Hormone Replacement Therapy and Cognition:

A Systematic Evidence Review

And Meta-Analysis

by

Erin S. LeBlanc, MD

Presented to the Department of Public Health
And the Oregon Health Sciences University
School of Medicine
In partial fulfillment of
Masters of Public Health
May 2000

School of Medicine
Oregon Health Sciences University

CERTIFICATE OF APPROVAL

This is certify that the MPH thesis of

Erin S. LeBlanc, MD

has been approved

[redacted]
Professor in Charge of thesis

[redacted]
Member

[redacted]
Member

Member

Member

STRUCTURED ABSTRACT

Objective: To evaluate data on the use of hormone replacement therapy (HRT) to prevent cognitive decline and dementia in healthy postmenopausal women.

Data Sources: All English language studies identified in MEDLINE, HealthSTAR, and Cochrane Library databases from 1966 to March 2000 or in PsychINFO from 1984 to January 2000. In addition, reference lists of key articles were reviewed for all related studies including those pre-dating the database search.

Study Selection: All studies with primary data on the effects of HRT on cognitive testing in nondemented postmenopausal women or the relationship between HRT and dementia.

Data Extraction: Twenty-seven studies meeting inclusion criteria were identified. Fifteen studies, including 9 randomized controlled trials and 6 cohort studies, looked at the effects of HRT on cognitive testing. Twelve studies, including 2 cohort studies and 10 case control studies, looked at the relationship between HRT use and risk of dementia.

Data Synthesis: The results of the studies that looked at cognitive testing could not be combined quantitatively because of heterogeneous study design. Some studies found that estrogen improved performance on verbal memory and vigilance tasks in symptomatic women. There were no benefits in asymptomatic women. A meta-analysis of studies on the relationship between HRT and dementia indicated that use of postmenopausal HRT was associated with a decreased risk of dementia (summary odds ratio = 0.66; 95% CI, 0.53-0.82). Studies did not contain enough information to assess the effects of various hormonal preparations and there was conflicting information about doses or duration of therapy.

Conclusions: Postmenopausal HRT may improve some aspects of cognition in symptomatic women but does not improve cognition in asymptomatic women. There appears to be a decreased risk of dementia in HRT users but this may be due to differences between users and nonusers and not secondary to HRT use.

TABLE OF CONTENTS

Acknowledgements: I would like to thank Dr. Linda Humphrey for her advice and enthusiasm for this project, Dr. Heidi Nelson for her practical guidance and support, and Dr. Motomi Mori for her statistical expertise. I would also like to acknowledge Benjamin Chan for his invaluable statistical help and skills with WinBUGS and Dr. Jeri Janowsky for her thorough review of this thesis and for her thoughtful comments and ideas. Finally, I would like to thank Ben because without his support and encouragement, this thesis could have never been completed.

INTRODUCTION

Background on Menopause and Hormone Replacement Therapy

Menopause refers to the cessation of ovarian function that occurs spontaneously in most women between the age of 50 and 52. Some postmenopausal women take hormone replacement (HRT), low dose estrogen with or without progesterone, short term for the relief of symptoms or more long term for the prevention of chronic diseases. According to a random-digit telephone survey, 37.6% of postmenopausal women in the US in 1995 were using HRT. The percentage was higher in women who had undergone hysterectomy (58.7%) and lower in those with natural menopause (19.6%). Use varied by location and education. Women in the South and West and college graduates were more likely to use hormone replacement therapy.[1] Surveys of selected populations reveal even higher usage patterns. For example, in health maintenance organizations, 50-70% of women aged 50-70 may be using HRT.[2,3]

The number of long term users is likely much lower than the number of short term users because many women discontinue HRT because of side effects that include vaginal bleeding, breast tenderness, or fears about cancer risk. In order to minimize these side effects, various types of estrogen and progesterone and different routes of administration have been developed. For women with a uterus, estrogen is given in combination with cyclic or continuous progesterone administration. Continuous administration is often used because it eventually produces amenorrhea instead of cyclic bleeding and many women do not want to continue having periods. However, in the first year, irregular bleeding is common with the continuous regimen and women may actually be more likely to discontinue therapy.[3] The multiple different types, forms, and combinations of

HRT make research in this area problematic. It is unknown whether results using one HRT regimen can be generalized to other formulations.

History of Hormone Replacement Therapy

During the menopausal transition, approximately 50% of women experience vasomotor symptoms (hot flashes and night sweats) and 20% seek care for these symptoms.[4] In the 1940's, it was first noted that these menopausal symptoms could be treated with estrogen and the marketing of postmenopausal estrogen became widespread.[3] However, the popularity of HRTamong female patients and their providers has changed throughout the years as more has been learned about its risks and benefits.

The use of postmenopausal estrogen declined in the mid 1970's as studies indicated that the use of unopposed estrogen (no addition of progesterone) was associated with an increased risk of endometrial cancer.[5,6] Although many women discontinued estrogen, others began to take progesterone in addition to the estrogen as this combined regimen was found to decrease the risk of endometrial cancer.

Women continued to take HRT because studies began to show that estrogen could prevent the bone loss that occurred in the immediate postmenopausal period.[3,6,7] Although it was originally hoped that therapy during the early postmenopausal period could confer lifelong bone protection, it has now been shown that rapid bone loss occurs when estrogen is discontinued.[8,9] Therefore, for osteoporosis prevention, women must continue taking HRT long term, especially during the period when they are at high risk of fracture. In the 1990's, it was shown that estrogen could actually increase bone density and reduce fracture risk even when taken by older women and HRTis now one of the major treatment modalities of osteoporosis.[10]

Cardiovascular disease is the most common cause of death among postmenopausal women. Although it had been shown in the 1970's that bilateral oophorectomy was associated with an increased risk of coronary heart disease (CHD) and that this risk was reduced by postmenopausal estrogen[11], interest in the association between CHD and HRT did not surge until the 1990's.[3] Observational studies have revealed that users of HRT have a significantly reduced risk of death from CHD.[6,12,13] Multiple mechanisms for estrogen's protective effects on CHD have been noted including its favorable effects on lipids, enhancement of endothelial-dependent vasodilation, inhibition of platelet aggregation, and improvement in intimal repair. However, a recent randomized controlled trial in women with known CHD did not find that estrogen reduced the risk of CHD events; in fact, there may have been an increase in events in the first few years which was offset by a decreased risk of CHD in later years.[14]

These findings may be explained by the prothrombotic effects of estrogen. In the last 5 years, studies have shown that postmenopausal HRT is associated with a 2 to 3 fold increased risk of deep vein thrombosis and pulmonary embolus.[14-19] The prothrombotic effects of estrogen may actually increase the risk of acute coronary events when first given, although its antiatherogenic properties may predominate after estrogen has been used for several years.[20]

Women overestimate their risk of breast cancer and many do not take HRT because of a possible association between postmenopausal estrogen and breast cancer. Over 50 observational studies have looked at the association between HRT and breast cancer and most have not shown a statistically significant association. Meta-analysis techniques have been used to try to summarize the disparate results and to increase the

power to detect a small increase in risk. Although the meta-analyses do not show a consistently increased risk among ever users of HRT, they do find a 20-30% increased risk of breast cancer after 5-10 years of use.[21-24] Recent studies suggest that the addition of progesterone to estrogen may increase the risk even further.[25,26]

According to a decision analysis, almost all women will benefit from HRT based on its positive effects of the cardiovascular system and bones.[7] However, the recent evidence about the possibly negative effects of HRT in women with known CAD, the increased risk of breast cancer in women who take combined HRT, and the increased risk of deep venous thrombosis and pulmonary embolus in HRT users have decreased the enthusiasm about HRT that abounded in the mid 1990's. Part of the loss of enthusiasm for HRT relates to the fact that there are other treatments for CHD and osteoporosis that may not have as many negative side effects. However, there are few preventive or treatment options for dementia, which is the most recently proposed benefit of HRT.

Dementia-Burden of Suffering

Dementia involves a general decline in cognitive function, behavioral disturbance, and/or interference with activities of daily living.[27] It is estimated that between 3 and 8 million people in the United States have dementia.[28] The most common type of dementia in the United States is Alzheimer's Disease (AD), which affects between 3 to 4 million people.[29,30] Alzheimer's disease is a progressive dementia and there is a loss of memory, language, visuospatial skills and personality changes.[31] The next most common causes of dementia include vascular causes (10-20%) and Parkinson's disease (5-10%).[31] The incidence of dementia is 1% per year in older individuals, although in the most elderly populations, this rate may be as high as 2-3%.[29] One community-based prevalence study

in East Boston estimated that almost 50% of those aged 85 and over suffered from dementia.[32] Most studies found that after accounting for differences in life expectancy, women have a significantly higher risk of AD than men.[29,32] AD is the fourth leading cause of death in the US.[29] The life expectancy of demented patients is greatly reduced. Those with early-onset AD have a median survival of 6.7 to 8.1 years, while survival in those with late-onset disease is 4.8 to 6 years.[29,34] In 1991, the annual cost of AD was estimated to be $67.3 billion.[35] 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.

Evidence for cognitive effects of estrogen

There is observational data that suggests a possible relationship between estrogen exposure and cognition. Short term estrogen levels have been associated with changes in cognitive testing. For example, women in the high estrogen phase of the menstrual cycle have been shown to do better on tests of motor skills compared to when they are in the high estrogen phase of the cycle. In contrast, spatial task performance appears to be enhanced during the time of low estrogen.[36-38]

More long term estrogen exposure has also been associated with cognitive function and risk of cognitive decline. In 87 women, lifelong estrogen exposure as measured by an index combining menstrual, reproductive, and physical markers associated with estrogen levels predicted performance on tests of verbal functions, attention, and global cognition.[39] Bone mineral density, which has been hypothesized to be a marker of cumulative estrogen exposure, has been negatively correlated with cognition function and risk of cognitive deterioration.[40] One cross sectional study of 124

patients with early-onset AD found that early age of menopause was associated with a significantly increased risk of AD. This did not apply to patients without a family history, however.[41]

There is biological plausibility for the association between estrogen and cognition. Estrogen readily crosses the brain barrier and there are estrogen receptors in several areas of the brain although the role of these receptors is still being studied.[52] Based mostly on rat models, there are several possible mechanisms for estrogen's effects on cognition. It may improve cholinergic activity in the brain similar to the effects of the anticholinesterase drug tacrine, which is used to treat AD.[42] In addition, it may decrease the activity of monoamine oxidase, which metabolizes catecholamines. Selegeline, one of the treatments of AD, also works through inhibition of monoamine oxidase activity. Estrogen may also promote neuronal circuitry through stimulating dendritic spine density.[52] In vitro, estrogen promotes the breakdown of precursors to B-amyloid, which is often found in the core of the neurofibrillary tangles and senile plaques of patients with AD.[42] Estrogen may also prevent cerebral atherosclerosis and vascular dementia (which is often difficult to differentiate from AD) by favorably affecting the lipid profile. Although these mechanisms may all contribute to estrogen's effects on cognition, actual studies in women are lacking.[52]

Because of the biological plausibility and the observational studies showing an association between both short term and long term estrogen exposure and cognition, many have proposed that estrogen replacement therapy after the menopause may prevent cognitive decline and the development of dementia. In this paper, data on the use of hormone replacement therapy (HRT) to prevent cognitive decline and dementia in

healthy postmenopausal women is examined. Specifically, the effects of HRT on cognitive testing in nondemented women and the effect of HRT on risk of dementia, focusing on Alzheimer's Disease, are discussed.

METHODS

Analytic framework and key questions

The analytic framework in Figure 1 shows the target populations, interventions, and health outcome measures that were examined. The accompanying key questions (Figure 2) correspond to the numbered arrows in the analytic framework and communicate the main questions guiding our literature review and that are addressed in the results section.

The focus was on the use of HRT as chemoprophylaxis and therefore focused on the use of HRT, which includes either estrogen alone or estrogen combined with progesterone, in healthy, nondemented postmenopausal women. I looked at whether HRT improved or stabilized cognitive function as measured by cognitive testing (arrow 1a). I also looked at whether HRT lowered the risk of AD or other dementias (arrow 1b).

Literature Review

To find articles on the relationship between HRT and cognition, MEDLINE, HealthSTAR, and Cochrane Library databases were searched for papers published in 1966 or later using the search strategy shown in Appendix 1. In addition, PsychINFO from 1984 to January 2000 was searched. A search was performed that would identify studies on the effects of estrogen on any cognitive or memory process or on any type of dementia diagnosis. Additional articles were obtained from reference lists of relevant

reviews. The search was updated monthly. As seen in Appendix 2, using this search strategy, 423 abstracts were identified for review.

A single reader reviewed all English abstracts. 312 papers were excluded from full review because they focused on animals, only studied men, were in a foreign language (unless a key article), did not address links in the analytic framework, or were reviews, letters, or editorials that did not seem to offer a new perspective or helpful reference list. From the original search and from the search of the reference lists of relevant reviews, 54 studies with primary data on the relationship between HRT and cognition in nondemented postmenopausal women were then abstracted by the same person.

In order to identify the most important studies for inclusion in the evidence table, a "best evidence" approach was used (Appendix 4). To address the key questions related to the association between HRT and cognitive testing (arrow 1a in analytic framework), only randomized, double blind, placebo controlled studies and cohort studies were included. Although randomized controlled trials (RCTs) are considered the most rigorous study design, RCTs on HRT and cognition mostly looked at younger, perimenopausal women that only used estrogen for short periods (months). Therefore, it was felt that cohort studies, which tended to look at an older population and to follow them for longer periods, would add important information. Nonrandomized trials, case control studies, and cross-sectional studies were excluded from the evidence tables because they were felt to be too subject to bias because estrogen users are believed to be substantially different than nonusers. Other studies have found that users are healthier, have healthier lifestyles, and are more highly educated than nonusers, for example.[8-12]

To address the key questions related to the association between HRT and dementia (arrow 1b in analytic framework), only cohort and case control studies were included in the evidence table. There were no RCTs 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, and because users differ from nonusers in numerous ways.[43-47]

For a case control study to be included, the study methodology had to provide details about how AD was determined in the cases and excluded in the controls. If two studies were done on the same population, the more recent study with the most updated data was included. Only articles published after peer review were included (no abstracts).

Data Extraction and Synthesis

1. HRT and cognitive testing

Data extraction. Randomized controlled trials and cohort studies that looked at the effects of HRT on cognition as measured by formal cognitive testing were reviewed. Abstracted data from the randomized controlled trials included the type of study and setting; a description of cases and controls including age and menopausal type; type of HRT and duration of use; whether subjects were symptomatic; and the effects of HRT on symptoms in cases and controls (to look for unblinding of the study). Study design issues that were recorded included exclusion/inclusion criteria, method of allocation, compliance and follow-up rates. In both the RCTs and cohort studies, the cognitive tests that were used and the results of the users and nonusers were recorded. Results of studies were based on different types of analyses. Analyses was recorded based on either 1)

comparisons of the change score (post-pre) of users with nonusers or 2) if the pretest scores were equal or appropriate adjustments were made, comparisons of the post scores of users with 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.

Data synthesis. Although the original goal was to quantitatively combine the results of the cognitive tests, the studies were felt to be 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[48] and expert opinion. Jadad scores were used to measure the quality of the randomized controlled trials.[49] For the other studies, methodologic limitations that could compromise the study's quality were noted and recorded in the evidence tables.

2. HRT and dementia

Data Extraction. Case control and cohort studies that looked at the relationship between HRT and risk of dementia of any type were reviewed. Abstracted data for both types of studies included the type of study and setting; a description of the cases and controls 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. The method of obtaining HRT exposure history was obtained because of the potential for recall bias and proxy bias. Demented women would be less likely than controls to remember previous exposure history. Although proxy respondents were used in several of the studies, they might not accurately remember exposure history or be 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 AD, are less prone to these biases. However, as AD is a insidious disease with a long latency period,[35] a long follow-up is needed to avoid finding a falsely low HRT usage rate in women with early cognitive decline.

I also recorded 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 Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA). This is the most widely applied criteria for defining AD clinically. The criteria require a clinical exam, standardized mental status testing, and neuropsychological testing. The inter-rater agreement for this criteria ranges from poor to good (Kappa statistic 0.36 to 0.65).[50,51] Using pathological diagnosis of AD 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.[51] Seventy percent of the errors in this study were false negatives.

The recorded outcome measures were adjusted odds ratios or relative risks with the associated 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.

Data synthesis. The odds ratios and relative risks of the case control and cohort studies of HRT and dementia were combined using Markov Chain Monte Carlo in Bayesian analysis (WINBUGS software). Using a p value of 0.10, a test of homogeneity was done. Both fixed effects and random effects models were done using a noninformative distribution as the prior probability. To determine how sensitive the

results were to the prior distribution, an alternative (t) distribution was used. In addition, the results of a previous meta-analysis were used as the prior distribution, and the results of the new studies were then combined.[52] Because one study did not contain enough information to calculate adjusted confidence intervals, unadjusted data was used in the analysis.[53] To see if this affected the summary estimates, the meta-analysis was repeated excluding this study as well as using confidence intervals obtained by a previous meta-analysis.[52] Because of concerns about combining different types of studies, the analysis was also done using just the cohort studies, the case control studies, the studies that looked at AD only, and the studies that used NINCDS-ADRDA criteria.

RESULTS

Effects on cognition (Figure 1, Arrow 1a)

1. Does the use of postmenopausal estrogen with or without progesterone improve or stabilize nondemented women's scores on cognitive testing?

The literature search identified 8 randomized controlled trials (Appendix 4, Evidence Table 1), and 6 cohort studies (Appendix 4, Evidence Table 2) that used formal testing to measure the effects of estrogen on the cognition of non-demented women.

