Published in final edited form as:
Brain Res. 2013 June 13; 1514: 3–11. doi:10.1016/j.brainres.2013.03.047.

The rationale, design, and baseline characteristics of the Women's Health Initiative Memory Study of Younger Women (WHIMS-Y)

Leslie Vaughan, Ph.D.[1], **Mark A. Espeland, Ph.D.**[1], **Beverly Snively, Ph.D.**[1], **Sally A. Shumaker, Ph.D.**[1], **Stephen Rapp, Ph.D.**[2], **Jill Shupe, M.S.**[3], **Jennifer Robinson, M.D., M.P.H.**[4], **Gloria E. Sarto, M.D., Ph.D.**[5], and **Susan M. Resnick, Ph.D.**[6] **for the Women's Health Initiative Memory Study of Younger Women (WHIMS-Y) Study Group**

[1]Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina, 27157

[2]Department of Psychiatry, Wake Forest University School of Medicine, Winston-Salem, North Carolina, 27157

[3]Fred Hutchinson Cancer Research Center, Seattle, WA 98109 (jshupe@whi.org)

[4]Departments of Epidemiology and Medicine, Division of Cardiology, University of Iowa, Iowa City, IA 52242 (jennifer-g-robinson@uiowa.edu)

Corresponding Author: Leslie Vaughan, Ph.D., Department of Social Sciences and Health Policy, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157; alvaugha@wakehealth.edu. (alvaugha@wakehealth.edu; mespelan@wakehealth.edu; bmellen@wakehealth.edu; sshumake@wakehealth.edu; srapp@wakehealth.edu)

WHIMS-Y Clinical Coordinating Center: (Wake Forest School of Medicine, Winston-Salem, NC): Sally Shumaker, Mark Espeland, Steven Rapp, Claudine Legault, Laura Coker, Maggie Dailey, Michelle Naughton, Beverly Snively, Iris Leng, Kaycee Sink. **WHIMS-Y Clinical Centers:** (Albert Einstein College of Medicine, Bronx, NY): Sylvia Wassertheil-Smoller; (Baylor College of Medicine, Houston, TX): Aleksandar Rajkovic; (Brigham and Women's Hospital, Harvard Medical School, Boston, MA): JoAnn Manson; (Emory University, Atlanta, GA): Lawrence Phillips; (Fred Hutchinson Cancer Research Center, Seattle, WA): Andrea LaCroix; (George Washington University Medical Center, Washington, DC): Richard Katz; (Harbor-UCLA Research and Education Institute, Torrance, CA): Rowan Chlebowski; (Kaiser Permanente Center for Health Research, Portland, OR): Frances Lynch; (Kaiser Permanente Division of Research, Oakland, CA): Bette Caan; (Medical College of Wisconsin, Milwaukee, WI): Jane Morley Kotchen; (MedSTAR Research Institute, Washington, DC): Barbara Howard; (Memorial Hospital of Rhode Island, Pawtucket, RI): Charles Eaton; (Northwestern University, Chicago, IL): Sandra Weintraub; (Cook County Hospital, Rush-Presbyterian St. Luke's Medical Center, Chicago, IL): Lynda Powell; (Stanford Center for Research in Disease Prevention, Stanford University, Stanford, CA): Marcia L. Stefanick; (State University of New York at Stony Brook, Stony Brook, NY): Dorothy Lane; (The Ohio State University, Columbus, OH): Rebecca Jackson; (University of Alabama at Birmingham, Birmingham, AL): Beth Lewis; (University of Arizona, Tucson/Phoenix, Tucson, AZ): Cynthia Thomason; (University of Buffalo, Buffalo, NY): Jean Wactawski-Wende; (University of California at Davis, Sacramento, CA): John Robbins; (University of California at Irvine, Orange, CA): Allan Hubbell; (University of California at Los Angeles, Los Angeles, CA): Lauren Nathan; (University of Cincinnati, Cincinnati, OH): Margery Gass; (University of Florida, Gainesville/Jacksonville, Gainesville, FL): Marian Limacher; (University of Hawaii, Honolulu, HI): Kamal Masaki; (University of Iowa, Iowa City/Davenport, Iowa City, IA): Jennifer Robinson; (University of Massachusetts, Worcester, MA): Judith Ockene; (University of Medicine and Dentistry of New Jersey, Newark, NJ): Norman Lasser; (University of Medicine and Dentistry of New Jersey, New Brunswick, NJ): John Kostis; (University of Minnesota, Minneapolis, MN): Richard Grimm; (University of Nevada, Reno, NV): Robert Brunner; (University of North Carolina, Chapel Hill, NC): Carol Murphy; (University of Pittsburgh, Pittsburgh, PA): Lewis Kuller; (University of Tennessee at Memphis, Memphis, TN): Karen Johnson; (University of Texas Health Science Center, San Antonio, TX): Donald Royall; (University of Washington, Seattle, WA): Shirley Beresford; (University of Wisconsin, Madison, WI): Gloria Sarto; (Wake Forest University Health Sciences, Winston-Salem, NC): Mara Vitolins; and (Wayne State University School of Medicine/Hutzel Hospital, Detroit, MI): Michael Simon.

