective Reminding-Intermediate	Delayed verbal recall									NS		p<0.01		NS				
Verbal Memory	Visual Paired Associates- Delayed	Delayed verbal recall									NS								
	Visual Paired Associates-Intermediate	Immediate verbal recall									NS								
	Benton Visual Retention	Visual memory		NS						NS				p=0.05					
	Figural Memory	Visual memory									NS								
	Visual Memory Span	Visual memory									NS								
	Visual Reproduction (WMS)	Visual memory			NS						NS				NS				

Appendix 7. Acronym/Abbreviation List

BID	Twice daily
BMI	Body mass index
BSO	Bilateral salpingo-oophorectomy
BVRT	Benton visual retention test
CASI	Cognitive abilities screening instrument
CCSE	Cognitive capacity screening examination
CE	Conjugated estrogen
CEE	Conjugated equine estrogen
CI	Confidence intervals
CVA	Cerebrovascular accident
FMRI	Functional magnetic resonance imaging
FSH	Follicle stimulating hormone
IM	Intramuscular
IVD	Ischemic vascular disease
MMSE	Mini mental status exam
MPA	Medroxy progesterone acetate
MRI	Magnetic resonance imaging
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer Disease and Related Disorders Association Criteria for Alzheimer Disease
OR	Odds ratio
pmol/u	Picomole per unit
RR	Relative risk
SOP	Self ordered pointing
USTM	User skills and task match
TAH	Total abdominal hysterectomy
WAIS	Wechsler adult intelligence scale
WMS	Wechsler memory scale
3MS	Modified mini-mental state examination

Appendix 8. Acknowledgements

This systematic evidence review was prepared for the Agency for Healthcare Research and Quality (contract #290-97-0018, task order no. 2) to be used by the U.S. Preventive Services Task Force. Task Force members Janet Allen, PhD, RN, and Steven Teutsch, MD, MPH, served as liaisons. Jeri Janowsky, PhD, Mark Helfand, MD, MS, and Eric Orwoll, MD provided scientific expertise. Oregon Health & Science University, Evidence-based Practice Center staff who contributed to this project included Peggy Nygren MA, research associate, Susan Wingenfeld, administrative assistant, and Patty Davies, MA, librarian.