The randomized controlled trials are dissimilar in several ways. Three of the studies used a crossover design[54-56] while the rest used separate experimental and placebo groups. The mean age of the women in the studies ranged from 45 to 80. Three studies used women immediately after a total abdominal hysterectomy/bilateral salpingo-oophroectomy[55,57,58] while 6 other studies used community volunteers with only a small

percentage of women having undergone surgical menopause.[54,56,59-62] Some of the studies used women with menopausal symptoms[55,57,60,62] while others used asymptomatic women.[58] Over 40 different cognitive tests were used as outcome measures in these studies, and 30 of these tests were used by only one investigator (Table 1). Only 3 tests were used by more than 2 studies and even when tests were repeated by several investigators, the test was often administered slightly differently. Only 2 studies, both done by the same research group, used the identical estrogen formulation and dose. The duration of use ranged from 40 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 between studies, results were not combined quantitatively. Instead, to look for any patterns, the tests were grouped according to the cognitive process they measured (Tables 1 and 2).

Memory. Although cross-sectional studies have suggested that estrogen may affect memory, especially verbal memory,[63-67] the results from the randomized controlled trials and cohort studies are conflicting. In a study of 18 women with menopausal symptoms, those given 1.5 mg of piperazine oestrone for 6 months did better than the placebo group on a memory battery, the Guild Memory Test.[61] However, the results from other studies that looked at specific components of memory did not report this finding.

Five studies, 2 randomized controlled trials and 3 cohort studies, used 3 tests of immediate verbal recall. The 2 randomized controlled trials, both done by the same investigators, found that 10 women randomized to intramuscular estradiol for 3 months performed better than 9-10 women given placebo on 2 tests of immediate verbal recall: paragraph recall and associate learning. However, the women in both of these studies

were immediately status post total abdominal hysterectomy/bilateral salpingo-oophorectomy and, although only measured in one study, likely had menopausal symptoms. Indeed, when the same research group used a cohort design to study elderly women (average age 73 to 74) 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 one study, 81 users' scores improved on this test, while the scores of the 646 nonusers declined.[68] However, in another study, 394 ever users of HRT in the Rancho Bernardo cohort did not perform better than the 406 nonusers on this same test.[69] The younger women in the Rancho Bernardo cohort performed significantly better than older women, suggesting this test was sensitive. In addition, the users in the Rancho Bernardo cohort had a longer duration of use than in the previous study (16.5 years versus 2.5 years), and were followed for a more extended period of time (15 years versus 2.5 years).[69] A third study 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.[59]

Estrogen exposure was associated with improvement in at least one test of delayed verbal recall in 2 out of 4 studies. The 2 randomized controlled trials found improvements in immediate paragraph recall, but did not find that these benefits extended to delayed recall of a paragraph.[55,57] Although one of these randomized controlled trials did find there was an effect on the delayed associate learning test,[57] 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 that users improved while nonusers declined.[59,68,69] The scores on the other tests of delayed verbal recall that were used in these cohort studies did not differ between exposure groups, however.[59,69]

Six studies used 6 tests of visual recall and did not, for the most part, find that estrogen was beneficial. Although 18 women who became users of postmenopausal estrogen during the course of follow up in the Baltimore Longitudinal Study of Aging performed better on the Benton Visual Retention test than matched nonusers[70] 2 randomized controlled trials did not find that women given estrogen for 3 months did better than those given placebo on this same test.[54,62] No study found that women exposed to estrogen did better on 3 other tests of visual recall.[57,59,69]

Attention. Some have suggested that estrogen may influence attention, especially working memory, because of its inhibiting effects on dopaminergic transmission.[54] Ten studies looked at various aspects of attention with disparate results. The most recent randomized controlled trial used nonstandardized tests to measure working memory. 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). They 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.[56] Another study also used nonstandardized tests to try to evaluate working memory. There was no difference in scores according to estrogen exposure and it is unlikely that the null results are secondary to a ceiling effect as the results in this study were measured in milliseconds.[54]

Performance on 2 measures of complex attention, the digit symbol and trail making test, were not affected by estrogen in any of the 5 studies that used them as cognitive measures.[54,58,62,69,71] However, one of these randomized controlled trials found that nuns given estrogen had borderline improvement on another test of attention, the spot pattern test (p=.08).[62] Another study found that estrogen treated women did better on a Swedish test of attention.[60] Both of these latter 2 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."[60]

Women using estrogen showed improvement compared to the placebo group on 2 of the 13 tests of mental tracking.[54,60,62,69] Although one randomized controlled trial found that estrogen exposure was related to performance on digit span, a test of mental tracking,[55] the women randomized to estrogen in 3 other randomized controlled trials did not perform better on this test.[54,58,62] Two of these studies used women who were post-op from total abdominal hysterectomy/bilateral salpingo-oophorectomy; however, the women had to have less than 4 hot flashes in a 2-week period in one of them.[58] The 24 women who were 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, although all of the pre-treatment scores being in the normal range may have precluded finding a difference (ceiling effect).[58] A crossover study of 62 women also did not find that women given estrogen for 3 months did better on the digit span test.[54] Since this study was the largest of the randomized trials and reported a power of 90%,[54] it is likely that the women in the other study 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' performance 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 had improvement on this tests.[60] However, these results were not confirmed by the later, larger study.[54]

Of the 3 studies that measured vigilance, 2 found that estrogen improved women's ability to sustain attention. In both of these studies, the women in the study were symptomatic with fatigue, sleep problems, hot flashes, and depression.[60,62] In contrast, a larger randomized crossover trial did not find that women given estrogen performed better on two sensitive tests of vigilance.[54]

Concept formation and reasoning. Concept formation and reasoning was tested in 2 studies with conflicting results. While a randomized crossover study found that subjects given estrogen improved in their abstract reasoning scores compared to when they were on placebo,[55] a New York based cohort study did not find that ever users scores changed over 2.5 years compared to never users.[68]

Motor speed. Motor speed, as measured by simple reaction time, was improved by postmenopausal estrogen in one study[60] but not another.[54] A randomized controlled trial that found estrogen improved reaction time by over 100 milliseconds included symptomatic women.[60] A larger, more recent trial did not find a difference in reaction time between exposure groups even though measurements were also in milliseconds.[54]

Women given estrogen had improvement in clerical speed and accuracy in another study.[55]

Mental status, verbal function, learning ability. No study found that women given estrogen had improvement in their mental status.[69,71,72] However, this is not unexpected given that the ease of the mental status exam might preclude finding subtle differences (ceiling effect). Neither verbal functions and language skills nor learning ability appeared to be affected by estrogen use.[59,68,69]

Influence of Symptoms. Table 3 summarizes the results of the reviewed studies according to whether the subjects in the studies were symptomatic. All of the studies that used women with various somatic complaints found that they had improvement in at least one cognitive test when given estrogen. The cognitive process that was most consistently improved in these women was verbal recall and vigilance, although complex attention, mental tracking, concept formation and reasoning, and motor speed were also affected in several studies. It may be that symptomatic women perform better on cognitive testing because of improved sleep, less fatigue, and less symptomatology. Alternatively, the subtle effects of estrogen on cognition may only be apparent in subjects that are not performing at maximum cognitive ability because of fatigue and loss of sleep secondary to menopausal symptoms.

The 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.[58] The largest study, a crossover 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. This study used sensitive tests (outcomes measured in milliseconds) and had a power of 90% to detect a difference between users and nonusers.[54]

Effects of Progesterone. 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%) of the women used unopposed oral conjugated equine estrogen.[59,68,69,72] None of these studies looked at subgroups that used progesterone. One small, nonrandomized trial of 19 symptomatic women used estradiol combined with progesterone. They concluded that progesterone 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 several previous randomized controlled trials of symptomatic women, immediate verbal recall was not enhanced by HRT exposure.[79]

2. What is the optimal formulation/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 2 of the randomized controlled trials used oral conjugated equine estrogen (either 0.625 or 1.25 milligrams) and neither found that the women randomized to estrogen performed better on several tests of cognition.[56,58] The study using transdermal estrogen also did not find a difference in cognitive test scores between the estrogen and placebo groups. Although the 2 studies that used intramuscular estradiol found that estrogen favorably affected women's performance on cognitive testing, these studies were done by the same author and used symptomatic women. The early studies that found beneficial effects on cognition used larger doses of oral estrogen than are currently prescribed. Although it is tempting to conclude that that

the different estrogen formulations or dosages may have contributed to the studies' disparate findings, there are too many other factors that varied between studies to draw any conclusions about which formulation or dose may be more protective.

The randomized controlled trials only lasted several months and so information about duration of use comes from the cohort studies. The one cohort study that looked at duration of use found that users of greater than 20 years scored one point higher on category fluency, a test of verbal functions and language skills. These long-term users did not perform better on any of the other 8 measures of cognition, however.[69] Another cohort study looked at recency of use and actually found that past users had more benefit than current users.[71]

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.
- Postmenopausal estrogen does not seem to enhance asymptomatic women's performance on cognitive testing.
- Only studies that used symptomatic women found that estrogen improved cognitive performance. The most consistent findings in these studies appeared to be on verbal memory and vigilance, although there were also effects on complex attention, mental tracking, concept formation and reasoning, and motor speed.
- There is insufficient evidence about the effects of the addition of progesterone to estrogen. One nonrandomized trial of a small number of women found that

progesterone 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 1, Arrow 1b)

1. Does the use of postmenopausal estrogen with or without progesterone lower the risk of Alzheimer's disease (AD) and other dementias?

Ten case control studies (Appendix 4, Evidence Table 3) and 2 cohort studies (Appendix 4, Evidence Table 4) on the association between postmenopausal estrogen use and risk of Alzheimer's disease were identified from the literature review and met the inclusion criteria.

The early case controls studies did not find an association between HRT and AD (Odds Ratio (OR) of 0.78 to 2.38).[53,74-76] HRT use was only one of many risk factors evaluated in these studies. These early studies all used proxy interviews to determine exposure to postmenopausal estrogen 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 (the Kappa values were 0.63 and 0.64).[53,74] The studies were relatively small with the number of users ranging from 8-21

or about 8-18% of the study population. All of the studies controlled for age, but only one study controlled for education.[53]

All but one of the case control studies done since 1990 have found a significantly decreased risk of AD among users of postmenopausal estrogen (OR of 0.33 to 1.1).[77-82] However, in some of these studies, 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.[78,80,82] As proxy informants would be expected to have less knowledge about HRT use, the rate of estrogen exposure in cases may have been underestimated. The largest case control study, which was nested in the Leisure World Cohort, used death certificate data to determine dementia outcomes, and defined HRT exposure with a self-administered questionnaire completed approximately 5 years before death. As dementia is insidious in onset and women with cognitive decline were not excluded at baseline, cases might have been less likely to remember previous estrogen use at the original questionnaire. This might have falsely lowered the percentage of users among demented subjects and lead to the finding of decreased risk of AD among HRTusers (OR 0.65; 95% CI 0.29,0.88).[80]

Three studies used more objective measures such as medical or pharmacy records to determine estrogen use, but their results are conflicting. One study stated that "medical records were the primary source material."[78] However, it did not specify how the material was abstracted and 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 AD (CI 0.26-1.16).

Another study enlisted subjects from the Group Health Cooperative of Puget Sound, a Health Maintenance Organization (HMO) in Seattle, Washington.[77] Cases were identified from the Alzheimer's Disease Patient Registry, which uses NINCDS-ADRDA criteria to diagnose AD. The HMO's computerized pharmacy records were then used to identify a subject's filling of 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 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 it 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, there was no decreased risk of AD 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 another study.[81] Dementia diagnosis and HRT use was 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 six months after the menopause but before the onset of AD was associated with a 68 percent reduction in AD.(OR 0.42; 95% CI 0.18-0.96).

The strongest evidence for an association between postmenopausal estrogen use and AD comes from two cohort studies (Appendix 4, Evidence Table 4). The Manhattan Study of Aging cohort was formed from Medicare listings and senior housing centers.[83] The average age of subjects was 74.2 years. One hundred fifty-six ever users of

postmenopausal estrogen and 968 nonusers were followed for 1-5 years for the development of AD as defined by the NINCDS-ADRDA criteria. After controlling for education, age, and ethnicity, users were significantly less likely to develop AD (Relative Risk (RR) 0.5; 95% CI 0.25-0.9). Users also had a later age of onset of AD. One problem with this study, however, was that subjects developed AD within 5 years of the initial interview. Given that a diagnosis of AD lags symptom onset by 3.5 to 5.5 years,[35] 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 16 years, which makes it less likely that the AD subjects had subtle memory problems at the beginning of the study.[84] They evaluated the subjects every 2 years for the development of AD as diagnosed by the NINCDS-ADRDA criteria. After controlling for age and education, the 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 5). The test of homogeneity indicated that the studies were homogeneous ($p>0.10$). When the studies were combined quantitatively using the random effects model, the summary odds ratio was 0.66 (CI 0.53-0.82) (Table 4). When case control or cohort studies were analyzed separately, the estimates did not change. Also, restricting the analysis to studies that only looked at AD or only used NINCDS-criteria also 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 AD in women exposed to postmenopausal estrogen, confounders may explain this inverse relationship. Women who use HRT are more educated[43,44,46] and formal education has been found to be protective against dementia.[29,34] Although several of the studies controlled for education and found that there was still a decreased risk of AD, there could be residual confounding.[79,81,82] HRT users are also healthier and have healthier lifestyles, and physical health status has been associated with cognitive changes with advancing age.[28] Users are also younger than nonusers and the most important risk factor for dementia is advancing age, although all of the studies controlled for age.

Only one study using pharmacy records, looked at the effect of progesterone on the risk of AD. Adding progesterone to the logistic regression model did not change the risk estimates, indicating that it was not a significant confounder.[77] Another study did not find that excluding women who reported a surgical menopause (who are usually taking unopposed estrogen) affected the results.[84]

Only one study looked at dementia other than Alzheimer's disease.[78] This study included women with dementia secondary to ischemic vascular disease (IVD). Women who used HRT had a 50 percent decreased risk of developing IVD, although the 95% confidence interval contained one (OR 0.50; CI 0.26-1.20).

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

The older case control studies do not specify the formulation of HRT that was used by the subjects.[53,74-76] The more recent case control studies define HRT exposure as

the use of any form (oral, IM, topical, suppository) of estrogen after the menopause.[77,79-81] The cohort studies only included women who used oral or transdermal forms of estrogen.[83,84] In the studies that looked at this information, 66-95% of the women used oral conjugated equine estrogen.[77,79,81] While one study found that only oral estrogen was associated with a decreased risk of AD,[77] another study found that oral, injectable, and/or cream were all associated with a decreased risk of dementia.[80]

Although one study found that taking at least 1.25 mg of oral conjugated estrogen was associated with greater risk reduction,[80] another study did not find that cumulative dose was associated with dementia risk.[81] The results for duration of use were also mixed. In the Manhattan Study of Aging, users with greater 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.[83] Another study also suggested that there was a threshold effect; only users for more than 6 months had a decreased risk of AD.[81] While one case-control study found that increasing duration of use was associated with a decreased risk of dementia, another case-control and a cohort study did not confirm this finding.[77,84]

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 of AD[78,79] while a third did not.[53] Another study found that the odds ratio for the risk of AD in current users was 0.6 (CI 0.3-1.2) compared to 1.7 (CI 0.9-3.2) in former users.[77] However, women with AD 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 44% decreased risk of AD among users of postmenopausal estrogen.
- However, the studies upon which this risk estimate is based have several flaws.

  -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 postmenopausal estrogen are healthier ("healthy user bias").
- There is insufficient data about the addition of progesterone.
- There is insufficient data about other forms of dementia.
- There is conflicting evidence about a dose response or duration effect.
- Although some studies have found current users had a decreased risk, this could be secondary to prescribing bias.

DISCUSSION

Conclusions

Although the study populations and outcome measures differ, the 9 randomized controlled trials and 6 cohort studies offer some provisional conclusions about the effects of postmenopausal estrogen on cognition. HRT does not appear to enhance asymptomatic women's performance on formal cognitive testing. In contrast, some studies have found that symptomatic women have improved cognitive performance,

especially in tests of verbal memory and vigilance, when given postmenopausal estrogen. There is insufficient evidence about whether progesterone attenuates these cognitive effects in symptomatic women. 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 two cohort studies suggest that HRT users have a 44% decreased risk of AD. However, there are several flaws in the studies upon which this estimate is based. The 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 are 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 AD not because of postmenopausal estrogen exposure, but because they are healthier and more educated. 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 progesterone to the regimen or whether specific dosages or formulations of estrogen are more protective.

Future Research Needs

Since women only have symptoms for a limited amount of 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 than

nonusers in many lifestyle and health behaviors, the ideal study would be a large, blinded, randomized controlled trial. The trial should last for at least several years in order that the effects of long term HRT can be studied and to increase the likelihood that subtle changes between treatment groups will be detected. Progesterone should be included as part of the intervention arm because of the possibility that it may attenuate some of estrogen's cognitive effects. Future studies also need to control for the psychological effects of estrogen to ensure any cognitive effects are not secondary to changes in depressive symptoms.

The cognitive tests used in future studies should measure aspects of verbal memory, vigilance, complex attention, mental tracking, concept formation and reasoning, and motor speed as these functions were enhanced in some studies of symptomatic women. Tests which do not have a ceiling value and that are sensitive to very small differences should be used because the effects of estrogen on cognition may be subtle in nondemented women. Most importantly, future studies should include measures of the ability to care for oneself, live independently, and complete activities of daily living because these are the real outcomes of interest.

Because of the methodological problems with the case control and cohort studies, a large, double blind randomized controlled trial lasting for several years is needed to determine if HRT actually reduces the risk of AD and other dementias. The study should include progesterone as many women take this with estrogen in order to reduce the risk of endometrial cancer. It should also include the various formulations and dosages of estrogen that are commonly used. Such a study is currently underway. The Women's Health Initiative Memory Study, which is a part of the Women's Health Initiative, is

currently looking at the effects of HRT on the development of AD and other dementias and will also look at the effect of HRT on cognitive function. If enough subjects continue to participate in the study, the effects of HRT on dementia and cognition may be known in the next several years.

REFERENCES

1. Keating NL, Cleary PD, Rossi AS, et al. Use of Hormone replacement therapy by Postmenopausal women in the United States. Ann Int Med. 1999;130:545-553.

2. Savor BG, Taylor TR, Woods NF. Use of hormone replacement therapy in Washington State: is prevention being put into practice? J Fam Prac. 1999;48:364-71.

3. Ettinger B. Overview of estrogen replacement therapy: a historical perspective. Proc Soc Exp Biol Med. 1998;217:2-5.

4. Johnson SR. Menopause and hormone replacement therapy. Med Clin N Am. 1998;82:297-320.

5. Grady D, Gebretsadik T, Kerlikowske K, et al. Hormone replacement therapy and endometrial cancer risk: a meta-analysis. Obstet gynecol. 1995;85:304-13.

6. Grady D, Rubin Sm, Petitti DB, et al. Hormone therapy to prevent disease and prolong life in postmenopausal women. Ann Int Med. 1992;117:1016-1037.

7. Col N, Eckman MH, Karas RH. Patient-specific decisions about hormone replacement therapy in postmenopausal women. JAMA. 1997;1140-1147.

8. Christiansen C, Christensen MS, Transbol I. Bone mass in postmenopausal women after withdrawal of oestrogen/gestagen replacement therapy. Lancet. 1981;1:459-461.

9. Lindsay R, Hart DM, MacLean A, et al. Decreased risk of fractures of the hip and lower forearm with postmenopausal use of estrogen. N Engl J Med. 1980;303:1195-8.

10. Lufkin EG, Wahner HW, O'Fallon WM, et al. Treatment of postmenopausal osteoporosis with transdermal estrogen. Ann Int Med. 1992;117:1-9.

11. Colditz G, Willett W, Stampfer M. menopause and the risk of coronary heart disease in women. N Eng J Med. 1987;316:1105-10.

12. Stampfer M, Colditz G. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiological evidence. Prev Med. 1991;20:47-63.

13. Sullivan JM. Coronary arterography in estrogen-treated postmenopausal women. Prog Cardiovasc Dis. 1995;38:211-222.

14. Hully S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. heart and

Estrogen/progestin replacement Study (HERS) Research Group. JAMA. 1998;280:605-613.

15. Daly E, Vessey MP, Hawkins MM, et al. Risk of venous thromboembolism in users of hormone replacement therapy. Lancet. 1996;348:977-80.

16. Grodstein F, Stampfer MJ, Goldhaber SZ, et al. Prospective study of exogenous hormones and risk of pulmonary embolism in women. lancet. 1996:348:983-7.

17. Varas-Lorenzo C, Garcia-Rodriguez LA, Cattaruzzi C, et al. Hormone replacement therapy and the risk of hospitalization for venous thromboembolism: a population-based study in southern Europe. Am J Epidemiol. 1998;147:347-90.

18. Perez Gutthann S, Garcia Rodriguez LA, Castellsague J, Duque Oliart A. Hormone replacement therapy and risk of venous thromboembolism: population based case-control study. BMJ. 1997;314:796-800.

19. Effects of estrogen or estrogen/progestin regimens on heart disease risk factors in postmenopausal women. the Postmenopausal Estrogen/Progestin Interventions (PEPI) Trial. The Writing Group for the PEPI Trial. JAMA. 1995;273:199-208.

20. Psatty BM, Heckbert SR, Atkins, D, et al. A review of the association of estrogens and progestins with cardiovascular disease in postmenopausal women. Arch Intern med. 1993;153:1421-1427.

21. Dupont WD, Page DL. Menopausal estrogen replacement therapy and breast cancer. Arch Intern med. 1991;151:67-72.

22. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Lancet. 1997;350:1047-1059.

23. Grady D, Rubin SM, Petitti DB, et al. Hormone therapy to prevent disease and prolong life in postmenopausal women. Ann Intern Med. 1992;117:1016-1037.

24. Steinberg KK, Thacker SB, Smith SJ, et al. A meta-analysis of the effect of estrogen replacement therapy on the risk of breast cancer. JAMA. 1991;265:1985-1990.

25. Ross RK, Paganini-Hill A, Wan PC, Pike MC. Effect of hormone replacement therapy on breast cancer risk: estrogen versus estrogen plus progestin. J Natl Cancer Inst. 2000;92:320-332.

26. Magnusson C, Baron JA, Correia N, et al. Breast-cancer risk following long-term oestrogen-and oestrogen-progestin-replacement therapy. Int J Cancer. 1999;81:339-44.

27. American Psychiatric Association Diagnostic and Statistical Manual, 4th ed, APA press, Washington DC, 1994.

28. Teri L, McCurry S, Logsdon R. Memory, thinking, and aging. What we know about what we know. *West J Med.* 1997;4:269-75.

29. Keefover R. The clinical epidemiology of Alzheimer's disease. *Neurol Clin.* 1996;2:337-51.

30. 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.

31. McBee WI, Kaily ME, Dugan E, Shumaker SA. Hormone replacement therapy and other potential treatments for dementias. Endoc Metab Clin N Am. 1997;26:329-45.

32. 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.

33. Jorm AF, Korten AE, Henderson AS. The prevalence of dementia: a quantitative integration of the literature. *Acta Psychiatr Scand.* 1987;76:465-79.

34. Breteler MM, Claus JJ, van Duijn CM, Launer LJ, Hofman A. Epidemiology of Alzheimer's disease. *Epidemiol Rev.* 1992;14:59-82.

35. Ernst RL, Hay JW. The US economic and social costs of Alzheimer's disease revisited. *Am J Public Health.* 1994;84:1261-4.

36. Chiarello C, McMahon MA, Schaefer K. visual cerebral lateralization over phases of the menstrual cycle: A preliminary investigation. Brain cogn. 1989;11:18-36.

37. Hampson E. Estrogen-related variations in human spatial and articulatory-motor skills. Psychoneuroendocrol. 1990;15:97-111.

38. Hampson E, Kimura D. reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. Behav Neruosci. 1988;102:456-459.

39. Smith CA, McCleary CA, Murdock GA, et al. Lifelong estrogen exposure and cognitive performance in elderly women. Brain cogn. 1999;39:203-218.

40. Yaffe K, Browner W, Cauley J, et al. Association between bone mineral density and cognitive decline in older women. JAGS. 1999;47:1176-1182.

41. van Duijn CM. Hormone replacement therpay and Alzheimer's disease. Maturitas. 1999;31:201-5.

42. Henderson V. Estrogen, cognition, and a woman's risk of Alzheimer's disease. Am J Med. 1997;103:11S-18S.

43. 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.

44. 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.

45. 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.

46. 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.

47. Barrett-Connor E. Postmenopausal estrogen and prevention bias. *Ann Intern Med.* 1991;115:455-6.

48. Lezak MD. *Neuropsychological assessment.* 3rd ed. New York: Oxford University Press; 1995.

49. 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.

50. 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.

51. 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.

52. Yaffe K, Sawaya G, Lieberburg I, Grady D. Estrogen therapy in postmenopausal women: effects on cognitive function and dementia. *JAMA.* 1998;279:688-95.

53. 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.

54. 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.

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

56. Shaywitz SE, Shaywitz BA, Pugh KR, et al. Effect of estrogen on brain activation patterns in postmenopausal women during working memory tasks. *JAMA*. 1999;281:1197-1202.

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

58. Ditkoff EC, Crary WG, Cristo M, Lobo RA. Estrogen improves psychological function in asymptomatic postmenopausal women. *Obstet Gynecol*. 1991;78:991-995.

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. 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.

61. Hackman BW, Galbraith D. Replacement therapy and piperazine oestrone sulphate ('Harmogen') and its effect on memory. *Curr Med Res Opinion*. 1976;4:303-6.

62. 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 menopause research.* Lancaster, Eng., MTP. 1976.

63. 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.

64. Resnick SM, Maki PM, Golski S, Kraut MA, Zonderman AB. Effects of estrogen replacement therapy on PET cerebral blood flow and neuropsychological performance. *Hormones Behavior*. 1998;34:171-82.

65. Kimura D. Estrogen replacement therapy may protect against intellectual decline in postmenopausal women. *Hormones Behavior*. 1995;29:312-21.

66. 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.

67. Kampen DL, Sherwin BB. Estrogen use and verbal memory in healthy postmenopausal women. *Obstet Gynecol*. 1994;83:979-83.

68. Jacobs DM, Tang MX, Stern Y, et al. Cognitive function in nondemented older women who took estrogen after menopause. *Neurology*. 1998;50:368-73.

69. Barrett-Connor E, Kritz-Silverstein D. Estrogen replacement therapy and cognitive function in older women. *JAMA*. 1993;269:2637-41.

70. Resnick SM, Metter EJ, Zonderman AB. Estrogen replacement therapy and longitudinal decline in visual memory. A possible protective effect? *Neurology*. 1997;49:1491-7.

71. Matthews K, Cauley J, Yaffe K, Zmuda JM. Estrogen replacement therapy and cognitive decline in older community women. *J Am Geriatr Soc*. 1999;47:518-523.

72. Funk JL, Mortel KF, Meyer JS. Effects of estrogen replacement therapy on cerebral perfusion and cognition among postmenopausal women. *Dementia*. 1991;2:268-272.

73. 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.

74. Graves AB, White E, Koepsell TD, et al. A case-control study of Alzheimer's disease. *Ann Neurol*. 1990;28:766-74.

75. Broe GA, Henderson AS, Creasey H, et al. A case-control study of Alzheimer's disease in Australia. *Neurology*. 1990;40:1698-707.

76. 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.

77. 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.

78. Mortel KF, Meyer JS. Lack of postmenopausal estrogen replacement therapy and the risk of dementia. *J Neuropsych Clin Neurosci*. 1995;7:334-7.

79. 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.

80. Paganini-Hill A, Henderson VW. Estrogen replacement therapy and risk of Alzheimer disease. *Arch Intern Med*. 1996;156:2213-7.

81. 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;52:965-970.

82. Harwood DG, Barker WW, Loewenstein DA, et al. A cross-ethnic analysis of risk factors for AD in white Hispanic and white non-Hispanics. *Neurology*. 1999;52:551-556.

83. 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.

84. 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.

Table 1: Summary of the cognitive test results organized by cognitive measure

Test	Overall Cognitive Measure	Subcategory	Sz	PK	PS	DF	Sw	FF	HG	VD	CS	Mw	JB	RK	BC	FK
Attention	Attention	Complex attention						p<.01								
Digit Symbol	Attention	Complex attention		NS		NS				NS		NS				
Spot pattern test	Attention	Complex attention								p=.08						
Trail Making Test-Part B	Attention	Complex attention										NS			NS	
Arithmetic-Groninger Intelligence Test	Attention	Mental tracking								NS						
Auditory Serial Addition Test	Attention	Mental tracking		NS												
Digit Span	Attention	Mental tracking		NS	NS	NS	p<.01			NS	NS					
Months backwards	Attention	Mental tracking													NS	
Serial sevens	Attention	Mental tracking													NS	
Stroop color word test	Attention	Mental tracking		NS				p<.01								
World backwards	Attention	Mental tracking													NS	
Letter cancellation	Attention	Vigilance		NS												
Multistep reaction time	Attention	Vigilance		NS												
Sorting (KVT)	Attention	Vigilance						p<.01								
Vigilance Test	Attention	Vigilance								p=.07						
Visual search	Attention	Vigilance						p<.01								
Nonverbal working memory test	Attention	Working memory	NS													
Subtraction test	Attention	Working memory		NS												
Verbal working memory test	Attention	Working memory	NS													
Verification test	Attention	Working memory		NS												
Similarities (WAIS-R)	Concept formation & Reasoning	Concept formation											NS			
Abstract Reasoning	Concept formation & Reasoning	Reasoning					p<.01									
Manual Labrynth of Rey	Learning ability	Learning ability								NS						
Guild Memory Test	Memory	Memory battery							p<.02							
CCSE	Mental Status	Mental status														NS
MMSE	Mental Status	Mental status										NS			NS	
Simple reaction time	Motor speed	Motor speed		NS				p<.01								
Clerical speed and accuracy	Motor speed	Speed of perception					p<.01									
Boston Naming Test	Verbal functions/ language skills	Verbal expression											NS			
Category naming/retrieval	Verbal functions/ language skills	Verbal fluency									NS				NS	

Sz: Shaywitz, PK: Polo-Kantolo, PS: Phillips & Sherwin, DF: Ditkoff, Sw: Sherwin, FF: Fedor-Freybergh, HG: Hackman & Galbraith, VD: Vanhulle & Demol, CS: Carlson & Sherwin, Mw: Matthews; JB: Jacobs, RK: Resnick, BC: Barrett-Connor, FK: Funk

Table 1: Summary of the cognitive test results organized by cognitive measure

Test	Overall Cognitive Measure	Subcategory	Sz	PK	PS	DF	Sw	FF	HG	VD	CS	Mw	JB	RK	BC	FK
5 minute recall	Verbal memory	Delayed verbal recall													NS	
Associate learning-delayed	Verbal memory	Delayed verbal recall			p<.05						NS					
Paragraph recall-delayed	Verbal memory	Delayed verbal recall			NS						NS					
Selective Reminding-Delayed	Verbal memory	Delayed verbal recall									p<.01		p<.001		NS	
Visual Verbal Learning-Delayed	Verbal memory	Delayed verbal recall														
Associate learning-Immediate	Verbal memory	Immediate verbal recall			p<.05						NS					
Paragraph recall-immediate	Verbal memory	Immediate verbal recall			p<.05		p<.01				NS					
Selective Reminding-Immediate	Verbal memory	Immediate verbal recall									NS		p<.01		NS	
Visual Paired Associates-Delayed	Visual memory	Delayed visual recall									NS					
Visual Paired Associates-Immediate	Visual memory	Immediate visual recall									NS					
Benton Visual Retention	Visual memory	Visual memory		NS						NS				p=.05		
Figural Memory	Visual memory	Visual memory									NS					
Visual Memory Span	Visual memory	Visual memory									NS					
Visual reproduction (WMS)	Visual memory	Visual memory			NS						NS				NS	

Sz: Shaywitz, PK: Polo-Kantolo, PS: Phillips & Sherwin, DF: Ditkoff, Sw: Sherwin, FF: Fedor-Freybergh, HG: Hackman & Galbraith, VD: Vanhulle & Demol, CS: Carlson & Sherwin, Mw: Matthews; JB: Jacobs, RK: Resnick, BC: Barrett-Connor, FK: Funk

Table 2. Summary of Cognitive Test Results

Cognitive function	Positive tests[1]/ total tests[2]	Subject profile in studies with positive tests	Explanation of results
Memory			
Memory battery	1/1	Symptomatic	
Immediate verbal recall	4/8	Symptomatic	2/3 studies found effects on paragraph recall; Positive results on other tests were not confirmed by other studies
Delayed verbal recall	3/8	Asymptomatic/Symptomatic	2/3 studies found effects on selective reminding; Positive results on the other test was not confirmed by a second study
Visual memory	1/9	Not stated	Two other studies did not confirm that users did better on this same test
Attention			
Working memory	0/4		
Complex attention	2/8	Symptomatic	One of the positive tests was only borderline significant (p=.08)
Mental tracking	2/13	Symptomatic	Positive results on these two tests were not confirmed by other studies
Vigilance	3/5	Symptomatic	Two of the positive tests were in same study; Third test was only borderline significant (p=.07)
Concept formation and reasoning	1/2	Symptomatic	Lack of sensitivity does not explain negative test
Motor speed	2/3	Symptomatic	One of the positive tests was not confirmed by another study; Second test used by only one study
Mental status	0/2		
Verbal functions and language	0/3		
Learning ability	0/1		

[1] Positive test indicates that women using estrogen scored significantly higher (at 0.10 significance level) than nonusers.
[2] 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.

Table 3. Influence of Symptoms on Results of RCTs

Subject profile	References	Results
Symptomatic-Subjects had various complaints, including hot flashes, sleep problems, depression, fatigue, chest pressure	Vanhulle-Demol, 1976; Hackman, 1976; Fedor-Freybergh, 1977; Sherwin, 1988; Phillips, 1988	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	Ditkoff, 1991	No improvement in cognition in women using estrogen
Not known if subjects were symptomatic	Polo-Kantolo, 1998; Shaywitz, 1999	Neither study found an improvement in cognition in women using estrogen

Table 4. Summary of Meta-analysis Results

Studies included	Fixed Effects		Random Effects		Test of Homogeneity X2 (p value)
	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	
All studies	0.66	0.55-0.78	0.66	0.53-0.82	11.92 (p>0.10)
Case control only	0.69	0.58-0.83	0.71	0.55-0.92	11.04 (P>0.10)
Cohort only	0.5	0.3-0.77			0.87 (P>0.10)
AD only-All			0.68	0.51-0.89	
AD by NINCDS-R criteria-All			0.67	0.46-0.92	
AD by NINCDS-R criteria-Case control only			0.77	0.46-1.16	
All studies except Heyman	0.65	0.55-0.77	0.65	0.52-0.80	
Using confidence intervals for Heyman from Yaffe	0.67	0.57-0.79	0.68	0.53-0.84	
Using t distribution			0.67	0.53-0.81	
Using Yaffe as prior distribution			0.67	0.51-0.87	

AD= Alzheimer's Disease

Figure 1. Hormone Replacement Therapy to Prevent Cognitive Decline

Analytic Framework

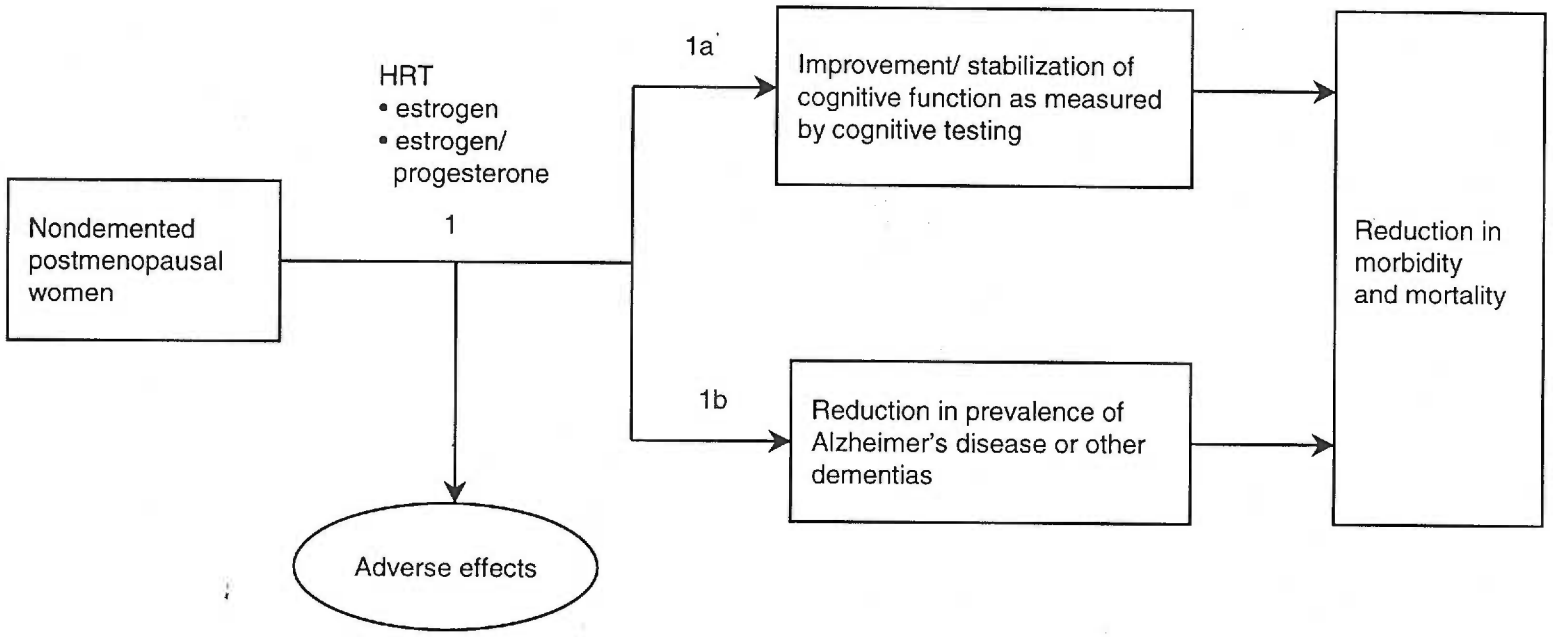


Figure 2. Key Questions in Analytic Framework: Hormone Replacement Therapy and Cognition

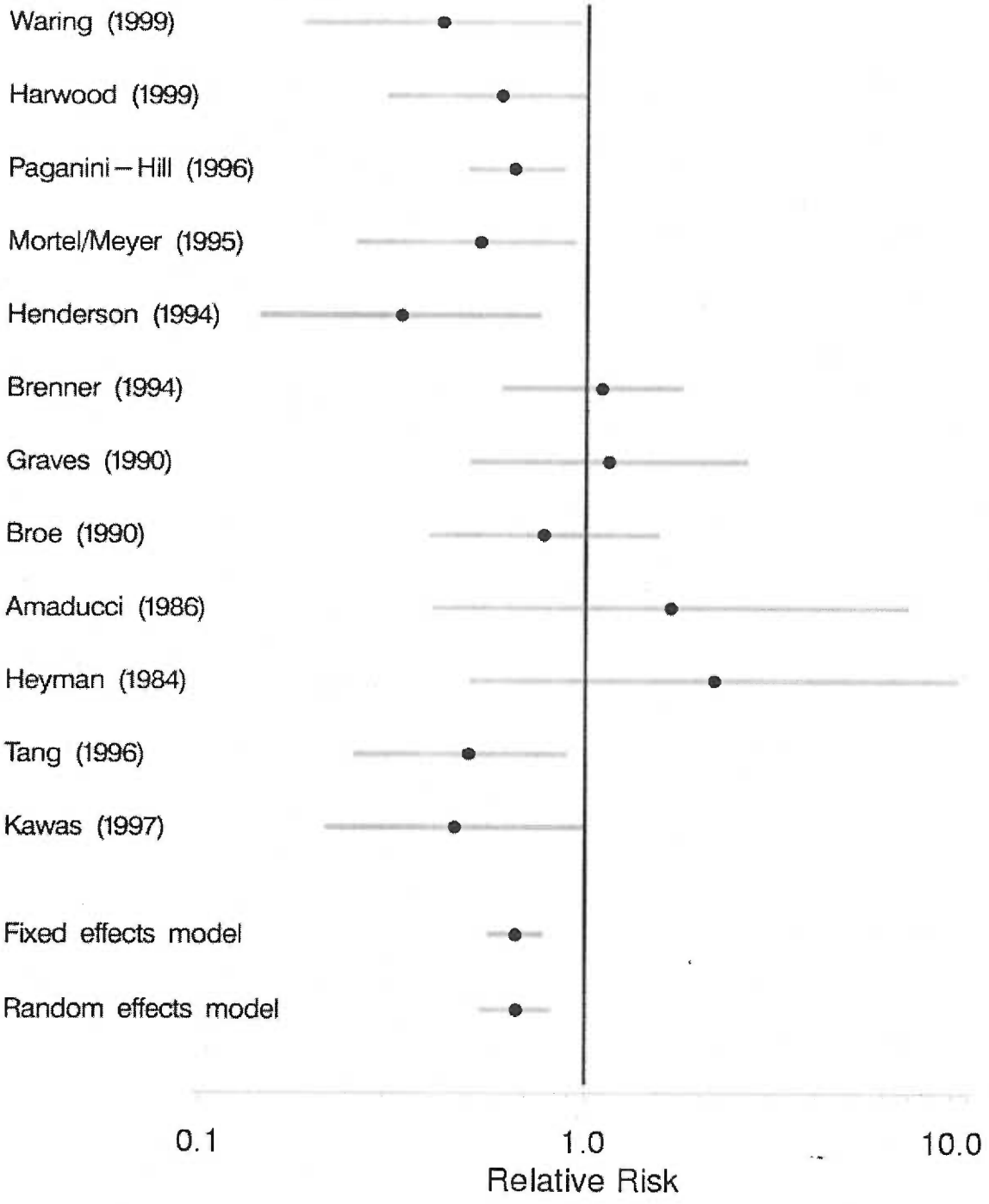
Arrow 1a. **Does HRT improve and/or stabilize cognition as measured by cognitive testing?**

1. Does the use of postmenopausal estrogen with or without progesterone improve or stabilize nondemented women's scores on cognitive testing?
2. What is the optimal dose and duration of use?

Arrow 1b. **Does the use of postmenopausal estrogen lower the risk of dementia?**

1. Does the use of postmenopausal estrogen with or without progesterone lower the risk of Alzheimer's Disease and other dementias?
2. What is the optimal dose and duration of use?

Figure 3. Results of Meta-analysis



Appendix 1. Search Strategy

Hormone Replacement Therapy Effect on Mental Processes

The topic of HRT and mental processes was searched in the Medline database including 1966 to January 2000.

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

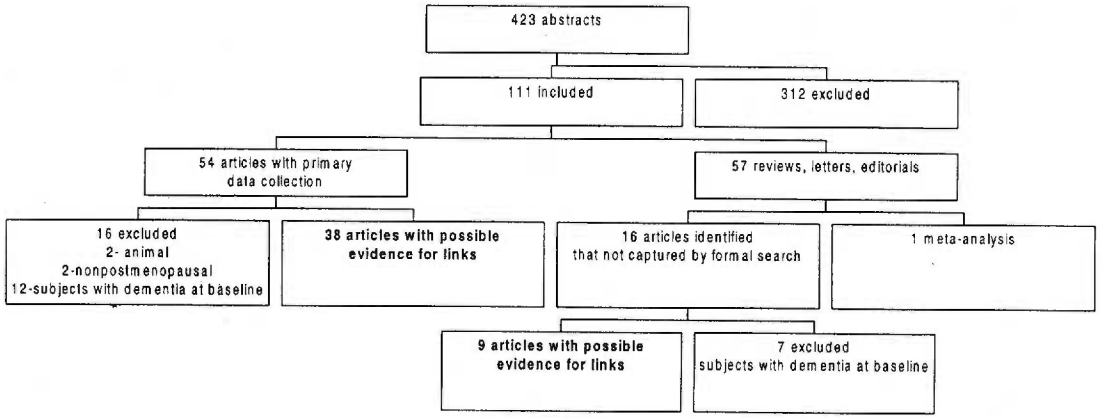
10 exp memory

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

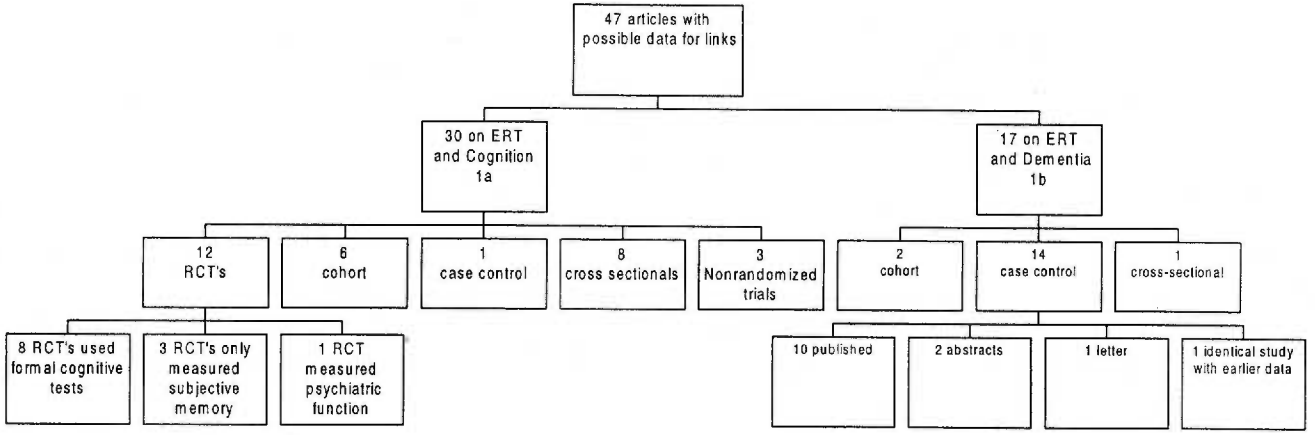
11 memory disorders

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 2a. HRT and Cognition--Search Results



Appendix 2b. HRT and Cognition--Articles Abstracted



Appendix 3. Inclusion/Exclusion Criteria

Title and Abstract Review--Exclusion Criteria

1. Non-human
2. Foreign language (unless key article)
3. Only looked 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 ERT and cognition (key question 1a) or ERT and dementia (any type of dementia) (key question 1b)
5. Review all meta-analyses

Evidence Tables—Inclusion Criteria "Best Evidence Approach"

1. For the association between ERT and Alzheimer's disease, studies which meet the following criteria:
   a. Cohort or case control study
   b. State dementia criteria (state how determined cases had dementia and controls did not)

2. For the association between ERT and cognition, studies which meet the following criteria:
   a. Randomized double blind placebo controlled study and cohort studies
   b. Objective measurement of cognition (not just subjective cognition)

3. If two studies of the same population, the most recent study with the most updated data will be included.

4. Peer reviewed published articles (no abstracts)

Appendix 4. Evidence Tables

Appendix 4 **Evidence Table 1. HRT and Cognition-Randomized Controlled Trials 1 of 10**

Author,year	Type of study	How recruited	Setting	Number of cases/ controls	Number originally randomized	Mean age (years)	Confirm menopause with FSH/ estradiol?	Percentage with surgical menopause	Symptomatic?
Shaywitz, 1999	Crossover	Paid volunteers from community	New Haven, CT	46	47	50.8 (33-61)	Yes	Not stated	Not stated
Polo-Kantola, 1998	Crossover	Volunteers-Newspaper ads	Finland	62	70	56.3 (47-65)	Yes	24% BSO[1]	Not stated
Phillips, 1992	Trial	Recruitment after TAH/BSO[1]	Canada	10 & 9	31	overall 48.2 +/- 4.7	s/p TAH[2]/BSO	100% s/p TAH/BSO	Yes-Placebo groups with more hot flashes; No difference in mood,depression, anxiety, hostility
Ditkoff, 1991	Trial	Not stated	East Los Angeles-Hispanic	24 & 12	Not stated	Overall 53 (45-60)	Yes	100% s/p TAH	No- <4 hot flash episodes for 2 weeks
Sherwin, 1988	Crossover	Recruitment after TAH/BSO (cases) or TAH (controls)	Canada	10 & 10 & 10 & 10	59	45.4	Yes	Cases-100% s/p TAH/BSO	Not specifically stated but most were likely symptomatic as s/p TAH/BSO

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Evidence Table 1. HRT and Cognition-Randomized Controlled Trials 2 of 10

Author,year	Type of study	How recruited	Setting	Number of cases/ controls	Number originally randomized	Mean age (years)	Confirm menopause with FSH/ estradiol?	Percentage with surgical menopause	Symptomatic?
Fedor-Freybergh, 1977	Trial	Recruitment at outpatient clinic	Gynecol-ogy clinic in Stock-holm	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	Trial	Not stated	England	9 & 9	Not stated	29-68	No	44% s/p TAH/BSO < 6 months ago	Yes
Vanhulle, 1976	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.

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Evidence Table 1. HRT and Cognition-Randomized Controlled Trials 3 of 10

Author,year	ERT form	Length of follow-up	Follow-up rate	Percentage compliant	Possible confounders?	Method of measuring outcome
Shaywitz, 1999	Oral-CEE[1] 1.25 mg/day	21 day course with 14 day break	100%	94%	Not measured	1.Interview 2. fMRI[2]
Polo-Kantola, 1998	Transdermal- If <56-Estrogel 0.6 mg/g (2.5 g/d), If >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	10 mg of estradiol valerate IM q month	2 months	100%	61%	No difference on anxiety, depression, hostility scores	Interview
Ditkoff, 1991	Oral CEE .625(12) or oral CEE 1.25(12) qd for 25 day/month	3 months	?100%	?100%	Users had improved depression (not controlled)	Interview
Sherwin, 1988	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	8 months (2-3 month treatment periods)	100%	85%	No difference in: baseline scores, education, occupation, personality inventory	Interview

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Author,year	ERT form	Length of follow-up	Follow-up rate	Percentage compliant	Possible confounders?	Method of measuring outcome
Fedor-Freybergh, 1977	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	1.5 mg piperazine oestrone sulphate BID	6 months	100%	Not stated	None	Interview
Vanhulle, 1976	4 mg estriol daily	3 months	100%	90%	Age	Interview

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Author,year	Outcome	Results	Analysis
Shaywitz, 1999	Verbal working memory tasks Nonverbal working memory tasks Brain activation	Not Significant Not Significant Significant	Intention to treat
Polo-Kantola, 1998	Simple reaction time Multistep reaction time Subtraction test Statement Verification Test Auditory Serial Addition Digit span Digit symbol Benton Visual Retention Letter cancellation Stroop color word test	Not Significant Not Significant Not Significant Not Significant Not Significant Not Significant Not Significant Not Significant Not Significant Not Significant	Only those that complete study
Phillips, 1992	Digit Span (WMS) Paragraph recall-immediate Paragraph recall-delayed Associate learning (WMS)-Immediate Associate learning (WMS)-Delayed Visual reproduction (WMS)	Not Significant p<.05 Not significant p<.05 p<.05 Not Significant	Only those that complete study
Ditkoff, 1991	Digit Span (WAIS) Digit Symbol (WAIS)	Not Significant Not Significant	Not clear
Sherwin, 1988	Digit Span Clerical Speed & Accuracy Paragraph Recall Test Abstract Reasoning	p<.01 p<.01 p<.01 p<.01	Only those that complete study and compliant

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Author,year	Outcome	Results	Analysis
Fedor-Freybergh, 1977	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 est vs placebo);p<.01 (for change est) p<0.01 (simple); p<.001 (w/ memory load) p<0.01 (simple);p<.001 (w/ interference) p<0.01 (time); p<0.001 (errors) p<0.05 (time); p<0.001 (errors)	
Hackman, 1976	Guild Memory Test	p<.02	Not reported
Vanhulle, 1976	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)	"no significant differences" "no significant differences" "no significant differences" "no significant differences" "no significant differences" "no significant differences" "no significant differences" p=.07 "no significant differences"	Only those that complete

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Author,year	Further explanation of main differences
Shaywitz, 1999	Increased activation of certain brain areas during verbal storage
Polo-Kantola, 1998	
Phillips, 1992	Immediate paragraph recall--users had improvement in score;no change in placebo Associate learning--users stayed the same; placebo had decline
Ditkoff, 1991	
Sherwin, 1988	Scores of all treatment groups higher than placebo; Scores of treatment groups dropped during the placebo month (p<.01)

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Author,year	Further explanation of main differences
Fedor-Freybergh, 1977	Estrogen users improved but placebo with no change
Hackman, 1976	Estrogen users improved but placebo with no change
Vanhulle, 1976	Compared difference in means of both groups (pretest-posttest). Estrogen users had greater improvement in scores in vigilance and attention.

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Author,year	Additional Information on Study Methodology	Jadad Score
Shaywitz, 1999	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	Mean serum estradiol levels lower with patch (190 vs 431 pmol/u)	4
Phillips, 1992	Compared pre and post scores. Did not compare scores of the estrogen and placebo groups	4
Ditkoff, 1991	Hispanic (American born); Compared pre and post scores. Did not compare scores of the estrogen and placebo groups	4
Sherwin, 1988	Only 10 women in estradiol treatment alone; Likely many women with symptoms	3

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Author,year	Additional Information on Study Methodology	Jadad Score
Fedor-Freybergh, 1977	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 est and placebo same unless specified	3
Hackman, 1976	Did not compare estrogen and placebo groups; No correlation between subjective improvement in memory and Guild Memory Test Score; 10/18 identified because of menopausal signs or symptoms; Sherwin reanalyzed and did not find statistically significant result; 3 estrogen users had large improvement and 1 had large deterioration; More variation in estrogen group than in nontreated group	2
Vanhulle, 1976	Compared difference in means of both groups (pretest-posttest). When compared post-tests, only attention was different ($p<.03$ for unadjusted analysis). Change in overall health score--no change in yes but did have change in no's;	3

1Bilateral salpingo-oophorectomy
2Total abdominal hysterectomy

Appendix 4 **Evidence Table 2: HRT and Cognition--Cohort Studies 1 of 8**

Author, year	Setting	ERT (n)	Non-user (n)	Eligibility	Mean Age (years) (User/ Nonuser)	Percentage with surgical menopause (User/ Nonuser)	Education (years) (User/ Nonuser)	Other differences
Carlson, 1999	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 socioeconomic status
Matthews, 1999	Study of Osteoporotic Fractures	Current-1325 Past-2612	5714	>64 yrs, Not Black, Able to walk w/o 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 smoke
Jacobs, 1998	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
Resnick, 1997	Baltimore Longitudinal Study of Aging	18	18	>39 yr; No dementia; Normal BVRT[1] at start; Short interval between ERT use and BVRT; No past use of ERT; No use of vaginal cream only	59.9/ 60.2	33.3/16.7	15.6/15.8	

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

Appendix 4 **Evidence Table 2: HRT and Cognition--Cohort Studies 2 of 8**

Author, year	Setting	ERT (n)	Non-user (n)	Eligibility	Mean Age (years) (User/ Nonuser)	Percentage with surgical menopause (User/ Nonuser)	Education (years) (User/ Nonuser)	Other differences
Barrett-Connor, 1993	Rancho Bernardo Cohort	394	406	>64 years, live in Rancho Bernardo	Overall 76.9	Not stated	2/3 completed college or more-not state if users differ	Users were less depressed
Funk, 1991	Veterans Administration longitudinal study of aging	30	77	Age 40-69, Caucasion	Overall 67	Not stated	Not stated	Users were more likely to smoke and drink more than 2 alcholic drinks per day

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

Appendix 4 **Evidence Table 2: HRT and Cognition--Cohort Studies 3 of 8**

Author, year	How determine use/ nonuse	Definition user	Definition nonuser	ERT form	Duration of ERT use	Length of follow-up	Follow-up rate	Counfounders controlled	Confounders not controlled
Carlson, 1999	Questionnaire	Current use	Not current use	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%	SES and years of education	Mood, Symptoms
Matthews, 1999	Interview	Ever use of oral estrogen at initial assessment	Never use of oral estrogen at initial assessment	Oral only	Current-14.3 yr Past-5.2 yr	4-6 years	77%	Age, Education, Activity limitations, Initial Performance	Mood, Symptoms, Medical problems
Jacobs, 1998	Questionnaire	Ever use of ERT	Never use of ERT	Any but most used unopposed oral CEE	Overall-4.55 yr	2.5 years	72%	Education, Age, Ethnicity	Mood, Symptoms, Medical problems
Resnick, 1997	Interview	Never user at first test and current user at time of follow-up test	Never user at both tests	Oral or transdermal	1--<=6 mo; 5--6 mo-1 yr; 10--1-5 yr; 2--5-10 yr	Not stated	Not stated	Age, Baseline BVRT score, Interval between assessments	Education, Mood, Symptoms, Med problems

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

Appendix 4 **Evidence Table 2: HRT and Cognition--Cohort Studies 4 of 8**

Author, year	How determine use/ nonuse	Definition user	Definition nonuser	ERT form	Duration of ERT use	Length of follow-up	Follow-up rate	Counfounders controlled	Confounders not controlled
Barrett-Connor, 1993	Interview; pill & prescription review	Ever use of oral estrogen at initial assessment	Never use of oral estrogen at initial assessment	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, Symptoms, Medical problems
Funk, 1991	Medical record review	Current use	Not current use	"Almost entirely unopposed" estrogen	Not stated	Maximum of 6 years	Not stated	Length of time since menopause	

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

Evidence Table 2: HRT and Cognition--Cohort Studies 5 of 8

Author, year	How outcome determined	Results[2]	Trends	Differences
Carlson, 1999	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	NS NS NS NS NS p<.01 NS NS NS NS NS NS NS	None studied	Users had improvement in scores on delayed selective reminding but non-users had decreased scores
Matthews, 1999	Modified MMSE[3] Trails B Digit Symbol	Not significant Not significant Not Significant	Past users had more benefit than current users	Only past users exhibited smaller decline in MMSE (p=0.03) and Trails B (p=0.02); Current users did not differ from nonusers
Jacobs, 1998	Immediate Selective Reminding Delayed Selective Reminding Similaries Subtest Boston Naming Test	p<=0.01 p<=0.001 Not Significant Not Significant	None studied	Selective Reminding Test-Users w/ improved scores while nonusers scores declined; No difference in scores over time on other tests
Resnick, 1997	BVRT[1]	p=0.05	None studied	Users with stable number of errors over time compared to increased number of errors in nonusers

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

Evidence Table 2: HRT and Cognition--Cohort Studies 6 of 8

Author, year	How outcome determined	Results[2]	Trends	Differences
Barrett-Connor, 1993	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<.01)	No difference in age related decrease in cognitive function in current or past users
Funk, 1991	Cognitive Capacity Screening Examination (CCSE)	Not Significant		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

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

Evidence Table 2: HRT and Cognition--Cohort Studies 7 of 8

Author, year	Comments
Carlson, 1999	Analysis based on only 10 users and 27 nonusers
Matthews, 1999	When only looked at those who were consistent current, past, or never users, results not changed
Jacobs, 1998	Level of depression did not differ by ERT use history
Resnick, 1997	

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

Evidence Table 2: HRT and Cognition--Cohort Studies 8 of 8

Author, year	Comments
Barrett-Connor, 1993	
Funk, 1991	Study included women with a history of transient ischemic attacks or reversible ischemic neurologic deficits 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 has history of RIND or TIA

1Benton Visual Retention Test
2Results are for current or ever users (not past users)
3Mini-mental status exam

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	Rochester, Minnesota-Population based	222/222	AD[1]	Diagnostic criteria "equivalent" to NINCDS-ADRDA[2]	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	AD center-Miami 30% Hispanic	White -229/139 Hispanic-133/53	AD	NINCDS-ADRDA	Were evaluated at AD center	Age >=65, Normal MMSE[3], Normal 4 trial recall of 3 words in MMSE
Paganini-Hill, 1996	Nested in Leisure World Cohort of 8877 women-Retirement Community California-High Socioeconomic status	248/1198	AD	Dementia diagnosis listed on death certificate; exclude multi-infarct dementia or dementia from another likely cause	Death certificate or National Death Index list AD, "senile dementia," "dementia," "senility"	No mention of dementia on death certificate
Mortel, 1995	Baylor College of Medicine and Houston VA	93/148	1. AD 2. IVD[4]	1.NINCDS-ADRDA 2. State of California Alzhemimer's Disease Diagnostic and Treatment Centers Criteria	Referral by local physicians and support groups	Neurological examination and neuropsychologic assessment
Henderson, 1994	AD research center California	143/92	AD	NINCDS-ADRDA (70 confirmed with autopsy)	Volunteers recruited from community outreach who meet criteria for AD by history, exam, lab	Neurological examination and detailed neuropsychologic assessment

[1]Alzheimer's disease
[2]National Institute of Neurological and Communicative disorders and Stroke and the Alzheimer's Disease and Related Disorders Association Criteria for Alzheimer's Disease
[3]Mini mental status exam
[4]Ischemic vascular disease

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	HMO, Washington	107/120	AD	NINCDS-ADRDA	AD patient registry in HMO	MMSE score of at least 28 and no evidence of AD on psychometric evaluation, chart review, judgement of study RN
Graves, 1990	AD referral center in Washington State	60/60 women (130/130 total)	AD	NINCDS-ADRDA	90% from AD center and 10% from VA	No memory loss (not stated how determine this)
Broe, 1990	AD referral center in Sydney, Australia	106/106 women (170/170 total)	AD	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	Neurology departments in Italy	60 cases /60 hospital and 50 community controls (116/116/97 total)	AD	Blessed dementia scale; 2 signs or symptoms of cognitive decline; no depression; no evidence for dementias other than AD by history, exam, testing	Admission to neurology deparments of seven centers	Blessed dementia scale
Heyman, 1984	Duke Medical Center, North Carolina	28 female cases/ 56 female controls (40/80 total)	AD	"Rigorous criteria"	Participants of another comprehensive study of Alzheimer's 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	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 years/ 12 years	Cases with less breast cancer; No difference in measures of endogenous estrogen exposure
Harwood, 1999	85% recruited for free memory screening; 9% evaluated at AD 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 hyptertension but not statistically significant
Paganini-Hill, 1996	Death certificates-Matched on year of death and year of birth	87.7/87.3	Not stated	Not stated	
Mortel, 1995	Friends and relatives	73.7(AD)/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
Henderson, 1994	Volunteers recruited from community outreach in whom AD 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	Stratified random sample of HMO matched within 2 yr	78.7/76.6	22/9	Percentage with > 12 yrs: 34.6/60.8	
Graves, 1990	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 h/o dementia
Broe, 1990	Clinic controls-matched for sex and age within 2 years	78.6/78.7 (men & women)	Not stated	Not stated	Cases more likely to have first-degree relative with h/o dementia
Amaducci, 1986	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	13.7 / 9.6 / 8.3 had oophorectomy	No significant association found between education/literacy and case/control status	Cases more likely to have first-or second-degree relative with h/o dementia
Heyman, 1984	Population controls--random digit dialing; matched for sex, race, and 5 year age interval	60.8 (51-71)	Not stated	49% of cases and 22% of controls had education beyond high school	Cases had greater h/o prior thyroid disease and h/o severe head disease

Author, year	Response rate	How ERT use determined	Definition of ERT use	Definition of Nonuser	ERT form/dose	Duration of ERT use
Waring, 1999	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 AD	Never use	90% used oral +/ - topical; most CEE[5]	Not stated
Harwood, 1999	Not stated	Cases- proxy interview Controls-self interview	Ever use	Never use	Not stated	Median 2 years
Paganini-Hill, 1996	61% response to questionnaire	Questionnaire prior to death-85% complete >=5 years before death	Ever use	Never use	Any	1981-95
Mortel, 1995	Not stated	Medical records, questionnaires and interviews-surrogate used for patient with dementia	Current user	Not current user	Not stated	Not stated
Henderson, 1994	Not stated	Cases- proxy interview Controls-self interview	Current user	Not current user	Any (>81% oral CEE)	Not stated

[5] Conjugated Equine Estrogen

Author, year	Response rate	How ERT use determined	Definition of ERT use	Definition of Nonuser	ERT form/dose	Duration of ERT use
Brenner, 1994	Not stated	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	Screened 800 medical records and 188 met criteria; 143 entered study	Proxy telephone interview (88% spouse of >10 years)	"Estrogen replacement" use prior to symptoms	No "estrogen replacement"	Not stated	Not stated
Broe, 1990	Screened 333 to obtain 170 cases; Screened 270 to obtain 170 controls	Proxy interview (>85% were spouse or 1st degree relative)	"Hormonal treatment"	No "hormonal treatment"	Not stated	At least 6 months
Amaducci, 1986	Screen 152 admissions to get 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	Not stated	Proxy interview	Current use of "estrogen replacement"	Not current user	Not stated	Not stated

Author, year	Counfounders controlled	Confounders not controlled	Number cases & controls that used ERT	Adjusted OR (95% CI)	Significance	Trends
Waring, 1999	Age, education, length of time in linkage system	Mood, Symptoms, Medical problems, Ethnicity	9 & 20	0.42 (0.18-0.96)	p=0.04	No decreased risk with use less 6 months; No cumulative dose effect
Harwood, 1999	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	Age, Weight, Blood pressure medication, Weight, Menopause type, Age, LMP, Age 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 (signif only for > 5 years)
Mortel, 1995	For analysis of AD and IVD: none For analysis of all dementia: age	Mood, Education, Symptoms, Ethnicity	11 (AD) & 7 (IVD) & 29 (Controls)	AD: 0.55(0.26-1.16) IVD: .0.50(0.26-1.2) All: 0.53(0.25-0.94)	Not significant	None stated
Henderson, 1994	Education, Age	Mood, Symptoms	7 & 18	0.33 (0.14-0.76)[6]	p=0.01	

[6]Odds ratios and confidence intervals are unadjusted and calculated from data in tables.

Author, year	Counfounders controlled	Confounders not controlled	Number cases & controls that used ERT	Adjusted OR (95% CI)	Significance	Trends
Brenner, 1994	Age, history of hysterectomy before & after age 55, education, ethnicity	Mood, Symptoms	52 & 58	1.1 (0.6-1.8)	Not significant	Oral only: 0.7 (0.4-1.5); Current 0.6 (0.3-1.2); No cumulative dose trend
Graves, 1990	Age	Education, Medical problems, Symptoms, Mood	11 & 10	1.15 (0.50-2.64)	Not significant	
Broe, 1990	Age	Education, Medical problems, Symptoms, Mood	14 & 18	0.78 (0.39-1.56)	p=0.48	
Amaducci, 1986	Age, area of residence	Education, Medical problems, Symptoms, Mood	6 & 4	1.67 (0.39-6.97)[7]	p=0.73	
Heyman, 1984	Age, race, education residence	Medical problems, Symptoms, Mood	4 & 4	2.17 (0.5-9.41)[8]	p>0.05	

[7]Matched odds ratios and confidence intervals calculated from data in paper using SAS.
[8]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
Waring, 1999	
Harwood, 1999	Study was of both men and women--info in table on both
Paganini-Hill, 1996	Did not exclude those with dementia at baseline; Earlier 1994 report had similar results)
Mortel, 1995	
Henderson, 1994	OR is unadjusted, calculated using data in table. Authors state that univariate analysis same as multivariate analysis.

Author, year	Comments
Brenner, 1994	
Graves, 1990	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; ERT only one of many risk factors studied; Kappa for agreement in reported ERT use between controls in the validation subsample and their surrogates was 0.64.
Broe, 1990	Study of both men and women; Not state if blinded interviewers; ERT only one of many risk factors studied; ; Kappa for agreement in reported ERT use between controls in the validation subsample and their surrogates was not specifically stated.
Amaducci, 1986	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; Does not define or specifically state estrogen replacement therapy (uses "hormonal therapy"); Only 52% of proxy respondents could answer question about estrogen use during menopause; Agreement in reported ERT use between controls in the validation subsample and their surrogates was not specifically stated but was greater than 60%.
Heyman, 1984	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; ERT only one of many risk factors studied; Kappa for agreement in reported ERT use between controls and their surrogates was 0.63.

Evidence Table 4. HRT and Alzheimer's Disease (AD)-Cohort Studies 1 of 3

Author, year	Cohort	How recruit participants	ERT (n)	Nonuser (n)	Mean Age (years) (User/ Nonuser)	Percentage with surgical menopause (User/	Education (years) (User/ Nonuser)	Other differences	Eligibility	Definition of user	How determine use
Kawas, 1997	Baltimore Longitudinal Study of Aging	Not stated	230	242	Overall 61.5-range 28-94/ No difference (data not given)	Not stated	No difference (data not given)	No difference in age menopause	Information on ERT/Full 16 yr follow-up	Ever use of oral or transdermal	Interview
Tang, 1996	Manhattan Study of Aging	Recruitment from Medicare and senior housing	156	968	Overall 74.2-users younger (p=0.01)	50/26.4 (p=0.001)	Users more educated (p=0.005)	Users: fewer Blacks, earlier menopause	No cognitive impairment at baseline/ Information on ERT	Ever use after menopause	Interview

1NINCDS-ADRDA-National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association Criteria for Alzheimer's Disease

Appendix 4

Evidence Table 4. HRT and Alzheimer's Disease (AD)-Cohort Studies 2 of 3

Author, year	ERT form	Duration of use	Neuro-psychol-ogical testing at baseline	Outcome	Criteria	How determine outcome	Length of follow-up	Follow-up rate	Counfounders controlled
Kawas, 1997	Oral (212)/ Patch (18)	Not stated	Yes	AD	NINCDS-ADRDA[1]	Multidisciplinary evaluations every 2 year including neuropsychological assessments	16 years	Not stated	Education, Age
Tang, 1996	"Majority"-Conjugated Equine Estrogen	2 mo-49 yrs (average 6.8 yr)	Yes	AD	NINCDS-ADRDA[1]	Medical records and imaging studies and data from initial and follow-up study examinations	1-5 years	84%	Education, Age, Ethnicity

1NINCDS-ADRDA-National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association Criteria for Alzheimer's Disease

Evidence Table 4. HRT and Alzheimer's Disease (AD)-Cohort Studies 3 of 3

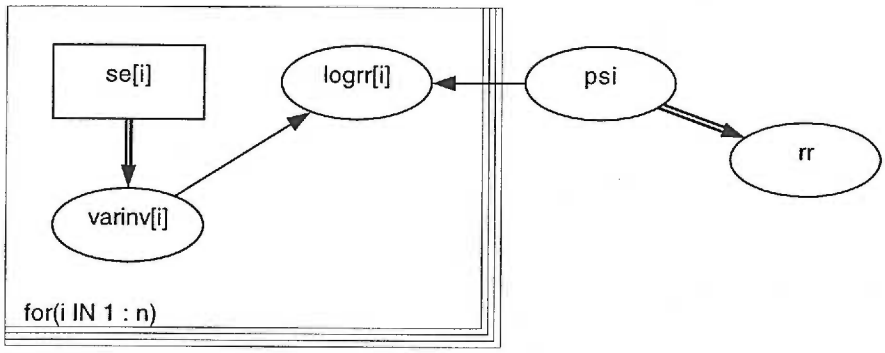
Author, year	Confounders not controlled	Number of users and nonusers with AD	Adjusted RR (95% CI)	Trends	Comments
Kawas, 1997	Mood, Symptoms, Medical problems	9 & 25	0.457 (0.209-0.997)	No duration effect	
Tang, 1996	Mood, Symptoms, Medical problems	9 & 158	0.5 (0.25-0.9)	RR 0.13 for users > 1 yr/ users with later age of onset	Women who did not remember ERT classified as nonusers

1NINCDS-ADRDA-National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association Criteria for Alzheimer's Disease

Appendix 5. Meta-Analysis Results

Fixed effects model

The fixed effects model assumes that each study is estimating one true population effect. This assumption can only be made if the studies are homogeneous. The X2 value for the test of homogeneity is 11.92 with 11 degrees of freedom, which corresponds to a p value of greater than 0.10. This suggests that the studies are homogeneous and the fixed effects model can be used. The graphical data model is shown in the diagram below. Psi is the population effect that each study is estimating. We assume that the log RR are drawn from a normal distribution. If y[i] denotes the log RR for the ith study, then y[i] is normally distributed with mean psi and varance SE[i]2.



The BUGS statements that correspond to the model are given below:

MODEL>>

```
model;
{
  for( i in 1 : n ) {
    logrr[i] ~ dnorm(psi,varinv[i])
  }
  for( i in 1 : n ) {
    varinv[i] <- 1 / (se[i] * se[i])
  }
  psi ~ dnorm( 0.0,1.0E-6)
  rr <- exp(psi)
}
```

DATA>>

```
list(
  n       =12,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, -0.69, -0.78, 0.77),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73,  0.30, 0.40, 0.75)
)
```

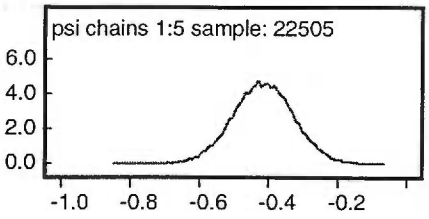
INITS>>

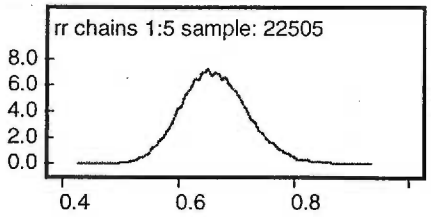
```
#chain 1
list(psi = 0)
#chain 2
list(psi = 2)
#chain 3
list(psi = -2)
#chain 4
list(psi = 8)
#chain 5
list(psi = -8)
```

Results: The results from 22500 draws (5 chains of 5000 iterations) after a burn in of 2500 draws (500 iterations) are shown below. A noninformative normal prior was used.

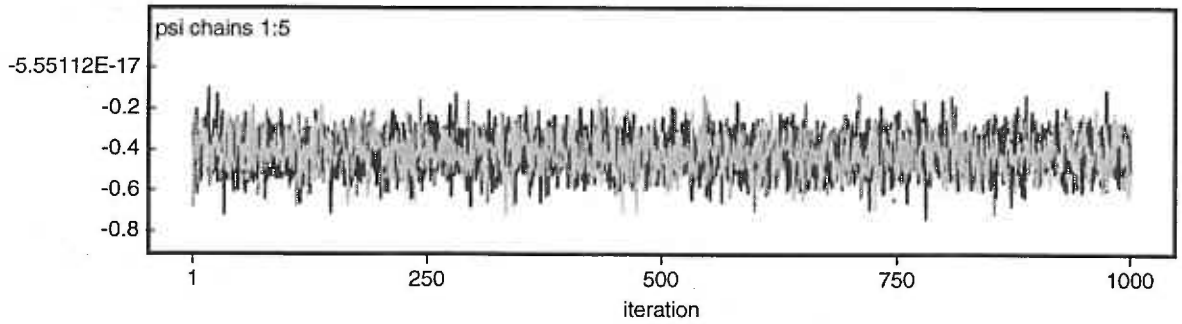
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.4147	0.08582	0.001305	-0.5831	-0.4148	-0.2456	500	22505
rr	0.663	0.05702	8.71E-4	0.5582	0.6605	0.7823	500	22505

The density plots of psi and rr are shown below. They show that all of the 22500 draws resulted in a odds ratio of less than one.

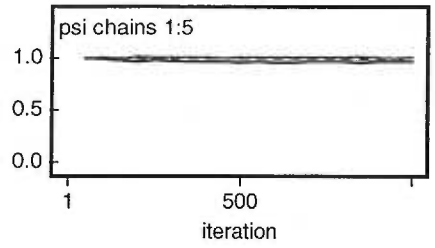


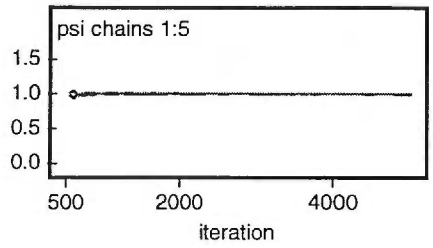


Sensitivity analysis: Initial values are required to start the analysis. We can examine if different starting vaues for psi result in similar posterior distributions. The goal is to have convergence to a solution that is independent of the initial values. As shown under the *INITS>>* statement above, we run five Markov chains from different initial values. The figure below shows the value of psi for the first 1000 iterations (including the burn in period).

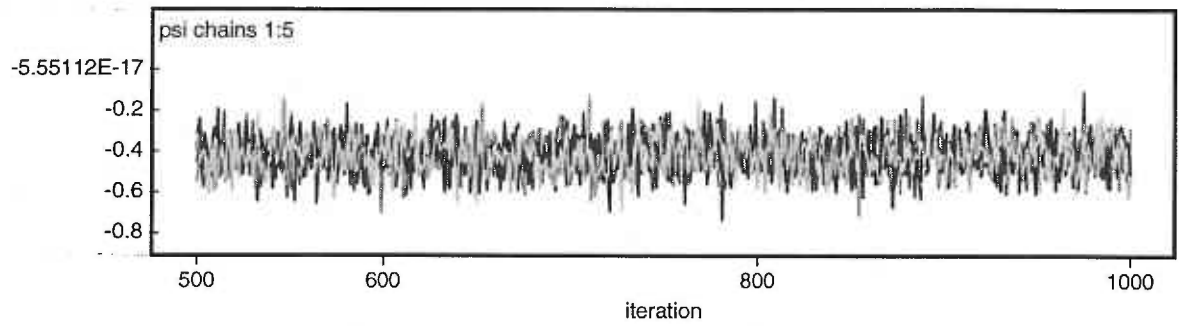


This figure seems to show good convergence, but if we look at the Gelman-Reiter diagram, convergence seems to occur by iteration 500:

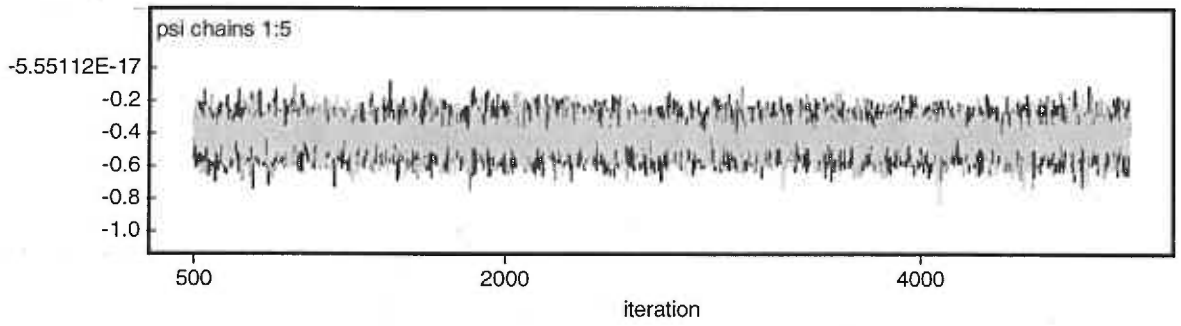




The figure below shows, in more detail, the chains for psi, exluding the 500 burn-in period. It shows that there is good convergence.



The entire history for the iterations that are used is shown below:



Therefore, even if we start with very different initial values for psi, we get siilar posterior distributions.

The confidence intervals for the Heyman 1984 study are uncertain because they were not given by the authors. The raw data was used to calculate the unadjusted odds ratio and confidence intervals and this data was used in the above analysis. However, a previous meta-analysis reported different confidence intervals that they state was obtained from the X2 value; however, we were unable to find this X2 value and unable to contact the original authors. To determine if these different values for the confidence intervals affected the results, the above analysis was repeated using the data from the previous meta-analysis. The data and results are shown below:

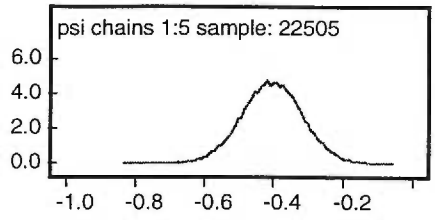
DATA>>

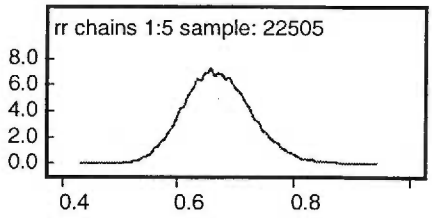
```
list(
  n       =12,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, -0.69, -0.78, 0.87),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73,  0.30, 0.40, 0.60)
)
```

Results>>

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.4038	0.0855	0.0013	-0.5716	-0.4039	-0.2353	500	22505

rr	0.6702	0.05743	8.772E-4	0.5646	0.6677	0.7903	500	22505





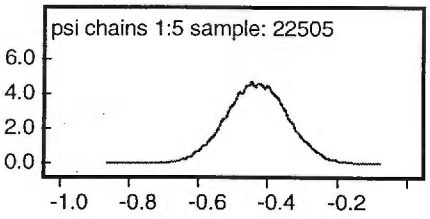
This shows that using the Yaffe data does not change the estimates significantly. The analysis was also done excluding the Heyman study. The data and results are shown below:

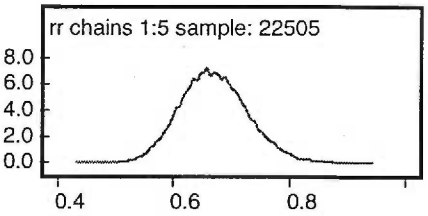
DATA>>

```
list(
  n      =11,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, -0.69, -0.78),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73,  0.30, 0.40)
)
```

Results>>

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.4307	0.0864	0.001313	-0.6002	-0.4308	-0.2604	500	22505
rr	0.6525	0.0565	8.63E-4	0.5487	0.65	0.7707	500	22505





This shows that excluding the study does not significantly change the results.

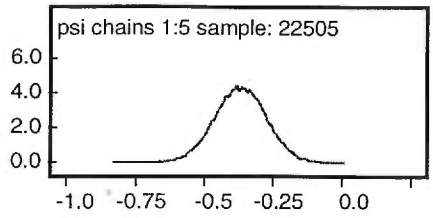
Because of concerns about combining the results of different study designs, the analysis was repeated using just case control studies. The data is shown below:

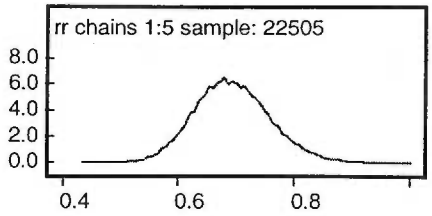
DATA>>

```
list(
  n      =10,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, 0.77),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73, 0.75)
)
```

Results: The results from 22500 draws (5 chains of 5000 iterations) after a burn in of 2500 draws (500 iterations) are shown below. These are results using a noninformative normal prior.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.3686	0.09202	0.001399	-0.5492	-0.3687	-0.1872	500	22505
rr	0.6947	0.06408	9.792E-4	0.5774	0.6916	0.8292	500	22505





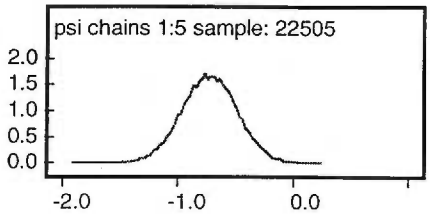
The density plots of psi and rr show that all of the 22500 draws resulted in a odds ratio of less than one.

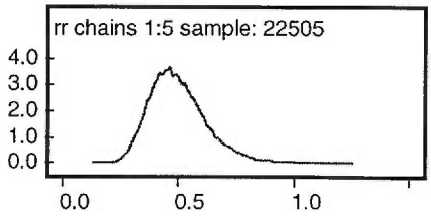
The analysis was also done on the two cohort studies using the data below:

```
list(
  n      =2,
logrr = c(-0.69, -0.78),
se    = c(0.30, 0.40)
)
```

The results are as follows:

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.7233	0.2378	0.003615	-1.19	-0.7237	-0.2548	500	22505
rr	0.499	0.1205	0.001854	0.3042	0.485	0.7751	500	22505





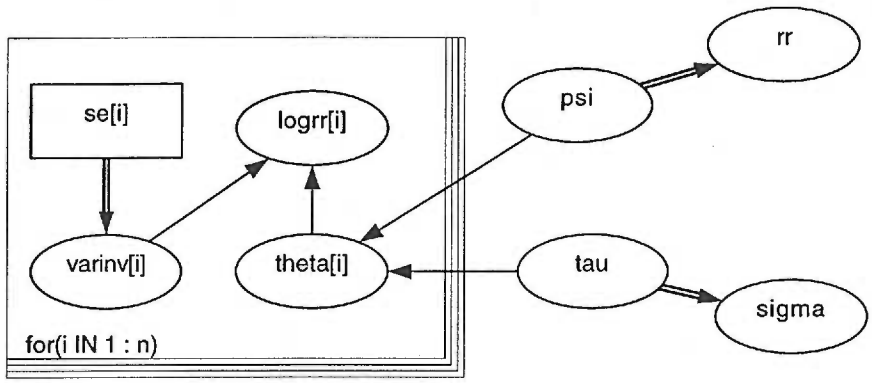
The density plot shows that a few of the draws resulted in an odds ratio of greater than one, although most were less than one and the 95% CI was less than one.

In summary the results of the fixed effects model are shown below:

Studies used	OR	Lower CI	Upper CI
All studies--Unadjusted Heyman data	0.66	0.55	0.78
All studies except Heyman	0.65	0.55	0.77
All studies using Heyman data fromYaffe	0.67	0.57	0.79
Case control only	0.69	0.58	0.83
Cohort only	0.5	0.3	0.77

Random effects model

Although the test of homogeneity suggests that the studies are homogeneous, a random effects model is helpful because it is more conservative and allows us to test the robustness of the fixed effects results. Unlike the fixed effects model which assumes a true population effect, a random effects model assumes that each study is estimating a separate effect from the other studies. The graphical data model is shown in the diagram below. Theta (i) represent the separate study effects and are considered random variates from a hyperdistribution. The hyperdistribution in this model is a normal distribution with mean, psi, and variance, 1/tau. We assume the log RR are drawn from a normal distribution. Therefore, if y[i] represents the log RR from the ith study, y[i] is normally distributed with mean theta [i] and variance SE[i]2. Theta[i] is also normally distributed with mean psi and variance 1/tau.



The BUGS statements that correspond to the model are given below:

MODEL>>

```
model;
{
  for( i in 1 : n ) {
    varinv[i] <- 1 / (se[i] * se[i])
  }
  for( i in 1 : n ) {
    logrr[i] ~ dnorm(theta[i],varinv[i])
  }
  for( i in 1 : n ) {
    theta[i] ~ dnorm(psi,tau)
  }
  psi ~ dnorm( 0.0,1.0E-6)
  tau ~ dgamma(0.001,0.001)
  rr <- exp(psi)
  sigma <- 1 / sqrt(tau)
}
```

DATA>>

```
list(
  n       =12,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, -0.69, -0.78, 0.77),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73,  0.30, 0.40, 0.75)
)
```

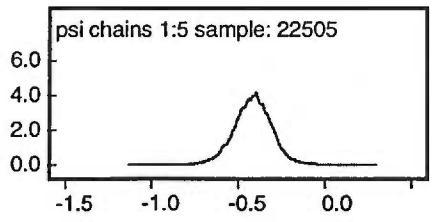
INITS>>

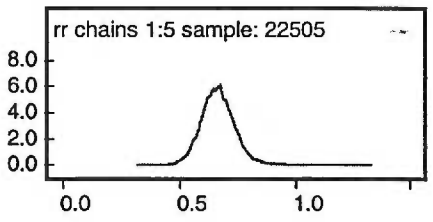
```
# chain 1
list(psi = 0,  tau = 10,  theta = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0))
# chain 2
list(psi = 15, tau = 1, theta = c(15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 15, 15))
#chain 3
list(psi = 5, tau = 1, theta = c(5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5))
#chain 4
list(psi = 4, tau = 10, theta = c(4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4))
#chain 5
list(psi = -4, tau = 10, theta = c(-4, -4, -4, -4, -4, -4, -4, -4, -4, -4, -4, -4))
```

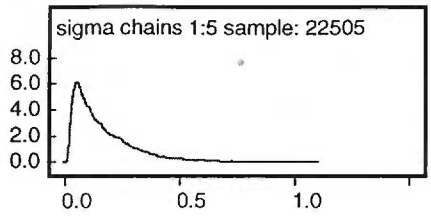
Results: The results from 22500 draws (5 chains of 5000 iterations) after a burn in of 2500 draws (500 iterations) are shown below. A noninformative normal prior on psi and a noninformative gamma prior on tau was used.

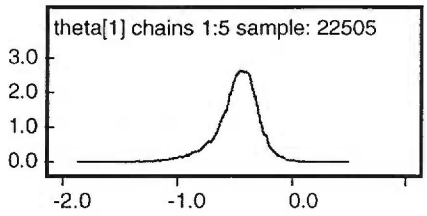
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.4141	0.1098	0.002538	-0.6324	-0.4128	-0.1974	500	22505
rr	0.6649	0.0737	0.001657	0.5313	0.6618	0.8209	500	22505
sigma	0.1602	0.1282	0.004152	0.02601	0.1215	0.5018	500	22505
theta[1]	-0.4808	0.195	0.004514	-0.9663	-0.4566	-0.1565	500	22505
theta[2]	-0.439	0.1521	0.003007	-0.7665	-0.4322	-0.1457	500	22505
theta[3]	-0.4211	0.1135	0.002331	-0.6494	-0.4201	-0.1977	500	22505
theta[4]	-0.4642	0.1658	0.003377	-0.8417	-0.4499	-0.1695	500	22505
theta[5]	-0.5152	0.212	0.005331	-1.054	-0.4784	-0.1956	500	22505
theta[6]	-0.2801	0.1899	0.004876	-0.5741	-0.3124	0.1805	500	22505
theta[7]	-0.3366	0.2021	0.004382	-0.6647	-0.3647	0.1648	500	22505
theta[8]	-0.3866	0.174	0.003353	-0.7194	-0.3939	-2.5E-4	500	22505
theta[9]	-0.3564	0.2214	0.004395	-0.7259	-0.3826	0.188	500	22505
theta[10]	-0.4748	0.1708	0.003699	-0.879	-0.4576	-0.1805	500	22505
theta[11]	-0.4717	0.1892	0.004123	-0.927	-0.4503	-0.1513	500	22505
theta[12]	-0.3437	0.2321	0.004654	-0.7134	-0.3758	0.2505	500	2250

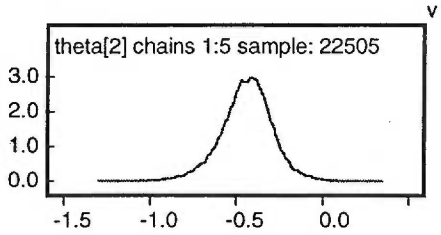
The results for psi and RR are similar to the means from the fixed effects model. As expected, the confidence intervals are slightly larger. The results show that the confidence interval for the rr did not cross one. The density plots of psi, rr, sigma (the standard deviation), and theta are shown below. The RR density plot shows that although most of the 22500 draws resulted in a odds ratio of less than one, a few draws had OR of up to approximately 1.3.

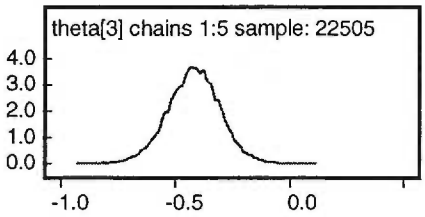


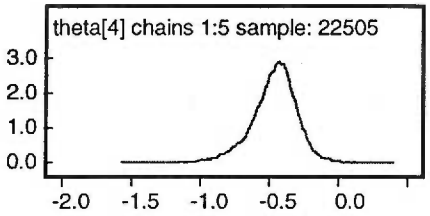


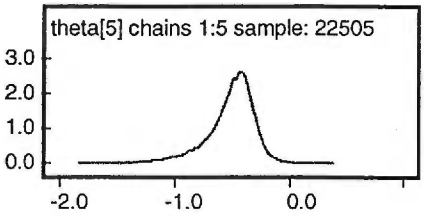

sigma chains 1:5 sample: 22505
8.0
6.0
4.0
2.0
0.0
0.0
0.5
1.0

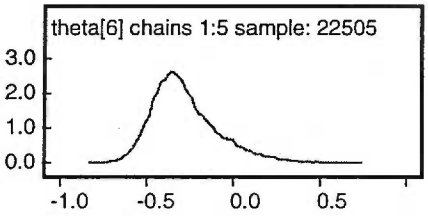

theta[1] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-2.0
-1.0
0.0

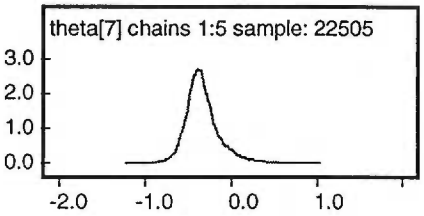

theta[2] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-1.5
-1.0
-0.5
0.0

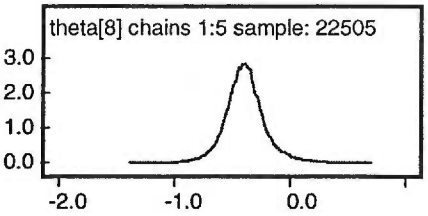

theta[3] chains 1:5 sample: 22505
4.0
3.0
2.0
1.0
0.0
-1.0
-0.5
0.0

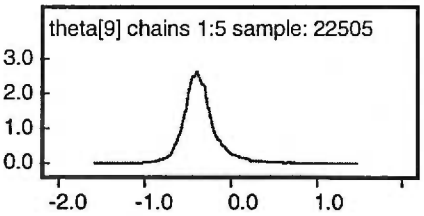

theta[4] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-2.0
-1.5
-1.0
-0.5
0.0

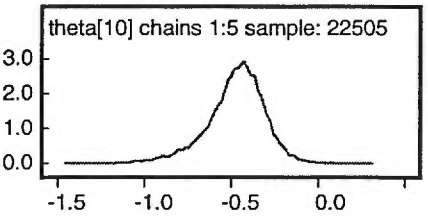

theta[5] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-2.0
-1.0
0.0

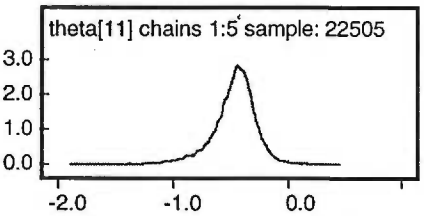

theta[6] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-1.0
-0.5
0.0
0.5

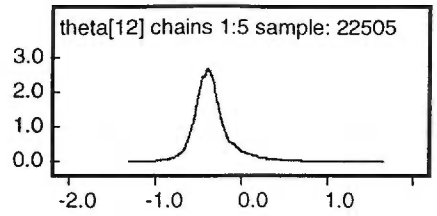

theta[7] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-2.0
-1.0
0.0
1.0


theta[8] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-2.0
-1.0
0.0

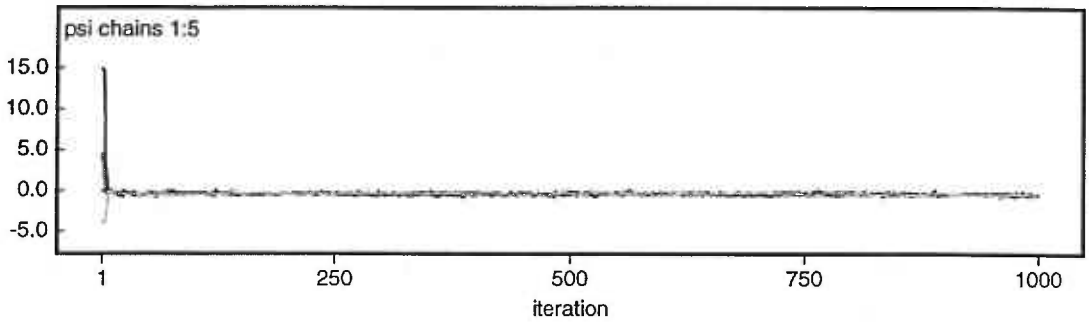

theta[9] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-2.0
-1.0
0.0
1.0


theta[10] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-1.5
-1.0
-0.5
0.0

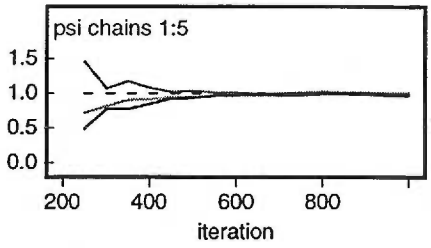

theta[11] chains 1:5 sample: 22505
3.0
2.0
1.0
0.0
-2.0
-1.0
0.0

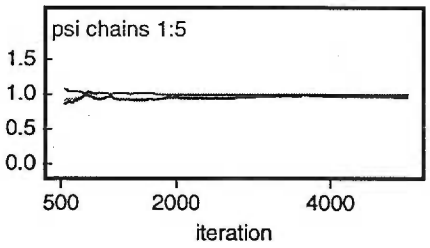


Sensitivity analysis: Initial values are required to start the analysis. We can examine if different starting vaues for psi result in similar posterior distributions. The goal is to have convergence to a solution that is independent of the initial values. As shown under the *INITS>>* statement above, we run five Markov chains from different initial values. The figure below shows the value of psi for the first 1000 iterations (including the burn in period).

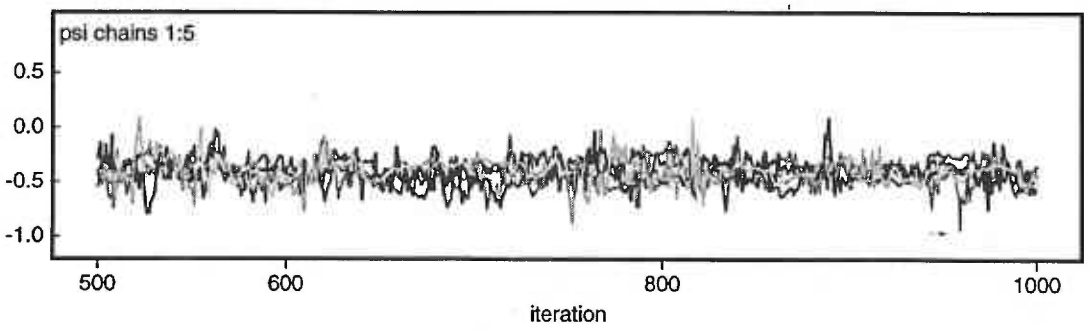


This figure seems to show good convergence, but if we look at the Gelman-Reiter diagram, convergence does not seem to occur until iteration 500:

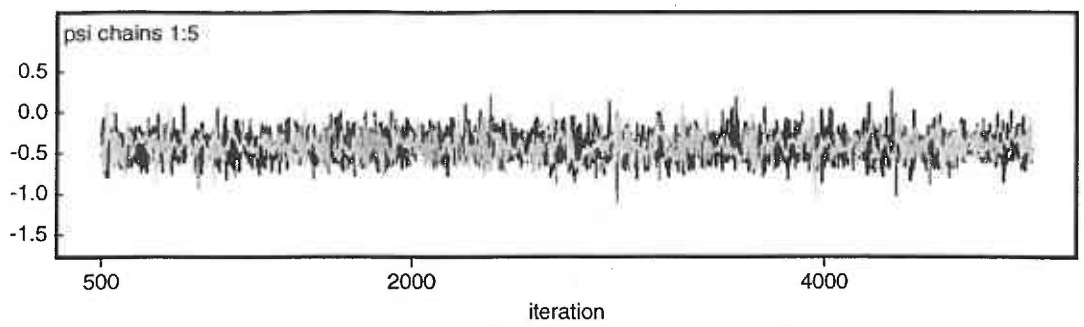




The figure below shows, in more detail, the chains for psi, exluding the 500 burn-in period. It shows that there is good convergence.



The entire history for the iterations that are used is shown below:



Therefore, even if we start with very different initial values for psi, we get similar posterior distributions.

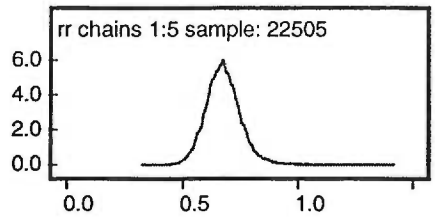
The confidence intervals for the Heyman 1984 study are uncertain because they were not given by the authors. The raw data was used to calculate the unadjusted odds ratio and confidence intervals and this data was used in the above analysis. However, a previous meta-analysis reported different confidence intervals that they state was obtained from the X2 value; however, we were unable to find this X2 value and unable to contact the original authors. To determine if these different values for the confidence intervals affected the results, the above analysis was repeated using the data from the previous meta-analysis. The data and results are shown below:

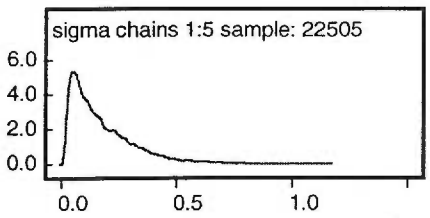
DATA>>

```
list(
  n       =12,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, -0.69, -0.78, 0.87),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73,  0.30, 0.40, 0.60)
)
```

Results>>

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.3987	0.1139	0.002469	-0.625	-0.3981	-0.1701	500	22505
rr	0.6756	0.07808	0.001646	0.5352	0.6716	0.8436	500	22505
sigma	0.1783	0.1428	0.00475	0.02668	0.1361	0.5556	500	22505
theta[1]	-0.479	0.2077	0.004722	-0.9986	-0.4508	-0.1356	500	22505
theta[2]	-0.4319	0.1582	0.002999	-0.7761	-0.4245	-0.1279	500	22505
theta[3]	-0.4143	0.1158	0.002279	-0.6481	-0.4124	-0.1872	500	22505
theta[4]	-0.4599	0.1737	0.003474	-0.8559	-0.4432	-0.1527	500	22505
theta[5]	-0.5188	0.2272	0.005674	-1.096	-0.475	-0.1806	500	22505
theta[6]	-0.2542	0.1977	0.005232	-0.5603	-0.2888	0.2171	500	22505
theta[7]	-0.312	0.2153	0.004773	-0.6586	-0.3454	0.2202	500	22505
theta[8]	-0.371	0.1827	0.003405	-0.7219	-0.3796	0.03663	500	22505
theta[9]	-0.3312	0.2403	0.004863	-0.7265	-0.3634	0.2694	500	22505
theta[10]	-0.4721	0.1797	0.003792	-0.8922	-0.4525	-0.165	500	22505
theta[11]	-0.4685	0.2008	0.004279	-0.9556	-0.4427	-0.1315	500	22505
theta[12]	-0.2715	0.2685	0.006369	-0.6441	-0.3296	0.4483	500	22505





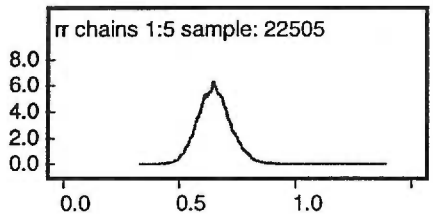
This shows that using the Yaffe data does not change the estimates significantly. The analysis was also done excluding the Heyman study. The data and results are shown below:

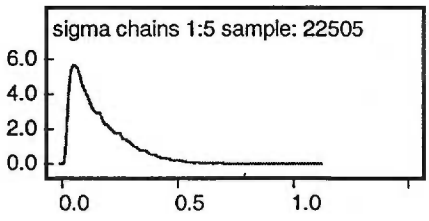
DATA>>

```
list(
 n       =11,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, -0.69, -0.78),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73,  0.30, 0.40)
)
```

Results>>

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.4342	0.112	0.002647	-0.6599	-0.432	-0.2207	500	22505
rr	0.6518	0.07307	0.001748	0.5169	0.6492	0.802	500	22505
sigma	0.162	0.1244	0.003398	0.02678	0.1261	0.4763	500	22505
theta[1]	-0.4991	0.1985	0.00401	-0.9782	-0.4736	-0.1717	500	22505
theta[2]	-0.4516	0.1534	0.003105	-0.7801	-0.4452	-0.1607	500	22505
theta[3]	-0.4293	0.1145	0.002531	-0.6564	-0.428	-0.2041	500	22505
theta[4]	-0.4764	0.1665	0.00341	-0.8552	-0.4627	-0.1756	500	22505
theta[5]	-0.5324	0.2162	0.004526	-1.076	-0.495	-0.2012	500	22505
theta[6]	-0.2867	0.1915	0.00479	-0.5833	-0.3199	0.1699	500	22505
theta[7]	-0.3503	0.2005	0.004266	-0.6851	-0.3762	0.137	500	22505
theta[8]	-0.4005	0.174	0.003422	-0.7367	-0.4075	-0.02167	500	22505
theta[9]	-0.3739	0.2196	0.004376	-0.7559	-0.3956	0.1588	500	22505
theta[10]	-0.4916	0.1734	0.003541	-0.8854	-0.4735	-0.1948	500	22505
theta[11]	-0.4847	0.1911	0.003779	-0.9427	-0.4645	-0.1548	500	22505





This shows that excluding the study does not significantly change the results.

Another study, Harwood et al, presented odds ratios for Whites and Hispanics separately. Because most of the other studies used White women, the White odds ratio was used in the above analysis. However, the analysis was repeated using the data for Hispanics. The WINBUGS data statement is shown below:

```
list(
 n       =12,
logrr = c(-0.87, -0.92, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, -0.69, -0.78, 0.77),
se    = c(0.42,  0.47,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73,  0.30, 0.40, 0.75)
)
```

The results are shown below:

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.426	0.119	0.002557	-0.6661	-0.424	-0.1925	500	22505
rr	0.6577	0.07885	0.001649	0.5137	0.6544	0.8249	500	22505
sigma	0.184	0.1453	0.004809	0.02716	0.1424	0.5639	500	22505
theta[1]	-0.5036	0.2127	0.004889	-1.03	-0.474	-0.1526	500	22505
theta[2]	-0.4991	0.2188	0.004618	-1.038	-0.4701	-0.1316	500	22505
theta[3]	-0.4268	0.1175	0.002291	-0.662	-0.4261	-0.1943	500	22505
theta[4]	-0.4803	0.177	0.003558	-0.8794	-0.4646	-0.1659	500	22505
theta[5]	-0.5452	0.2327	0.005892	-1.13	-0.5004	-0.1975	500	22505
theta[6]	-0.2657	0.2025	0.005287	-0.5802	-0.2996	0.2123	500	22505
theta[7]	-0.3294	0.2201	0.004757	-0.6868	-0.3624	0.2084	500	22505
theta[8]	-0.3895	0.187	0.003445	-0.7502	-0.3977	0.02673	500	22505
theta[9]	-0.3525	0.2454	0.004803	-0.7612	-0.3837	0.2547	500	22505
theta[10]	-0.4933	0.1828	0.003874	-0.9229	-0.4742	-0.1781	500	22505
theta[11]	-0.4921	0.2053	0.004411	-0.9831	-0.4665	-0.1474	500	22505
theta[12]	-0.3364	0.2581	0.005162	-0.7443	-0.3752	0.3263	500	22505

These results show that the estimates for RR and sigma do not change significantly when the Hispanic data is used.

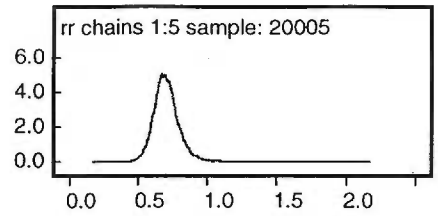
Because of concerns about combining the results of different study designs, the analysis was repeated using just case control studies. The data are shown below:

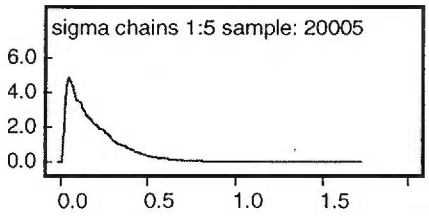
DATA>>

```
list(
 n      =10,
logrr = c(-0.87, -0.51, -0.43, -0.63, -1.11, 0.10, 0.14, -0.25, 0.51, 0.77),
se    = c(0.42,  0.26,  0.15,  0.29,  0.42,  0.25, 0.42,  0.35,  0.73, 0.75)
)
```

Results: The results from 20000 draws (5 chains of 5000 iterations) after a burn in of 5000 draws (1000 iterations) are shown below. A burn in of 5000 draws was used because a longer time was needed to reach convergence. A noninformative normal prior on psi and a noninformative gamma prior on tau was used.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.3548	0.1294	0.00301	-0.6024	-0.3585	-0.08647	1000	20005
rr	0.7073	0.09461	0.002177	0.5475	0.6987	0.9172	1000	20005
sigma	0.1954	0.159	0.005393	0.02696	0.1509	0.5966	1000	20005
theta[1]	-0.4542	0.2236	0.004931	-0.9915	-0.4251	-0.08071	1000	20005
theta[2]	-0.4073	0.1688	0.003412	-0.7747	-0.3973	-0.09257	1000	20005
theta[3]	-0.3945	0.1209	0.002691	-0.6433	-0.3921	-0.1634	1000	20005
theta[4]	-0.436	0.1838	0.004013	-0.8571	-0.4192	-0.1122	1000	20005
theta[5]	-0.4996	0.2459	0.005791	-1.117	-0.4527	-0.1327	1000	20005
theta[6]	-0.2081	0.203	0.006251	-0.5265	-0.2421	0.2661	1000	20005
theta[7]	-0.2665	0.2284	0.005817	-0.6355	-0.3016	0.2981	1000	20005
theta[8]	-0.3346	0.1933	0.004059	-0.7091	-0.3432	0.09118	1000	20005
theta[9]	-0.2815	0.2635	0.005924	-0.7086	-0.319	0.387	1000	20005
theta[10]	-0.2652	0.2758	0.006358	-0.695	-0.3103	0.4599	1000	20005





The rr density plot shows that most of the 20000 draws resulted in a odds ratio of less than one, although some of the draws found rr of greater than two.

The analysis could not be done on the two cohort studies because the random effects model requires two sources of variation and there is only one souce of variation, between study variation, with just two studies.

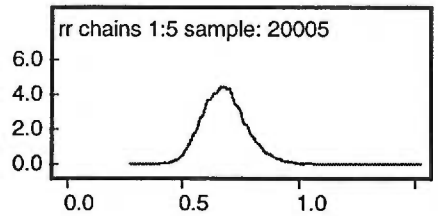
To better evaluate the effects of estrogen on AD only, the analysis was done using the risk estimates for the association between HRT and AD (excluding results for other dementias). The data and results are shown below:

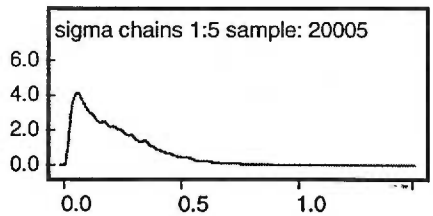
DATA>>

```
list(
 n      =11,
logrr = c(-0.87, -0.51, -0.60, -1.11, 0.10, 0.14, -0.25, 0.51, 0.77, -0.69, -0.78),
se    = c(0.42,  0.26,  0.38,  0.42, 0.25, 0.42,  0.35, 0.73, 0.75, 0.30,  0.40)
)
```

Results: The results from 20000 draws (5 chains of 5000 iterations) after a burn in of 5000 draws (1000 iterations) are shown below. A burn in of 5000 draws was used because a longer time was needed to reach convergence. A noninformative normal prior on psi and a noninformative gamma prior on tau was used.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.3906	0.142	0.003235	-0.6679	-0.3887	-0.112	1000	20005
rr	0.6835	0.0981	0.002246	0.5128	0.678	0.894	1000	20005
sigma	0.2163	0.1638	0.005448	0.02906	0.1773	0.6163	1000	20005
theta[1]	-0.496	0.2405	0.005581	-1.066	-0.4654	-0.09021	1000	20005
theta[2]	-0.4321	0.1787	0.0038	-0.8092	-0.4236	-0.0909	1000	20005
theta[3]	-0.4406	0.2166	0.004374	-0.9162	-0.4269	-0.03751	1000	20005
theta[4]	-0.542	0.2617	0.006791	-1.17	-0.4978	-0.1301	1000	20005
theta[5]	-0.2115	0.2179	0.005796	-0.5751	-0.2387	0.2784	1000	20005
theta[6]	-0.2778	0.2449	0.005362	-0.6878	-0.3097	0.2992	1000	20005
theta[7]	-0.3544	0.2061	0.003998	-0.7504	-0.3609	0.08291	1000	20005
theta[8]	-0.3033	0.279	0.005915	-0.778	-0.3359	0.3745	1000	20005
theta[9]	-0.2789	0.2927	0.006436	-0.7426	-0.3241	0.4542	1000	20005
theta[10]	-0.4848	0.2058	0.004742	-0.9414	-0.465	-0.127	1000	20005
theta[11]	-0.4761	0.2301	0.005181	-1.016	-0.4516	-0.07956	1000	20005





The risk estimate is similar to the estimate when the data for all dementias are combined which is not surprising given that most of the studies only looked at AD.

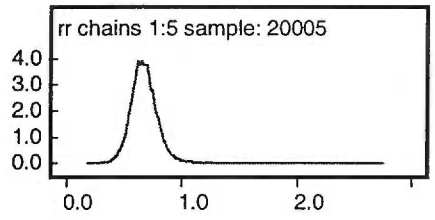
As the NINCDS criteria is felt to be the most rigorous way to diagnose AD, only studies which used the NINCDS criteria were combined. The data and results are shown below:

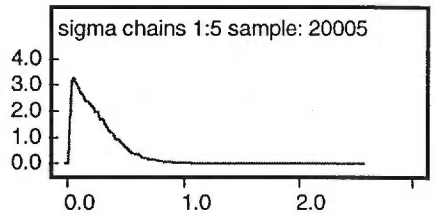
DATA>>

```
list(
 n       =7,
logrr = c( -0.51,  -1.11, 0.10, 0.14, -0.25, -0.69, -0.78),
se    = c( 0.26,   0.42,  0.25, 0.42,  0.35,  0.30,  0.40)
)
```

Results: The results from 20000 draws (5 chains of 5000 iterations) after a burn in of 5000 draws (1000 iterations) are shown below. A burn in of 5000 draws was used because a longer time was needed to reach convergence. A noninformative normal prior on psi and a noninformative gamma prior on tau was used.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.4117	0.1732	0.003112	-0.7636	-0.4099	-0.07615	1000	20005
rr	0.6726	0.1195	0.002149	0.466	0.6637	0.9267	1000	20005
sigma	0.2509	0.1976	0.005193	0.03123	0.2049	0.7396	1000	20005
theta[1]	-0.4491	0.1908	0.003684	-0.8513	-0.4442	-0.08318	1000	20005
theta[2]	-0.5806	0.2872	0.005842	-1.273	-0.5333	-0.1352	1000	20005
theta[3]	-0.1974	0.2298	0.005446	-0.578	-0.2218	0.307	1000	20005
theta[4]	-0.2707	0.2622	0.005208	-0.7065	-0.3038	0.3521	1000	20005
theta[5]	-0.361	0.222	0.003832	-0.7863	-0.3719	0.1217	1000	20005
theta[6]	-0.5057	0.2174	0.004036	-0.9888	-0.4884	-0.1196	1000	20005
theta[7]	-0.5075	0.2493	0.004605	-1.074	-0.4826	-0.07033	1000	20005





The risk estimate is similar using only the results for the studies which used the NINCDS criteria.

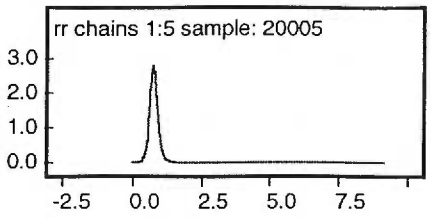
Finally, just case control studies that used NINCDS criteria were included in the meta-analysis. The results are shown below:

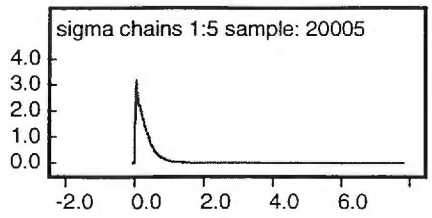
DATA>>

```
list(
 n       =5,
logrr = c( -0.51,  -1.11, 0.10, 0.14, -0.25),
se    = c( 0.26,   0.42,  0.25, 0.42,  0.35)
)
```

Results: The results from 20000 draws (5 chains of 5000 iterations) after a burn in of 5000 draws (1000 iterations) are shown below. A burn in of 5000 draws was used because a longer time was needed to reach convergence. A noninformative normal prior on psi and a noninformative gamma prior on tau was used.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.2921	0.2408	0.004714	-0.7752	-0.2862	0.1485	1000	20005
rr	0.769	0.2202	0.004332	0.4606	0.7511	1.16	1000	20005
sigma	0.3044	0.2941	0.006929	0.03082	0.2235	1.05	1000	20005
theta[1]	-0.3814	0.2155	0.004637	-0.8464	-0.3695	0.005023	1000	20005
theta[2]	-0.5247	0.3419	0.007512	-1.332	-0.461	-0.01145	1000	20005
theta[3]	-0.1115	0.2226	0.004631	-0.5036	-0.1284	0.3691	1000	20005
theta[4]	-0.1595	0.2826	0.00575	-0.6468	-0.1909	0.5047	1000	20005
theta[5]	-0.2726	0.2387	0.004518	-0.7512	-0.2713	0.2069	1000	20005





Although the risk estimate is in the same range as previously, the confidence intervals are much wider and some of the draws resulted in risk estimates of over 7.5, likely secondary to the small number of studies in this analysis.

More Sensitivity Analysis: To further address the robustness of the results, further sensitivity analysis was done using the random effects model.

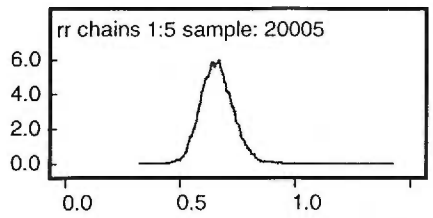
Distributions of theta:

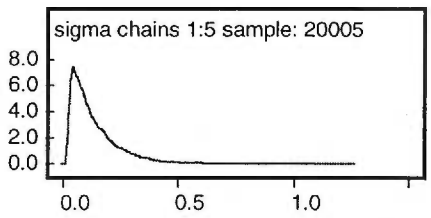
The model was run with theta having a t distribution instead of a normal distribution. This results in more variability because the thetas have heavier tails. A t-distribution has a variance of (v/v-2))sigma; v are the degrees of freedom. 4 degrees of freedom was used so that the thetas would have twice the variance of a normal random variable. The WINBUGS statement is shown below:

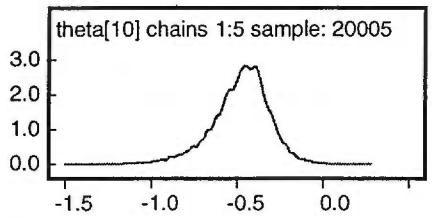
```
model;
{
  for( i in 1 : n ) {
    varinv[i] <- 1 / (se[i] * se[i])
  }
  for( i in 1 : n ) {
    logrr[i] ~ dnorm(theta[i],varinv[i])
  }
  for( i in 1 : n ) {
    theta[i] ~ dt(psi,tau,4)
  }
  psi ~ dnorm( 0.0,1.0E-6)
  tau ~ dgamma(0.001,0.001)
  rr <- exp(psi)
  sigma <- 1 / sqrt(tau)
}
```

The results are shown below. A burn in period of 5000 draws (1000 iterations) was used because it took longer for convergance.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
psi	-0.42	0.1082	0.003398	-0.6283	-0.4193	-0.2105	1000	20005
rr	0.6609	0.07181	0.002254	0.5335	0.6575	0.8102	1000	20005
sigma	0.1325	0.1037	0.003471	0.02509	0.1013	0.4024	1000	20005
theta[1]	-0.4848	0.1978	0.004401	-0.9665	-0.4606	-0.1528	1000	20005
theta[2]	-0.4411	0.1489	0.003558	-0.7546	-0.4356	-0.1557	1000	20005
theta[3]	-0.4259	0.1123	0.002992	-0.6501	-0.4258	-0.2048	1000	20005
theta[4]	-0.4659	0.1621	0.003888	-0.8335	-0.4538	-0.1786	1000	20005
theta[5]	-0.5246	0.2238	0.00578	-1.103	-0.4845	-0.1931	1000	20005
theta[6]	-0.2707	0.2008	0.005536	-0.5727	-0.3062	0.2044	1000	20005
theta[7]	-0.3372	0.2092	0.004826	-0.667	-0.3684	0.198	1000	20005
theta[8]	-0.389	0.1706	0.003927	-0.7089	-0.3985	-0.004758	1000	20005
theta[9]	-0.3567	0.2374	0.005034	-0.7301	-0.3891	0.2406	1000	20005
theta[10]	-0.4789	0.1678	0.0044	-0.8693	-0.4617	-0.1852	1000	20005
theta[11]	-0.4749	0.1869	0.00433	-0.9212	-0.4554	-0.1602	1000	20005
theta[12]	-0.3433	0.2501	0.005741	-0.7145	-0.3811	0.31	1000	20005







Altering the distribution did not significantly change the estimate for RR but sigma decreased slightly.

Prior distribution:

A previous meta-analysis by Yaffe was done in 1996. We therefore used the Yaffe study to provide prior information rather than letting the prior distribution be noninformative. Two additional studies were done after the Yaffe study and these were used in the analysis. The data statement for this procedure is shown below:

```
model;
{
  for( i in 1 : n ) {
    varinv[i] <- 1 / (se[i] * se[i])
  }
  for( i in 1 : n ) {
    logrr[i] ~ dnorm(theta[i],varinv[i])
  }
  for( i in 1 : n ) {
    theta[i] ~ dnorm(psi,tau)
  }
  psi ~ dnorm(-0.34,42.21)
  tau ~ dgamma(0.001,0.001)
  rr <- exp(psi)
  sigma <- 1 / sqrt(tau)
}
list(
 n       =2,
logrr = c(-0.87, -0.51),
se    = c(0.42,  0.26)
)
# chain 1
list(psi = 0,  tau = 10,  theta = c(0, 0))
# chain 2
list(psi = 15, tau = 1, theta = c(15, 15))
#chain 3
list(psi = 5, tau = 1, theta = c(5, 5))
#chain 4
list(psi = 4, tau = 10, theta = c(4, 4))
#chain 5
list(psi = -4, tau = 10, theta = c(-4, -4))
```

The results are shown below.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample

psi	-0.4118	0.1375	0.00228	-0.672	-0.4135	-0.1373	500	22505
rr	0.6688	0.09306	0.001527	0.5107	0.6613	0.8717	500	22505
sigma	0.3044	0.652	0.01044	0.02725	0.1487	1.477	500	22505
theta[1]	-0.5307	0.2598	0.004419	-1.168	-0.4962	-0.1099	500	22505
theta[2]	-0.454	0.1816	0.002764	-0.8294	-0.4487	-0.1043	500	22505

They show that the RR is about the same although sigma is slightly larger.

In summary, then, the estimates for rr do not change significantly when the random effects model is used compared to the fixed effects. Multiple different analysis usings different study estimates, different distributions, and different prior probabilities also do not significantly change the results suggesting the summary estimates are robust. The summary is shown below:

Studies used	OR	Lower CI	Upper CI
All studies--Unadjusted Heyman data	0.66	0.53	0.82
All studies except Heyman	0.65	0.52	0.80
All studies using Heyman data from Yaffe	0.68	0.53	0.84
Case control only	0.71	0.55	0.92
AD only	0.68	0.51	0.89
AD-NINCDS criteria only	0.67	0.46	0.92
AD-NINCDS criteria CC only	0.77	0.46	1.16
Using t distribution	0.67	0.53	0.81
Using Yaffe as prior distribution	0.67	0.51	0.87