[5]University of Wisconsin Center for Women's Health Research, University of Wisconsin-Madison, Madison WI 53715 (gsarto@wisc.edu)

[6]Laboratory of Personality and Cognition, Intramural Research Program, National Institute on Aging, NIH, 5600 Nathan Shock Drive, Baltimore, MD 21224 (susan.resnick@nih.gov)

Abstract

The Women's Health Initiative Memory Study-Younger (WHIMS-Y) was designed to assess the effect of prior random assignment to hormone therapy (HT) (conjugated equine estrogen (CEE) alone or CEE plus medroxyprogesterone acetate (MPA)) on global cognitive function in younger middle-aged women relative to placebo. WHIMS-Y was an ancillary study to the Women's Health Initiative (WHI) HT trial and enrolled 1361 women who were aged 50-54 years and postmenopausal at WHI enrollment. WHIMS-Y will examine whether an average of 5.4 years of HT during early menopause has longer term protective effects on global cognitive function and if these effects vary by regimen, time between menopause and study initiation, and prior use of HT. We present the study rationale and design. We describe enrollment, adherence to assigned WHI therapy, and compare risk factor characteristics of the WHIMS-Y cohort at the time of WHI enrollment to similar aged women in the WHI HT who did not enroll in WHIMS-Y. Challenges of WHIMS-Y include lower than expected and differential enrollment. Strengths of WHIMS-Y include balance in baseline risk factors between treatment groups, standardized and masked data collection, and high rates of retention and on-trial adherence and exposure. In addition, the telephone-administered cognitive battery showed adequate construct validity. WHIMS-Y provided an unprecedented chance to examine the hypothesis that HT may have protective effects on cognition in younger postmenopausal women aged 50-54 years. Integrated into the WHI, WHIMS-Y optimized the experience of WHI investigators to ensure high retention and excellent quality assurance across sites.

Keywords

Postmenopausal hormone therapy; Cognitive function; Aging

1. Introduction

Although early termination and publication of the Women's Health Initiative Memory Study (WHIMS) primary results showed increased risk of probable dementia and no protection of global cognitive function following initiation of post-menopausal hormone therapy (HT) (Espeland et al., 2004; Rapp et al., 2003; Shumaker et al., 2003; 2004), there is continued speculation that HT may still protect cognitive function in women if initiated during the peri-menopausal or recent post-menopausal period (Craig et al., 2005; Joffe et al., 2006). This speculation has led to recent calls for studies that examine the possible protective effects of HT on cognitive function in peri-menopausal and recent post-menopausal women. (Henderson et al., 2007; Lord et al., 2008; Resnick & Henderson, 2002). The WHIMS Study of Younger Women (WHIMS-Y) provides a unique and cost-effective opportunity to evaluate the impact of HT on cognitive function in younger postmenopausal women enrolled in the Women's Health Initiative (WHI) HT trials at ages 50-54 years. It will assess the long-term impact of randomized assignment to HT among these women, and thus will provide critical information regarding the clinical treatment of younger post-menopausal women and potential mechanisms for how HT may affect cognitive function.

1.2 Objectives of WHIMS-Y

The primary objective of WHIMS-Y is to test the hypothesis that conjugated equine estrogen (CEE)-based HT (CEE-Alone or CEE + MPA (medroxyprogesterone acetate)) in postmenopausal women aged 50-54 years has a long-term effect on women's global cognitive function. Specifically, WHIMS-Y tests whether randomized assignment to CEE +MPA and/or CEE-alone in younger postmenopausal women may confer the proposed protection relative to placebo.

Secondary objectives are to determine whether effects on cognitive function vary according to prescription of unopposed or opposed CEE, years between menopause and the initiation of study-prescribed therapy, and prior use of HT. WHIMS-Y will also identify incident cases of probable dementia (PD) and mild cognitive impairment (MCI), however it is not expected to provide sufficient statistical power to detect differences in the rates of these.

In this paper that focuses on the design of WHIMS-Y, we describe the enrollment of the cohort and compare selected characteristics of the cohort at the time of their WHI HT enrollment with characteristics of similar-aged women in the WHI HT who declined enrollment in WHIMS-Y. We also describe adherence to assigned WHI therapy, using pill counts and length of enrollment in the WHIMS-Y cohort.

2. Results

Here, we report on the design of WHIMS-Y, including enrollment of the cohort, the comparison of cognitive risk factors and adherence patterns of WHIMS-Y enrollees and non-enrollees and CEE and CEE/MPA groups at WHI enrollment, and baseline cognitive characteristics of WHIMS-Y women at the initiation of cognitive testing. In addition, we describe the results of factor analytic analyses of the cognitive battery. In year 1, $N = 1732$ currently active participants of the WHI Extension Study agreed to initial contact by the WHIMS coordinating center, and $N = 1361$ (78.6%) agreed to participate. Of these, $N = 1264$ (93.1%) completed the test battery in Year 1, with a small percentage lost to follow-up after eight attempts to contact. An additional $N = 62$ participants included in the analyses completed the test battery for the first time in years 2 or 3.

In the comparison of WHIMS-Y enrollees and non-enrollees at the time of their WHI enrollment, a number of risk factors were examined; including age, age at last menstrual period, education, race and ethnicity, smoking status, alcohol intake, body-mass index (BMD), hypertension status, prior cardiovascular disease (CVD), hysterectomy, years since last regular menstrual period, prior HT at recruitment, and adherence. As seen in Table 1, at WHI enrollment there was no difference in the distributions of important potential confounds between women in the placebo and the HT groups. When we compared WHIMS-Y enrollees to non-enrollees, there were significant or marginal differences in several variables, including: age at last menstrual period, education, race and ethnicity, alcohol consumption, BMI, years since last regular menstrual period, prior HT at WHI recruitment, and adherence. Overall, WHIMS-Y enrollees reported being slightly older at their last menstrual period ($M = 45.1$, $SD = 6.2$) than non-enrollees ($M = 44.4$, $SD = 6.5$), $p = 0.04$. Enrollees reported a lower percentage having only a high school education or less (15.9%) than non-enrollees (25.1%), $p < 0.001$. A lower percentage of enrollees were African American (12.5%) than non-enrollees (20.1%), and Hispanic (4.4%) than non-enrollees (9.9%), $p = < 0.001$ for race overall. A higher percentage of enrollees reported < 1 drink per day (66.1%) than non-enrollees (59.7%), $p = 0.008$. A higher percentage of enrollees (28.5%) than non-enrollees (23.3%) had BMI's of 20-25 kg/m^2, $p = 0.06$ overall. For enrollees, years since last regular menstrual period for women with prior hysterectomy were somewhat fewer ($M = 12.6$, $SD = 6.1$) than non-enrollees (M = 13.6, SD = 5.8), $p = 0.05$.

There was a greater percentage of enrollees (54.6%) than non-enrollees (51.8%) who were 0-5 years since their last regular menstrual period, and a greater percentage of enrollees (21.5%) than non-enrollees (17%) who were 6-10 years since their last regular menstrual period, and a smaller percentage of enrollees (23.9%) than non-enrollees (30.1%) who were 11 plus years since last regular menstrual period, p=.003 overall.Prior HT at WHI recruitment was less in enrollees (34.4%) than in non-enrollees (38.6%), p = 0.001. On-trial adherence and exposure based on average pill counts was greater in enrollees (M = .82, SD = .21) than in non-enrollees (M = .79, SD = .23), p = 0.003, as was length of enrollment in number of years on study, (M = 5.43, SD = 2.48) and (M = 5.08, SD = 2.54), p = 0.002.

In the comparison of WHIMS-Y enrollees by arm (CEE vs. CEE+MPA), there were significant differences in age, age at last menstrual period, education, race, BMI, hypertension status, years since last regular menstrual period, prior HT at WHI screening, and years of adherence. Women in the CEE group were slightly younger (M = 51.9, SD = 1.4) than in the CEE+MPA group (M = 52.2, SD = 1.3), p = .002, were younger at their last menstrual period (M = 39.1, SD = 6.0) than in CEE+MPA (M = 48.2, SD = 3.4), p < .001, and had less education (77.9% with some college) than in CEE+MPA (87.4% with some college), p = .001. A higher percentage of African American women were in the CEE group (17.4%) than CEE+MPA (9.1%), and a slightly lower percentage of Hispanic women were in the CEE group (3.8%) than CEE+MPA (4.4%), p = .007 for race overall. The percentage of CEE women with BMI 20-25 kg/m^2 was lower in enrollees (24.8%) than non-enrollees (33.2%), p < 0.001. A higher percentage of CEE women had hypertension (26.9%) than CEE+MPA (17.5%), p = .003. A lower percentage of CEE women were 0-5 years since their last regular menstrual period (13.7%) than CEE+MPA (76.6%), and a higher percentage of CEE women were 6-10 years since their last regular menstrual period (28.0%) than CEE+MPA (17.3%) and 11+ years since their last regular menstrual period (58.2%) than CEE+MPA (6.1%), p < .001 overall. In the CEE group, a smaller percentage of women (45.1%) reported never having had prior HT at screening than in CEE+MPA (56.4%), p = .01 overall. Years of adherence was greater in the CEE group (M = 6.12, SD = 2.69) than CEE+MPA (M = 5.11, SD = 2.21), p < .001, although the percentage of women adhering did not differ (p = .70).

Baseline cognitive characteristics of the WHIMS-Y participants at the first test are included in Table 2. These include measures of memory: the East Boston Memory Test (EBMT) total score and the EBMT recall score; attention and executive function: the Trails Part A and Part B seconds; working memory: Digits Forward and Digits Backward span scores; and verbal fluency: Verbal Fluency Animals (VFA) word counts.

A confirmatory factor analysis (CFA) of the cognitive measures was performed using Mplus version 6.2 (Muthén & Muthén, 2011) to examine the underlying construct as well as the amount of variance in cognition explained by the different measures (see Table 3). First, the overall omnibus fit of the model (e.g., whether or not the sample variance-covariance matrix S is similar to the population variance-covariance matrix Σ) was tested. Typically a non-significant chi-square is the gold standard for determining model fit, but chi-square is sensitive to sample size. With large sample sizes (>1000), the chi-square values may be inflated (statistically significant), and erroneously imply a poor data-to-model fit (Schumacker & Lomax, 2004). A number of additional indices were examined to determine model fit, including the Root Mean Square Error of Approximation (RMSEA), the Standardized Root Mean Square Residual (SRMR), and the Comparative Fit Index (CFI). The RMSEA and SRMR were evaluated for overall fit. The RMSEA, widely reported, estimates the amount of error of approximation (the lack of fit between the hypothesized model and the population covariance matrix) per model degree of freedom, taking sample size into account (Kline, 2005). A value of less than or equal to 0.05 indicates a good fit.

The SRMR is the standardized difference between observed and predicted covariances. Low SRMR values indicate that the residual matrix of S used to estimate Σ is near zero in a good model. A CFI close to 1.0 indicates a high degree of fit. A one-factor model of cognitive function fit the data, $X^2 = 17.36$, $df = 11$, $p = 0.10$ (RMSEA = .02; 90% CI = .000 - .039; SRMR = .02; CFI = .99). All of the fit indices reported here indicate that model fit is good to excellent. The significance of the individual parameters (path estimates) was also evaluated for their contribution to the model. Each factor loading was significant (all p's < 0.001), and each measure explained a significant amount of variance in the model, all p's < 0.005, indicating that the construct of cognitive function was well-measured by the telephone-administered cognitive battery.

3. Discussion

The analyses presented in this paper reveal several challenges to the WHIMS-Y study. Recruitment was lower than projected, potentially resulting in less statistical power, and limiting the ability to make comparisons between study arms. Post hoc assessments of power may be important to interpret results. There were marked differences between women who did and did not enroll in WHIMS-Y with respect to characteristics at the start of the WHI. As in many other settings (e.g., Lovato et al., 1997), enrollees tended to be more highly educated, of majority ethnicity, and differed with respect to several health-related characteristics. If enrollment over-sampled women who had high levels of cognitive function, the ability to detect intervention effects may be limited. However, we are encouraged in this regard from results of the post-trial follow-up of the WHIMS cohort, which found that on-trial treatment-related deficits in the women were largely maintained during post-trial follow-up (Espeland et al., 2010). Importantly, with respect to the factors that we considered, the original balance between active and placebo treatment groups afforded by randomization appeared to be maintained. This balance would be expected to extend to cognitive function at baseline, which was not measured. The time between age at last regular period and age at screening varied greatly among women and, for women with prior hysterectomy, averaged over 12 years. This, coupled with use of HT during this period that was not controlled by study design, may complicate how WHIMS-Y study results can be applied to a window of opportunity hypothesis. We would have greater power (and require fewer participants) if we had baseline measures of cognitive function to use as covariates (or to produce change scores) in our analyses. WHIMS-Y, however, has been designed to provide adequate power, as described above, and the WHI collected data on many baseline factors shown to be strong correlates of cognitive function, which we can use as covariates in analyses to reduce variability. These include age, hypertension, education, ethnicity, smoking, alcohol intake, diabetes status, and history of stroke. Including these measures in analyses increases the alignment of groups with respect to baseline levels of cognitive function and increases power. An additional non-statistical concern is that our results may be discounted by some because of the lack of baseline data to express changes. We will address this issue directly in publications, citing examples of the demonstrated validity of post-randomization add-on measures, including the many WHI post-trial publications.

4. Methods and Materials

4.1 Design of the Women's Health Initiative (WHI) hormone therapy (HT) trials

The WHI randomized trials were designed to evaluate postmenopausal hormone therapy and prevention of disease, with coronary heart disease as the primary outcome and breast cancer as the primary adverse outcome (Wactawski-Wende, 1998). Secondary outcomes included hip fracture, other fractures, other cardiovascular diseases, and endometrial, colorectal, and other cancers. Postmenopausal women, aged 50 - 79 at baseline, were randomized to either

placebo or oral CEE 0.625 mg/day if they did not have a uterus, and women with a uterus were assigned randomly to either placebo or oral CEE plus MPA 2.5 mg per day (PremPro, Wyeth Pharmaceuticals, Philadelphia, PA). Enrollment began in 1993 and continued through 1998. Both trials were scheduled to continue through 2005, however study medications were terminated on July 9, 2002 in the CEE plus MPA trial after 5.2 years of follow-up due to an increase in the global risk (increased risks of invasive breast cancer, coronary heart disease, stroke and pulmonary embolism) (WHI, 2002) and study medications in the CEE only trial were terminated February 29, 2004 after a mean of 6.8 years of follow up due to a lack of benefit and a significantly increased risk for stroke (Anderson et al., 2004).

At that time, all participants enrolled in one or more of the original WHI study components (including the Hormone Therapy Trials) who were willing to provide written informed consent could join the first WHI Extension Study for an additional 5 years of follow-up. In 2010, the current participants were invited to continue for an additional 5 years for the second Extension Study. The longer follow-up continues to provide important information on outcomes that might be affected by study treatments years after the initiation of intervention, and on outcomes that were too uncommon for clear results to emerge during the initial follow-up period.

4.2 Design of the Women's Health Initiative Memory Study-Younger (WHIMS-Y)

4.2.1 Participants—WHIMS-Y volunteers had enrolled in the Women's Health Initiative HT Trial (WHI, 1998) when they were aged 50-54 years. To be eligible for WHIMS-Y, they were required to be currently active participants of the WHI Extension Study, allow a friend or family member to be contacted, have adequate hearing acuity to participate in the telephone interviews, and agree to undergo annual 60 minute telephone-based assessments of their cognitive function. Enrollment began in January, 2009 and continued through September, 2011. All women provided written informed consent and all protocols were approved by local Institutional Review Boards.

4.2.2 Procedures—A brief hearing screening test, performed annually, determined the participant's ability to hear over the telephone. Each woman was asked standard questions (e.g., Do you have difficulty hearing others in a quiet room?) and administered a brief hearing screening test requiring participants to repeat several phrases over the phone (e.g., "I have a dog so all I need is a cat."). Women who reported no or minimal hearing difficulty and who repeated the phrases correctly were enrolled and given the telephone-based cognitive testing. Women who were unable to successfully complete the hearing screening test after attempts to compensate for their hearing loss were ineligible to participate.

4.2.3 Cognitive test battery—The telephone cognitive battery included tests of global cognitive functioning, verbal memory, attention, executive function, verbal fluency, and working memory plus self-report measures of perceived memory problems, depressive symptoms, sleep disturbance and health-related quality of life. A detailed description of each of the measures and their outcomes is included in Supplementary Data. Test-retest reliability, concurrent validity, and relative bias associated with telephone administration has been evaluated in older women and it has been found to be both reliable and valid (Rapp et al., 2012).

4.2.4 Adjudication of Mild Cognitive Impairment and Dementia—The Dementia Questionnaire (DQ) (Kawas et al., 1994) is a semi-structured interview designed for a knowledgeable proxy to provide information needed to make dementia and mild cognitive impairment diagnoses and to identify selected causes of cognitive impairment. It covers six

domains: memory and cognition, verbal expression, daily functioning, recognition of problems/insight, other medical and psychiatric problems, and medical contacts. Knowledgeable friends or family members also estimate the years of symptom onset. The DQ has been validated against the gold-standard of a clinical evaluation with sensitivities and specificities >90% and inter-rater (face to face vs. phone) agreement of >94% (Kawas et al., 1994; Ellis et al., 1998; Khachaturian et al., 2000). Proxy data from the DQ (Kawas et al., 1994) were collected on participants scoring below the cut-point (<31) on the Telephone Interview for Cognitive Status – modified (TICSm) (Welsh et al., 1993; DeJager et al., 2003a).

The TICSm with the DQ has been previously validated for identifying dementia cases in community samples in older adults age 65 and over (Kawas et al., 1994; Khachaturian et al., 2000; Fritsch et al., 2005) with sensitivity of 83%, specificity of 100%, and agreement with face-to-face clinical evaluation of 89% (Crooks et al., 2005). Additional data on the utility of the TICS to detect cognitive impairment in older adults > 60 showed a normal distribution of test scores and less constraint from ceiling effects than the MMSE and the CAMCOG (de Jager, Budge, & Clarke, 2003). A scoring algorithm based on selected DQ responses was used to pre-classify all participants as potential *probable dementia* or *minor cognitive impairment* cases. Final adjudication of the pre-classified cases was made by a panel of specialists comprised of clinicians with recognized expertise in dementia. The adjudication panel made the final determination in cases of conflicting classifications.

4.2.5 Sample size justification—WHIMS-Y targeted the recruitment of approximately 2,200 women. Power was projected for comparing mean differences between treatment groups on the primary outcome (TICSm) across the first two annual administrations. Data from an alternative measure of global cognitive function, the Modified Mini-Mental State Exam (3MS) (Teng et al., 1987) collected by the WHIMS (Shumaker et al., 2003), was substituted for TICSm to project power. Variance estimates were derived from 3,408 WHIMS enrollees who were aged 65-69 at enrollment. Of these, 3,223 (94.6%) provided a one-year cognitive assessment. The standard error associated with a contrast of mean differences in 3MS scores over the two assessments between women assigned to HT versus placebo was 0.125 units. Based on this, 2,200 women would be expected to provide a standard error of 0.150 units. For z-tests of this contrast (2-sided Type 1 error = 5%), powers of 76%, 91%, and 98% power were projected to detect a mean difference of 0.4, 0.5, and 0.6 units, respectively, which corresponded to approximately 0.08 to 0.12 standard deviation units.

4.2.6 WHIMS-Y analysis plans—The primary outcomes for WHIMS-Y are global cognition scores from the TICSm, which will be collected twice, approximately one year apart. The primary contrast for these data will be the mean difference in TICSm scores over time between women grouped by WHI treatment assignment (i.e. HT versus placebo), based on general linear models [Littell, 1996]. Time between WHI enrollment and the date of WHIMS-Y enrollment is a covariate. Type I error will be set to be 0.05 for the primary comparison of these mean differences. Analyses will be supported by examining patterns of missing data (i.e. individuals for whom the second exam was not conducted), to assess whether these are balanced across intervention groups. In addition, the characteristics of women who chose to enroll in WHIMS-Y will be compared to non-enrollees to assess whether there may have been differential enrollment between intervention groups, which is described in this paper.

Data from the remaining cognitive assessment instruments collected by WHIMS-Y will be analyzed in a similar manner. To protect against Type I error and to demarcate clearly the primary outcome measure, these secondary outcomes will be tested against a significance

level of 0.01. Results from central adjudications to detect cases of MCI and PD will be summarized, however it is anticipated that the power for any comparisons between treatment groups will be low for these outcomes.

Three subgroup comparisons were pre-specified to examine the consistency of any effects of HT among women grouped by 1) age of menopause, 2) prior use of HT, and 3) WHI trial (i.e. CEE+MPA versus CEE-Alone), using tests of interaction terms.

5. Conclusions

WHIMS-Y provides a critical opportunity to test the hypothesis that younger postmenopausal women aged 50-54 may benefit from hormone therapy. This data will have clinical significance for the timing and administration of HT as it relates to weighing the health risks versus benefits of the intervention. The integration of WHIMS-Y into the WHI suite of studies offers a cost-effective way to examine this hypothesis. In addition, it capitalizes on the experience of WHI and WHIMS investigators and staff in ensuring high retention and excellent quality assurance across sites. Important strengths of the WHIMS-Y design include a distribution of risk factors for cognitive impairment at the time of WHI enrollment unrelated to treatment assignment, as well as the construct validity of the telephone-administered cognitive battery. Related strengths include a high retention rate and better than average on-trial adherence and exposure in enrollees. Challenges of conducting WHIMS-Y include lower than projected enrollment, as well as differential enrollment between WHIMS-Y enrollees and non-enrolles (although differential enrollment favored this study in certain instances (e.g., lower levels of prior HT, greater on-trial adherence and exposure, and more years on study).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

WHIMS-Y has been funded in whole or in part with Federal funds from the National Institute on Aging, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN-271-2011-00004C, and the National Heart, Lung, and Blood Institute, National Institutes of Health, under Contract No. N01-WH-4-4221. The WHI program is funded by the National Heart, Lung and Blood Institute, U.S. Department of Health and Human Services.

References

Abraham E, Axelrod BN, Ricker JH. Application of the Oral Trail Making Test to a mixed clinical sample. Archives of Clinical Neuropsychology. 1996; 11:697–701.

Albert M, Smith LA, Scherr PA, Taylor JO, Evans DA, Funkenstein HH. Use of brief cognitive tests to identify individuals in the community with clinically diagnosed Alzheimer's disease. International Journal of Neuroscience. 1991; 57:167–178. [PubMed: 1938160]

Anderson GL, Limacher M, Assaf AR, Bassford T, Beresford SA, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: The Women's Health Initiative randomized controlled trial. Journal of the American Medical Association. 2004; 291:1701–12. [PubMed: 15082697]

Benton A. Differential behavioral effects in frontal lobe disease. Neuropsychologia. 1968; 6:53–60.

Brandt J, Spencer M, Folstein MF. The telephone interview for cognitive status. Neuropsychiatry, Neuropsychology, & Behavioral Neurology. 1988; 1:111–117.

Burke WJ, Roccaforte WH, Wengel SP. The short form of the Geriatric Depression Scale: A comparison with the 30-item form. Journal of Geriatric Psychiatry and Neurology. 1991; 4:173–78. [PubMed: 1953971]

Craig MC, Maki PM, Murphy DG. The Women's Health Initiative Memory Study: Findings and implications for treatment. Lancet Neurology. 2005; 4:190–194. [PubMed: 15721829]

Crooks VC, Clark L, Petitti DB, Chui H, Chiu V. Validation of multi-stage telephone-based identification of cognitive impairment and dementia. BMC Neurology. 2005; 5:8. [PubMed: 15829005]

De Craen AJ, Heeren TJ, Gussekloo J. Accuracy of the 15-item geriatric depression scale GDS-15 in a community sample of the oldest old. International Journal of Geriatric Psychiatry. 2003; 18:63–66. [PubMed: 12497557]

De Jager CA, Budge MM, Clarke R. Utility of TICS-M for the assessment of cognitive function in older adults. International Journal of Geriatric Psychiatry. 2003; 18:318–324. [PubMed: 12673608]

Ellis RJ, Jan K, Kawas C, Koller WC, Lyons KE, Jeste DV, Hansen LA, Thal LJ. Diagnostic validity of the dementia questionnaire for Alzheimer disease. Archives of Neurology. 1998; 55:360–365. [PubMed: 9520010]

Espeland MA, Brunner RL, Hogan PE, Rapp SR, Coker LH, Legault C, Granek I, Resnick SM. Long-term effects of conjugated equine estrogen therapies on domain-specific cognitive function: Results from the Women's Health Initiative study of cognitive aging extension. Journal of the American Geriatric Society. 2010; 58:1263–1271.

Espeland MA, Rapp SR, Shumaker SA, Brunner R, Manson JE, Sherwin BB, Hsia J, Margolis KL, Hogan PE, Wallace R, Dailey M, Freeman R, Hays J. Conjugated equine estrogens and global cognitive function in postmenopausal women: The Women's Health Initiative Memory Study. Journal of the American Medical Association. 2004; 291:2959–68. [PubMed: 15213207]

Folstein MF, Folstein SE, McHugh PR. 'Mini Mental State': a practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatry. 1975; 12:189–198.

Fritsch T, Smyth KA, McClendon MJ, Ogrocki PK, Santillan C, Larsen JD, Strauss ME. Associations between dementia/mild cognitive impairment and cognitive performance and activity levels in youth. Journal of the American Geriatric Society. 2005; 53:1191–1196.

Gallo JJ, Breitner JC. Alzheimer's disease in the NAS-NRC Registry of aging twin veterans, IV: Performance characteristics of a two-stage telephone screening procedure for Alzheimer's dementia. Psychological Medicine. 1995; 25:1211–1219. [PubMed: 8637951]

Grodstein F, Chen J, Pollen DA, Albert MS, Wilson RS, Folstein MF, Evans DA, Stampfer MJ. Postmenopausal hormone therapy and cognitive function in healthy older women. Journal of the American Geriatric Society. 2000; 48:746–752.

Grodstein F, Chen J, Wilson RS, Manson JE. Type 2 diabetes and cognitive function in community-dwelling elderly women. Diabetes Care. 2001; 24:1060–1065. [PubMed: 11375371]

Henderson VW, Popat RA. Effects of endogenous and exogenous estrogen exposures in midlife and late-life women on episodic memory and executive functions. Neuroscience. 2011; 191:129–138. [PubMed: 21664950]

Henderson VW, Sherwin BB. Surgical versus natural menopause: Cognitive issues. Menopause. 2007; 14:572–579. [PubMed: 17476147]

Jarvenpaa T, Rinne JO, Raiha I, Koskenvuo M, Lopponen M, Hinkka S, Kaprio J. Characteristics of two telephone screens for cognitive impairment. Dementia and Geriatric Cognitive Disorders. 2002; 13:149–155. [PubMed: 11893836]

Joffe H, Hall JE, Gruber S, Sarmiento IA, Cohen LS, Yurgelun-Todd D, Martin KA. Estrogen therapy selectively enhances prefrontal cognitive processes: A randomized, double-blind, placebo-controlled study with functional magnetic resonance imaging in perimenopausal and recently postmenopausal women. Menopause. 2006; 13:411–22. [PubMed: 16735938]

Kawas C, Segal J, Stewart WF, Corrada M, Thal LJ. A validation study of the Dementia Questionnaire. Archives of Neurology. 1994; 51:901–906. [PubMed: 8080390]

Khachaturian AS, Gallo JJ, Breitner JC. Performance characteristics of a two-stage dementia screen in a population sample. Journal of Clinical Epidemiology. 2000; 53:531–540. [PubMed: 10812327]

Kline, RB. Principles and practice of structural equation modeling. Guilford Press; New York, NY: 2005.

Levine DW, Kripke DF, Kaplan RM, Lewis MA, Naughton MJ, Bowen DJ, Shumaker SA. Reliability and validity of the Women's Health Initiative Insomnia Rating Scale. Psychological Assessment. 2003; 15:137–148. [PubMed: 12847774]

Lines CR, McCarroll KA, Lipton RB, Block GA. Telephone screening for amnestic mild cognitive impairment. Neurology. 2003; 60:261–266. [PubMed: 12552041]

Littell, RC.; Milliken, GA.; Stroup, WW.; Wolfinger, RD. SAS System for Mixed Models. SAS Institute, Inc; Cary, NC: 1996.

Lord C, Buss C, Lupien SJ, Pruessner JC. Hippocampal volumes are larger in postmenopausal women using estrogen therapy compared to past users, never users and men: A possible window of opportunity effect. Neurobiology of Aging. 2008; 29:95–101. [PubMed: 17030472]

Lovato LC, Hill K, Hertert S, Hunninghake DB, Probstfield JL. Recruitment for controlled clinical trials: Literature summary and annotated bibliography. Controlled Clinical Trials. 1997; 18:328–52. [PubMed: 9257072]

Muthén, LK.; Muthén, BO. Mplus user's guide. 5th ed. Los Angeles, CA: 1998–2007.

Osborn DP, Fletcher AE, Smeeth L, Stirling S, Nunes M, Breeze E, Siu-Woon Ng E, Bulpitt CJ, Jones D, Tulloch A. Geriatric Depression Scale Scores in a representative sample of 14,545 people aged 75 and over in the United Kingdom: Results from the MRC Trial of Assessment and Management of Older People in the Community. International Journal of Geriatric Psychiatry. 2002; 17:375–382. [PubMed: 11994893]

Plassman B, Newman TT, Welsh KA, Helms MJ, Breitner J. Properties of the telephone interview for cognitive status. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. 1994; 7:235–241.

Rankin MW, Clemons TE, McBee WL. Correlation analysis of the in-clinic and telephone batteries from the AREDS cognitive function ancillary study. AREDS Report No. 15. Ophthalmic Epidemiology. 2005; 12:271–277. [PubMed: 16033748]

Rapp SR, Espeland MA, Shumaker SA, Henderson VW, Brunner RL, Manson JE, Gass ML, Stefanick ML, Lane DS, Hays J, Johnson DC, Coker LH, Dailey M, Bowen D. Effect of estrogen plus progestin on global cognitive function in postmenopausal women: The Women's Health Initiative Memory Study: A Randomized Controlled Trial. Journal of the American Medical Association. 2003; 289:2663–2772. [PubMed: 12771113]

Rapp SR, Legault C, Espeland M, Resnick S, Hogan P, Coker L, Dailey M, Shumaker S. Validation of a cognitive assessment battery administered by telephone. Journal of the American Geriatric Society. 2012; 60:1616–1623.

Reitan, R. Trail Making Test: Manual for administration and scoring. Reitan Neuropsychological Laboratory; Tucson, AZ: 1992.

Resnick SM, Henderson VW. Hormone therapy and risk of Alzheimer disease: A critical time. Journal of the American Medical Association. 2002; 288:2170–2172. [PubMed: 12413378]

Resnick SM, Maki PM, Rapp SR, Espeland M, Brunner R, Coker L, Granek I, Hogan P, Ockene J, Shumaker SA. Effects of combination estrogen plus progestin hormone treatment on cognition and affect: The Women's Health Initiative Study of Cognitive Aging (WHISCA). Journal of Endocrinology and Clinical Metabolism. 2006; 91:1802–1810.

Resnick SM, Espeland MA, An Y, Maki PM, Coker LH, Jackson R, Stefanick SML, Wallace R, Rapp SR. Effects of conjugated equine estrogens on cognition and affect in postmenopausal women with prior hysterectomy. Journal of Clinical Endocrinology and Metabolism. 2009; 94:4152–4161. [PubMed: 19850684]

Schumacker, RE.; Lomax, RG. A beginner's guide to structural equation modeling. 2nd ed. Erlbaum; Mahwah, NJ: 2004.

Shumaker SA, Legault C, Kuller L, Rapp SR, Thal L, Lane DS, Fillit H, Stefanick ML, Hendrix SL, Lewis CE, Masaki K, Coker LH. Conjugated equine estrogens and incidence of probable dementia and mild cognitive impairment in postmenopausal women: The Women's Health Initiative Memory Study. Journal of the American Medical Association. 2004; 291:2947–58. [PubMed: 15213206]

Shumaker SA, Legault C, Rapp SR, Thal L, Wallace RB, Ockene JK, Hendrix SL, Jones BN, Assaf AR, Jackson RD, Kotchen JM, Wassertheil-Smoller S, Wactawski-Wende J. Estrogen plus

progestin and the incidence of dementia and mild cognitive impairment in postmenopausal women: The Women's Health Initiative Memory Study: A Randomized Controlled Trial. Journal of the American Medical Association. 2003; 289:2651–2662. [PubMed: 12771112]

Sloan FA, Wang J. Disparities among older adults in measures of cognitive function by race or ethnicity. The Journals of Gerontology: Series B Psychological Sciences and Social Sciences. 2005; 60:242–250.

Teng EL, Chui H. The Modified Mini-Mental State (3MS) examination. Journal of Clinical Psychiatry. 1987; 48:314–318. [PubMed: 3611032]

Wactawski-Wende J, et al. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Controlled Clinical Trials. 1998; 19:61–109. [PubMed: 9492970]

Wechsler, D. Weschler Adult Intelligence Scale - Revised. Psychological Corporation; New York: 1981.

Wechsler, D. The Wechsler Memory Scale. 3rd Edition. Psychological Corporation, