Published in final edited form as:

Clin Trials. 2008 ; 5(5): . doi:10.1177/1740774508096313.

A multi-center, randomized, double blind placebo-controlled trial of estrogens to prevent Alzheimer's disease and loss of memory in women: design and baseline characteristics

Mary Sano^a, Diane Jacobs^b, Howard Andrews^c, Karen Bell^d, Neill Graff-Radford^e, John Lucas^f, Peter Rabins^g, Karen Bolla^g, Wei-Yan Tsai^h, Peter Cross^c, Karen Andrews^c, Rosann Costaⁱ, and Xiaodong Luo^a

^aMount Sinai School of Medicine, Department of Psychiatry, New York, NY, USA

^bConsulting Neuropsychologist, San Diego, CA, USA

^cNew York State Psychiatric Institute, Data Coordinating Center, New York, NY, USA

^dColumbia University College of Physicians & Surgeons, Department of Neurology, New York, NY, USA

^eMayo Clinic, Department of Neurology, Jacksonville, FL, USA

^fMayo Clinic, Departments of Psychiatry and Psychology, Jacksonville, FL, USA

^gJohns Hopkins University School of Medicine, Department of Psychiatry, Baltimore, MD, USA

^hColumbia University College of Physicians & Surgeons, Department of Biostatistics, New York, NY, USA

ⁱColumbia University College of Physicians & Surgeons, Gertrude H. Sergievsky Center, New York, NY, USA

Abstract

Background—Observational studies and small clinical trials suggested that hormone replacement therapy (HRT) decreases risk of cognitive loss and Alzheimer's disease (AD) in postmenopausal women and may have value in primary prevention.

Purpose—A clinical trial was designed to determine if HRT delays AD or memory loss. This report describes the rationale and original design of the trial and details extensive modifications that were required to respond to unanticipated findings that emerged from other studies during the course of the trial.

Methods—The trial was designed as a multi-center, placebo-controlled primary prevention trial for women 65 years of age or older with a family history of dementia. Recruitment from local sites was supplemented by centralized efforts to use names of Medicare beneficiaries. Inclusion criteria included good general health and intact memory functioning. Participants were randomized to HRT or placebo in a 1:1 ratio. Assignment was stratified by hysterectomy status and site. The primary outcomes were incident AD and memory decline on neuropsychological testing.

Results—Enrollment began in March 1998. In response to the Women's Health Initiative (WHI) May 2002 report of increased incidence of heart disease, stroke, pulmonary embolism, and breast

cancer among women randomized to HRT, participants were re-consented with a revised consent form. Procedural modifications, including discontinuation of study medication and a modification of the planned primary outcome based on a final enrollment below the target enrollment ($N = 477$), were enacted in response to the subsequent WHI Memory Study report of increased risk of dementia and poorer cognitive function with HRT. The mean length of treatment exposure prior to discontinuation was 2.14 years. Participants' mean age at baseline was 72.8; mean education was 14.2 years. Minority participation was 19% and 34% had a hysterectomy. The study continues to follow these participants for a total of 5 years blind to the original medication assignment.

Limitations—Results reported from the WHI during the course of this study mandated extensive procedural modifications, including discontinuing recruitment before completion and halting study medication. Alternative strategies for study redesign that were considered are discussed.

Introduction

This report describes the design and implementation of the PREPARE (PREventing Postmenopausal Alzheimer's Disease with Replacement Estrogens) Trial, the first Alzheimer's Disease (AD) primary prevention trial to be funded by the National Institute of Aging. The PREPARE Trial is a multi-center, randomized, double blind placebo-controlled trial of replacement estrogens to prevent AD and memory loss in women at increased risk for dementia by virtue of a positive family history. It was initiated at a time when few clinical trials addressed cognitive health in aging.

The rationale and structure for the PREPARE Trial as initiated is described and modifications made in response to new information reported by other studies during the course of the trial are discussed. Specifically the Women's Health Initiative (WHI) and the Women's Health Initiative Memory Study (WHIMS) found unexpected adverse effects of estrogen use and necessitated significant redesign of the PREPARE trial. These modifications included treatment discontinuation, redefinition of outcomes, and new analytic plans. The experience in addressing these issues provides lessons for managing future studies and is discussed in detail.

Background and rationale for the PREPARE trial

Scientific rationale for the trial

Societal impact of dementia—Dementia is now the number one cause of disability in aging; it robs quality of life from individuals with the disease and from those who care for them, and is estimated to cost over \$148 billion annually in direct and indirect costs [1]. AD is the most common cause of dementia. Cognitive decline, and memory loss in particular, is of great concern to the aging public and enthusiasm to ameliorate deterioration has high relevance. Therefore, prevention strategies that would completely prevent or delay onset are a high priority.

Rationale for estrogen replacement for dementia prevention: historical perspective—At the time of trial planning, compelling evidence from longitudinal and cross-sectional epidemiological studies, small clinical trials and laboratory research in model systems suggested that estrogen replacement therapy may reduce the risk of AD in women. Several prospective epidemiological studies and a systematic meta-analysis of observational studies indicated that estrogen use was associated with a reduced risk of dementia in general and AD specifically [2-4]. These reports were relatively consistent in showing approximately a 30% reduction in dementia risk associated with hormone replacement therapy use. Subsequent population-based studies [5,6] and a second meta-analysis [7] found similar associations. The most commonly used form of estrogen replacement included in

these investigations was conjugated equine estrogens (CEE). In some observational studies estrogen appeared to help to maintain or improve cognitive abilities in postmenopausal women. For example, Resnick *et al.* [8] found better performance on a visual memory test in women using estrogen compared to those not receiving replacement in the Baltimore Longitudinal Study of Aging. In a community-based sample in New York City, Jacobs *et al.* [9] observed preservation of verbal memory, abstract reasoning, and language skills among women reporting estrogen use during the postmenopausal period. Carlson *et al.* likewise observed a benefit over time among elderly women with past but not current estrogen exposure [10].

Several small, short-term clinical trials also demonstrated a benefit on cognitive outcomes with CEE and estradiol. Sherwin [11] found improved performance on tests of attention (digit span), abstract reasoning, clerical speed and accuracy, and immediate paragraph recall among women treated with IM estradiol valerate after hysterectomy. Similarly, in a later report, Phillips and Sherwin [12] demonstrated preservation of verbal paired-associate learning and improved immediate paragraph recall among surgically post-menopausal women treated with estradiol valerate. Shaywitz *et al.* [13] demonstrated a benefit of CEE in postmenopausal nondemented women with cognitive complaint. These same findings had provided the rationale for the WHIMS study, a cognitive add-on to the WHI, which was underway at the initiation of PREPARE, but no results were yet available.

The WHIMS [14] was an ‘add-on’ protocol to examine cognitive outcomes among participants in the WHI [15]. Initiated in 1992, WHI’s primary goal was to examine the effect of estrogen replacement on coronary heart disease and breast cancer outcomes in an 8.5 year long trial. Treatment regimen for the combination hormone therapy was 0.625 mg/day CEE and 2.5 mg. of medroxyprogesterone acetate (MPA) for women with a uterus and the CEE alone for those who had a hysterectomy. WHIMS began in May of 1996 and enrolled 92% of the eligible women (i.e., aged 65 or older and free of probable dementia at baseline), from the placebo controlled, randomized WHI trial. The 2-stage evaluation for WHIMS consisted of screening with a brief global measure of mental status, followed by a detailed clinical evaluation for those who scored below pre-set cut points. The primary outcome measures of this sub-study were incident dementia and mild cognitive impairment (MCI) defined as performance at or below the 10th percentile in at least one area of cognitive function.

Enriched design with family history—The incidence of AD ranges from 0.25% at age 65 to 2% at age 80 [16]. Thus, one of the challenges of designing a prevention trial is to enroll enough subjects to test the hypothesis that incidence rate is lowered. One known strategy is to recruit an enriched cohort to increase statistical power and shorten study duration by decreasing the requisite follow-up interval prior to incident events. Known factors that increase the risk of AD include advanced age, head trauma, limited education, family history of AD, and the epsilon 4 (e4) allele of the apolipoprotein-E (APOE) gene. Although advancing age is a strong risk factor for AD, the 2–4% estimates of incidence through the 9th decade are low, and a cohort of that age would be difficult to recruit. Head trauma, another risk for dementia, is relatively uncommon, particularly among women. Low educational attainment as a risk factor would be difficult to identify within the US since the cohort over 60 has a mean of nearly 12 years of education. Family history of AD in a first-degree relative increases the lifetime risk of AD for an individual by 3.5-fold [18]. The epsilon 4 (e4) allele of the apolipoprotein-E (APOE) gene has been shown to increase the risk of late-onset AD, however there are ethical concerns in using genotyping as a method to identify individuals at risk for AD. Hence, it was decided to use the report of family history in a first-degree relative (i.e., parent, sibling, child, or an extended family member if such

information was unknown in a first-degree relative due to early death) was used to establish an 'at risk' population.

Selecting the estrogenic agent—The treatment regimen for the PREPARE Trial paralleled that of WHI (i.e., combination hormone therapy was 0.625 mg/day CEE and 2.5 mg. of MPA for women with a uterus and the CEE alone for those who had hysterectomy). While recent results from the WHI have raised questions about the selection of this specific form of estrogen it is worth noting that the strongest human data at the time the PREPARE study was designed were epidemiologic studies in which the most commonly used estrogen was (CEE). The use of progestins came into favor with the results of the PEPI trial, in which unopposed estrogen use was associated with endometrial hyperplasia [17]. CEE plus medroxyprogesterone acetate (CEE + MPA) became the only combination medication with FDA approval for hormone replacement in women with an intact uterus. As such, the information on safety was available for this regimen. Given that our outcome measures emphasize cognitive functioning, consideration was given to using an agent with greater CNS bioavailability (e.g., beta-estradiol), but to optimize the ecological validity of the results the current standard of care (i.e., CEE and CEE + MPA) was implemented. Thus, based on maximal information about safety, efficacy, and generalizability, CEE and CEE + MPA were selected as the hormone replacement agents of choice for the protocol.

Hypothesis testing—PREPARE hypothesized that oral estrogens (alone or in combination with progesterone) would significantly reduce the risk of AD among women with a family history of AD in a first-degree relative. It also hypothesized that women randomized to oral estrogen would perform better on memory tasks over time than women on placebo. The specific aims were to (1) compare the 5-year cumulative incidence rate of clinically diagnosed probable or possible AD among women randomized to either estrogens or placebo; (2) compare the 5-year rate of memory change among women randomized to either estrogens or placebo, and; (3) compare the 5-year cumulative incidence rates of adverse effects and all cause mortality among women randomized to either estrogens or placebo while examining tolerability. In addition, 4 domains of secondary outcome measures, biological, functional, affective, and other cognitive domains, were included to test these hypotheses and the baseline characteristics of the subjects recruited into the study are described below.

Methods

Subjects

This study planned to enroll 900 healthy (defined as no illnesses associated with mortality within 5 years) women aged 65 or older who were free of dementia or significant cognitive impairment (as determined by performance on the neuropsychological battery) and had a family history of AD in a first-degree relative as determined by a standardized interview [18]. Inclusion criteria, designed to maximize safety for the use of estrogen, included a normal mammogram and gynecological examination. Exclusion criteria included the presence of neurological or psychiatric illness associated with cognitive change, recent (within 1 year) use of oral or trans-dermal estrogen, unstable cardiac disease, history of gynecological or breast cancer, thromboembolic events, liver or gallbladder disease or unstable life threatening medical condition. All subjects gave signed informed consent in accordance with local Institutional Review Boards prior to any screening assessments.

Administrative structure—Study oversight was accomplished by a 3 tiered administrative structure which included an administrative and data coordinating Center (DCC) at the parent site (Columbia University), three study centers and multiple sites.

Data coordinating center—The DCC along with administrative and executive staff from the parent institution were responsible for the development of the case report forms, secure data collection and monitoring, biological sample management, randomization and drug dispensing through a centralized pharmacy, and management of regulatory documents. They also took a major role in coordinating the part of the recruitment effort which used the Medicare data base (see subsequently).

Centers—The study centers were the home institutions of the three senior investigators at the start of the study (Columbia University; Johns Hopkins University; and Mayo Clinic, Jacksonville) and each was responsible for one third of the recruitment. Each was a tertiary medical center with well established clinical and research expertise in AD. All were NIA designated Alzheimer's Disease Centers and had expertise in neurology or psychiatry, and neuropsychology with extensive experience and infrastructure to conduct clinical trials. Each center had an expert clinician and a senior neuropsychologist. To accomplish the assigned recruitment, the three centers enrolled from multiple sites and took responsibility to provide ongoing training and supervision to the sites to insure standardized assessments were conducted in accordance with pre-stated trial procedures. Protocols were developed to standardize training for clinical and neuropsychological assessments.

Study sites—Study sites were selected for clinical and research expertise in AD and memory loss with resources available for evaluation and treatment. These were usually within Neurology or Psychiatry services. They were approved for participation by the staff of the centers. (Performance sites are listed in the acknowledgements.)

Recruitment strategy—Initial recruitment focused on site-specific resources such as local patient or research data bases, recruitment from educational talks or advertising and by approaching participants in other ongoing research studies. Initial projections from each center for available subject pools were estimated at three times the needed sample size. However, soon after the study began it became apparent that this strategy could not recruit adequate numbers of subjects, primarily because of the requirement of a family history of dementia, and thus additional recruitment methods were sought.

CMS (Centers for Medicare and Medicaid Services) lists as subject pool—

Recruitment was supplemented with a population-based methodology using listings of Medicare and Medicaid beneficiaries in the geographic areas surrounding the study sites. These files include over 95% of the resident population aged 65+ years, thereby providing a representative sample. Files consist of the beneficiary's full name, date of birth, gender, mailing address (including ZIP and county codes), and race/ethnicity.

A request to the Health Care Financing Administration (now the Centers for Medicare and Medicaid Services or 'CMS') was made for the records of all persons in selected counties surrounding study sites. Three million records were received. After the exclusion of males and of those with incomplete or incorrect addresses, a list of 1.6 million females aged 65 or over remained. Since CMS files do not include the beneficiary's phone number; commercial firms were used and a phone number was found for over 60% of the addresses.

An initial letter from CMS allowed each selected recipient to refuse to participate further in the study; <1% refused. A second letter, printed on university letterhead and signed by the principal investigator, explained the study in some detail and again allowed prospective participants to opt out (refusals here, too, were rare). Subsequently, participants were phoned directly by centralized study staff from the coordinating center and provided information about the study and its eligibility criteria. Women who expressed interest and a

willingness to participate came to a study site near them to complete informed consent and more detailed screening.

Screening and baseline evaluations—Consented subjects were assessed for entry criteria with a family history questionnaire; a medical history questionnaire, which included information regarding co-morbid conditions and concurrent medications; physical, neurological, breast and gynecological examinations; and a mammogram. The Beck Depression Inventory was administered to assess current depression. A neuropsychological examination was completed at screening to exclude individuals with dementia or significant memory impairment and to establish baseline performance; the battery is described in Table 1.

Randomization—Subjects who met entry criteria were randomized to receive estrogen (0.625 mg/day of CEE) or placebo in a 1 : 1 assignment. Specifically, those with an intact uterus received the estrogen in combination with MPA which was prepared as a single tablet. Those who had undergone hysterectomy received 0.625 mg CEE alone. The randomization scheme was stratified by hysterectomy status such that 50% of those with an intact uterus and 50% of those with hysterectomy received active treatment. Blocking by site was also included in the randomization scheme.

Follow-up—In the original proposal, participants were to be maintained on active treatment or placebo for 3 years. After randomization, subjects were seen annually and were assessed with a physical exam, neuropsychological assessment and also received all of the secondary outcome measures. Concurrent medications and intercurrent medical conditions were recorded or updated and medication compliance was assessed via pill count. To maximize retention, participants also received a brief phone evaluation at 6 month intervals in which concurrent medication use was updated and intercurrent illnesses were recorded. Study medication was collected and counted to assess compliance semi-annually. New medication was dispensed with a review of instructions for administration.

At each annual in-person assessment a significant impairment in memory performance could ‘trigger’ a ‘Dementia Evaluation’ that was scheduled within a month of the annual visit. The ‘trigger’ was defined as performance at or below 1.5 standard deviations of the demographically adjusted norm for immediate and delayed verbal recall. Subjects were considered to have ‘triggered’ if they performed 1 SD below the norm and these scores represented a decline from baseline testing, and were also scheduled for a dementia evaluation. The dementia evaluation included a neurological examination and clinical interview focusing on clinical evidence of cognitive, functional, and behavioral impairment. If warranted, further work-up including laboratory and imaging studies was done on a clinical basis and the results were provided to the study as stated in the consent. All subjects at study completion or at permanent discontinuation underwent a dementia evaluation. This provided data on potential dementia cases that did not meet the neuropsychological criteria to ‘trigger’ a dementia evaluation. It also provided data on which to assess the sensitivity and specificity of the screening procedure.

Primary outcomes—The study was designed with two co-primary outcomes. The first was incident AD defined by a consensus diagnosis of probable or possible AD. Each case that underwent a Dementia Evaluation was reviewed through a centrally-adjudicated mechanism via teleconference that was staffed by senior co-investigators from at least two sites and always included both a physician and a neuropsychologist. The operational criteria were based on NINDS-ADRDA criteria for Probable or Possible AD and DSM-IV criteria for AD. Specific criteria for dementia in this study consisted of impairment in memory and

nonmemory cognitive domains, and impairment in activities of daily living in the absence of other exclusionary etiologies.

The second primary outcome was memory decline assessed by rate of change on a composite memory score using four measures of memory: immediate and delayed recall of verbal and non-verbal memory. This summary score was created from neuropsychological measures collected at each time interval.

To assess safety, study emergent adverse medical events were reported as adverse events (AE). Gynecological and general physical examination, pap smear, and mammogram were repeated annually as an ongoing safety monitoring measure.

Data and safety monitoring board—An independent Data and Safety Monitoring Board (DSMB) was established to review the ongoing safety data of all patients enrolled in the trial. Areas of expertise among the DSMB members included clinical trials, hormonal and gynecological issues and neurological status along with statistical support. No investigator involved in the trial could be a member of the DSMB. The DSMB reviewed the safety data including adverse events, clinical evaluations, and laboratory studies on a routine basis. The DSMB received all serious adverse events and was immediately notified of fatal or life-threatening events. The DSMB received safety reports at regular intervals, and reviewed this information in meetings or teleconference calls. After each review, the DSMB submitted a written report to the NIA program director and the PI of this study, with recommendations regarding the conduct of the study.

Initial power estimates and analytic plan—The initial power calculation was based on several assumptions. Overall incidence of dementia in those over 65 was estimated at 1–2% per year and family history was thought to increase this 3.5-fold [18] to a range of 3.5–7%. Thus, an overall rate of 5% per year was considered to be a reasonable point estimate within the range. Given the plan to use co-primary outcomes, the usual alpha level of 0.05 was adjusted to 0.025 for each primary outcome. The enthusiasm for estrogen use at the time of study initiation led to an estimation for both a drop-in rate (10–20%) and drop-out rate (20–30%). No interim analyses were planned. To estimate power and sample size, the Log Rank test was proposed for the primary efficacy analysis for incident dementia. Based on these assumptions it was estimated that a sample size of 900 would yield 80% power to see a 50% reduction ($RR = 0.5$) in incident dementia over a 3-year period. Generalized Estimating Equations (GEE) were proposed to model change in the composite memory score over a 3-year period and a comparable sample size would yield 80% power to observe a difference between groups in memory decline of 2% per year.

The analytic plan to compare the active treatment groups to the placebo group included stratification for center and hysterectomy status. The point estimate and 95% confidence interval of the relative risk of treated subjects compared to placebo subjects was obtained using partial likelihood method from the Cox model. The *a priori* plan was to use analysis of variance or Chi-square tests, as appropriate. Other *a priori* variables to be considered in additional analyses include demographic variables (i.e., age, education, ethnicity), clinical history variables such as history of estrogen use or number of first degree relatives with AD, and biological markers. The plan was to examine them as predictors of the primary outcome and if significant to include them in the secondary statistical analysis. Time to the primary outcome estimates would be based on the survival curves generated from the Cox model. To assess safety, the prevalence of individual adverse reactions would be analyzed using Chi-square methods and Log Rank tests.

Results

The protocol was initiated as described in the preceding sections. Enrollment began in the first quarter of 1998 with the first subject randomized on March 24, 1998. Cumulative recruitment over time is shown in Figure 1. However, several interceding events, detailed below, required modification of the trial.

Event driven procedural modifications

Table 2 summarizes modifications occurring throughout the trial. First, in May 2002, the WHI study arm examining the effects of combination therapy (CEE + MPA) was discontinued after 5.6 years due to a finding that this treatment regimen increased risk of heart disease, stroke, pulmonary embolism, and breast cancer [19]. The PREPARE DSMB reviewed the WHI results and conducted a re-analysis of the adverse events within the PREPARE study. Based on this review, the DSMB recommended (1) a new consent form that summarized the WHI findings and (2) re-consenting previously consented participants. No modifications to entry criteria or study implementation were recommended.

Since participants were to be re-consented to inform them of the WHI findings, it was decided to simultaneously propose a lengthening of study treatment from 3 to 5 years in order to extend our observations and potentially increase power. Re-consenting and extending the treatment interval was met with high rates of retention (418 of 426). In re-estimating power, the criteria for ‘trigger’ were used as the outcome measure rather than incident dementia. The advantage of using ‘trigger’ is that it captures not only those participants meeting criteria for dementia, but also those potentially in a prodromal phase of dementia (i.e., relatively poor memory performance compared to peers and an absolute decline from baseline scores, but no functional impairment). The effect size estimate was also scaled down to a RR of 0.7. Finally the plan to split alpha for two relatively highly correlated outcomes seemed overly conservative and we re-estimated the power for each outcome at $\alpha = 0.05$. These modifications yielded a comparable power of 81% for this longer protocol assuming the sample of 900 could be achieved.

In May 2003, results from the WHIMS revealed an increased risk of dementia and poorer cognitive function among participants randomized to CEE + MPA. In June 2004, a similar but nonsignificant trend was reported among those on CEE. These findings led the PREPARE DSMB to recommend the cessation of study treatment with incomplete enrollment. The mean length of treatment exposure prior to discontinuation was 2.14 years with a minimum of 0.56 years and a maximum of 3 years. The alternatives reviewed by the DSMB at this point were to discontinue the study, break the blind and inform subjects of their treatment, or continue collecting data on the assembled cohort and maintain the blind in the absence of continued treatment. While the relatively short treatment exposure that many participants experienced during the trial was less than ideal, it actually approximated use patterns in epidemiological surveys and this possible benefit was weighed together with the fact that informing of the treatment assignment was noninformative with respect to clinical care or management. Re-estimates of power based on the available cohort at the time that treatment was halted ($N = 477$) and 5-year follow-up suggested very limited power for the outcome of trigger (power = 29%; $\alpha = 0.05$; two-sided test relative risk = 0.7). Power for the memory composite score was 64% ($\alpha = 0.05$; two-sided test; 1% change at end of 5 years) and the power to detect at least one significant difference (either trigger or composite score) was 74% ($1 - (1 - 0.29)(1 - 0.64)$). Though the power estimates were not ideal, the DSMB accepted the scientific value of continuing data collection and the protocol went forth as a blinded observational study for 5 years of data collection.

Description of the final cohort

Approximately 53% ($N = 477$) of the projected subjects ($N = 900$) were successfully recruited before the trial was stopped. Baseline demographic and clinical variables are presented in Table 3. The overall mean age was 72.8 (± 5.2) and the education level was 14.2 (± 3.1). 19% of participants were nonwhite. Approximately 34% of the cohort had undergone hysterectomy. Forty-six percent of participants reported using estrogen replacement prior to study enrollment and the average duration of hormone therapy was 5.7 years ($SD = 8.2$). The average interval between last use of hormone therapy and study enrollment was 13.4 years ($SD = 10.6$). APOE genotyping was available for 449 participants (94%). 142 participants (30%) had one or more APOE e4 alleles; 6 participants were e4 homozygous. e4 positive participants were significantly younger than e4 negative participants (e4 pos: 71.8 years; e4 neg: 73.1 years; $t(447) = 2.4$; $p < 0.02$).

Discussion

This report describes the design and data collection methods of a clinical trial for the prevention of Alzheimer's disease. While the WHIMS was the first 'add-on' trial to assess dementia prevention, PREPARE was the first prospectively designed trial powered to prevent AD to be funded by the National Institute of Aging. Although interest in estrogen for prevention of dementia has waned in the wake of the WHI and WHIMS findings, the PREPARE cohort, methodology, and methods of adapting to new information remain of interest. First, the PREPARE cohort was pre-selected to have intact cognition whereas nearly 10% of participants in the WHI had cognitive screen scores below the cut-off at entry [20]. Second, at the point of discontinuation of study medication in the PREPARE trial, subjects had been exposed to a mean of 25 months of study drug. This length of exposure is relatively comparable to that reported in population-based studies of hormone replacement therapy which demonstrated benefit [2-5], although the point in time for use of estrogens differs in that many of the population based studies include perimenopausal and early postmenopausal exposure to estrogens while PREPARE use was many years post menopause. However, randomization will permit us to determine if any effect of active drug on these outcomes is more than coincidence. Additionally, mirroring epidemiological studies, much of the benefit of HRT has been observed in cohorts with postmenopausal exposure, albeit at an earlier age than those in this randomized trial, followed by subsequent discontinuation. While these features were not part of the original design, maintaining the follow-up and the blind permits us to model the observed epidemiological phenomena upon which the study was originally justified. Equally important in the post WHIMS era will be documenting the absence or presence of distal negative effects after this brief exposure since estrogen remains an effective treatment for menopausal symptoms and is often used only briefly. Knowing the long-term consequence on cognition could have significant impact on acceptance of hormone use for these other conditions.

The PREPARE cohort was selected based on a family history of dementia. The data presented here suggest that doing so increases the rates of having an APOE e4 allele 2- to 3-fold compared to population rates. Nonetheless, the majority of the PREPARE cohort is not positive for this allele. Positive family history may also be associated with a differential effect of hormone replacement on cognition and the continued follow-up of the cohort off study medication will permit an assessment of this effect.

While the PREPARE design of selecting those with no evidence of memory deficit may lower the rate of cognitive decline and conversion to dementia, it insures a population with no evidence of a dementia at baseline. This will be particularly important for addressing future criticism that the neuropathology is likely to exist before treatment begins. However,

there is no method currently to confirm the presence or absence of neuropathology prior to symptom onset.

Great effort and expense went into the use of the CMS records for recruitment and the result was a sample highly representative of the US population. Education and ethnicity are comparable to United States population based estimates as per the 2000 census [21]. Use of hormones and the frequency of hysterectomy are comparable to rates observed in other reports. Thus, the data gleaned from PREPARE will provide valuable information about the natural history of cognitive function over time that will be generalizable to the breadth of the US population of aging women.

Value of administrative structure

The strength of the DCC is critical to multi-center studies to insure centralized data collection and administration. In the original proposal, recruitment was proposed as a set of site specific activities. However, over time this activity was centralized for several reasons. First, it became clear that local strategies were insufficient to accomplish the large recruitment. Second, senior staff changes occurred at two study centers and this altered the ability of the center to maintain oversight for local large recruitment efforts. Finally, additional sites were recruited and this provided a wider geographic area from which to enroll, although the new sites did not all have comparable expertise in recruitment of healthy subjects. This led to further centralization of the CMS recruitment effort and allowed for staff at individual sites to be trained and supervised for calling potential subjects and directing those interested to local sites for participation.

While these events may be study specific, staff changes are not uncommon, particularly in long trials, and insufficient recruitment resources are almost universal. Overall the experiences highlight the value of a single administrative coordinating center which offers the flexibility to make modifications as the need is identified.

Impact of trial modification

Modifications made during the course of the trial in response to emerging evidence from WHI and WHIMS made it necessary to shift the outcome from incident dementia to discrete measures of cognitive change and to focus on the continuous measures of memory performance to accommodate the smaller sample size and the truncated treatment interval. While these outcomes may not be as widely accepted as incident dementia, they remain important and consistent with recent interest in milder forms of cognitive deficit that may predict incident dementia. In fact, this new clarification allows us to add to the growing scientific and clinical interest in early cognitive changes and the identification of prodromal conditions such as MCI and cognitive impairment, not dementia (CIND). The particular advantage of this approach is that it permits detection of incident change rather than making the assumption of change based on normative data that are currently implicit in the MCI designation.

The decision to continue the blind after discontinuing treatment raised ethical issues. The modification permitted only modest power to see an effect on either the memory score or the change in cognition but this was balanced against the fact that notification of specific arm assignment could not direct any change in health behavior. The compromise position was to immediately remove all subjects from treatment, comprehensively inform subjects of the newly reported risks of estrogen, and to ask them to participate in the observational follow-up. This was met with high enthusiasm and participation, supporting the notion that the idea was in equipoise with community understanding of the situation and with the claim that many research participants do so because they perceive their participation as contributing to

a research effort that has the potential to benefit society as a whole rather than as individuals, who might experience a benefit from having dementia prevented or delayed in them. It is also possible that this cohort with a family history of dementia was more motivated than the general public to assess long-term cognitive effects.

Like many other trials, new information from other relevant ongoing studies required modification of study design and methods. This highlights the need to consider potential modifications *a priori* in order to maximize information from a trial. Clinical trials with other interventions [22] have reported findings in contradiction to prediction from basic science and epidemiologic studies. Designs which consider such paradoxical results and incorporate contingent procedural changes at the start may provide opportunities to maximize information.

Acknowledgments

This study was funded by R01 AG15922.

Participating Sites: Columbia University (PI: Karen Bell); Mayo Clinic Jacksonville (PI: Neill Graff-Radford); Johns Hopkins Medical Center (PI: Peter Rabins); Butler Hospital (PI: Steven Salloway); University Medical and Dental College of New Jersey (PI: William Reichman); Eastern Virginia Medical School (PI: Stefan Gravenstein); Medical University of South Carolina (PI: Jacobo Mintzer); University of California Irvine (PI: Ruth Mulnard); University of Alabama (PI: David Clark); Sun Health Research Institute (PI: Marwan N. Sabbagh); The New England Center for Headache (PI: Alan Pepper); The Winifred Masterson Burke Medical Research Institute, Inc. (PI: John Blass); West Florida Regional Medical Center (PI: B.D. Miller); North Broward Medical Center (PI: Murray Todd); Lee Memorial Health System (PI: Douglas A. Newland); St. Mary's Medical Center (PI: Carl Sadowsky).

References

1. Alzheimer's Association. Alzheimer's disease facts and figures. Alzheimer's Association. 2008;22.
2. 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. [PubMed: 8709781]
3. 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. *Neurology*. 1997; 48:1517–21. published correction appears in *Neurology*. 1998;51:654. [PubMed: 9191758]
4. Yaffe K, Sawaya G, Lieberburg I, Grady D. Estrogen therapy in postmenopausal women. *JAMA*. 1998; 279:688–95. [PubMed: 9496988]
5. Waring SC, Rocca WA, Petersen RC, et al. Postmenopausal estrogen replacement therapy and risk of AD: a population-based study. *Neurology*. 1999; 52:965–970. [PubMed: 10102413]
6. Harwood DG, Barker WW, Loewenstein DA, et al. A cross-ethnic analysis of risk factors for AD in white Hispanics and white non-Hispanics. *Neurology*. 1999; 52:551–56. [PubMed: 10025786]
7. LeBlanc ES, Janowsky J, Chan BK, Nelson HD. Hormone replacement therapy and cognition: systematic review and meta-analysis. *JAMA*. 2001 Mar 21; 285(11):1489–99. [PubMed: 11255426]
8. Resnick SM, Metter EJ, Zonderman AB. Estrogen replacement therapy and longitudinal decline in visual memory: A possible protective effect? *Neurology*. 1997; 49:1491–97. [PubMed: 9409335]
9. Jacobs DM, Tang M-X, Stern Y, et al. Cognitive function in nondemented older women who took estrogen after menopause. *Neurology*. 1998; 50:368–73. [PubMed: 9484355]
10. Carlson MC, Zandi PP, Plassman BL, et al. Cache County Study Group. Related Articles, Links Hormone replacement therapy and reduced cognitive decline in older women: the Cache County Study. *Neurology*. 2001; 57(12):2210–16. [PubMed: 11756599]
11. Sherwin BB. Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. *Psychoneuroendocrin*. 1988; 13:345–57.
12. Phillips SM, Sherwin BB. Effects of estrogen on memory function in surgically menopausal women. *Psychoneuroendocrin*. 1992; 17:485–95.

13. Shaywitz SE, Naftolin F, Zeltermann D, et al. Better oral reading and short-term memory in midlife, postmenopausal women taking estrogen. *Menopause*. 2003; 10:420–26. [PubMed: 14501603]
14. Shumaker SA, Reboussin BA, Espeland MA, et al. The Women's Health Initiative Memory Study (WHIMS): a trial of the effect of estrogen therapy in preventing and slowing the progression of dementia. *Control Clin Trials*. 1998; 19:604–21. [PubMed: 9875839]
15. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials*. 1998; 19(1):61–109. [PubMed: 9492970]
16. Brookmeyer R, Gray S. Methods for projecting the incidence and prevalence of chronic diseases in aging populations: applications to Alzheimer's disease. *Stat Med*. 2000; 19:1481–93. [PubMed: 10844713]
17. The Writing Group for the PEPI Trial. Effects of hormone replacement therapy on endometrial histology in postmenopausal women. The Postmenopausal Estrogen/Progestin Interventions (PEPI) Trial. *JAMA*. 1996; 275(5):370–75. [PubMed: 8569016]
18. Devi G, Marder K, Schofield PW, et al. Validity of family history for the diagnosis of dementia among siblings of patients with late-onset Alzheimer's disease. *Genet Epidemiol*. 1998; 15(3): 215–23. [PubMed: 9593109]
19. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the women's health initiative randomized controlled trial. *JAMA*. 2002; 288:321–33. [PubMed: 12117397]
20. Espeland MA, Rapp SR, Shumaker SA, et al. Conjugated equine estrogens and global cognitive function in postmenopausal women: Women's Health Initiative Memory Study. *JAMA*. 2004; 291(24):2959–68. [PubMed: 15213207]
21. United States Census Bureau. United States Census 2000: U S Summary. Jul.2002
22. ADAPT Research Group. Naproxen and celecoxib do not prevent AD in early results from a randomized controlled trial. *Neurology*. 2007; 68:1800–08. [PubMed: 17460158]
23. Buschke H, Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology*. 1974; 24:1019–25. [PubMed: 4473151]
24. Wechsler, D. Wechsler Memory Scale – Revised. The Psychological Corporation; San Antonio, TX: 1987.
25. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental State': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–98. [PubMed: 1202204]
26. Wechsler, D. Wechsler Adult Intelligence Scale-Revised. The Psychological Corporation; New York, NY: 1981.
27. Kaplan, E.; Goodglass, H.; Weintraub, S. Boston Naming Test. Lea & Febiger; Philadelphia, PA: 1983.
28. Benton, AL.; Hamsher, K deS. Multilingual Aphasia Examination. University of Iowa; Iowa City, IA: 1976.
29. Grooved Pegboard Test. Lafayette Instruments; Lafayette, IN:

Figure 1.
Cumulative recruitment over the enrollment period from first quarter 1998 (March, 1998) to second quarter of 2003 (May of 2003). Total recruitment is 477

Table 1

Neuropsychological measures

Domain	Test	Description/Source
Memory	Selective reminding test [23]	12 item 6-trial version with delayed free recall after a 15-min interval followed by a recognition trial
	Visual reproduction test [24]	Wechsler Memory Scale – Revised subtest with immediate recall of 4-line drawings and delayed recall after a 30-min interval
Orientation	Mini-mental state exam [25]	Subset of 10 items for orientation to time and place
Abstract Reasoning	Similarities [26]	Wechsler adult intelligence scale-revised subtest requiring identification of relevant superordinate category of word pairs
Language	Boston naming test [27]	15 item version
	Verbal fluency test [28]	Fluency for three letters and three categories
Attention	Cancellation test	Two timed paper-and-pencil tests using letter triad and shapes
	Digit symbol [26]	Timed task of matching symbols to numbers
Motor Speed	Grooved Pegboard [29]	Assesses manipulative dexterity with dominant and non-dominant hand

Table 2

Event driven procedural modifications

Date	Event
1998	PREPARE recruitment begins.
2002	Results of WHI cardiovascular outcomes are published. Initiate plan to reconsent PREPARE subjects with safety information.
2003	Results of WHIMS CEE + P trial released.
2004	Result of WHIMS CEE trial released.
2004	PREPARE DSMB recommends full halt to trial. Initiate plan to reconsent subjects with safety information and to participate in blinded observational follow-up.

Table 3Mean (SD) cohort characteristics at baseline¹

	Overall	No hysterectomy	Hysterectomy
<i>N</i>	477	314	163
Age	72.8 (5.2)	72.8 (5.1)	72.9 (5.2)
Education ²	14.2 (3.1)	14.5 (3.0)	13.7 (3.0)
Ethnicity (%)			
White	1.0	81.2	80.4
Black	12.8	11.8	14.7
Hispanic	4.6	5.4	3.1
Other	1.7	1.6	1.8
APOE (%)			
E4 Positive	31.8	32.8	29.3
E4 Negative	68.2	67.2	70.7
Blessed ADL	0.3 (0.5)	0.3 (0.5)	0.2 (0.5)
CDR (12-item)	10.9 (7.6)	10.9 (7.6)	10.8 (7.7)
Beck depression inventory	4.5 (3.7)	4.4 (3.8)	4.6 (3.4)
MMSE total score	28.9 (1.5)	28.9 (1.5)	28.8 (1.5)
HRT prior to enrollment (N) ³	218 (46%)	124 (40%)	94 (58%)
Time since last HRT use (years) ⁴	13.4 (10.6)	12.0 (9.8)	15.4 (11.3)

¹ ANOVA was used to compare hysterectomy groups on continuous variables; Chi-square for categorical comparisons.² Hysterectomy groups differ significantly ($p < 0.01$).³ Hysterectomy groups differ significantly ($p < 0.0001$).⁴ Hysterectomy groups differ significantly ($p < 0.05$).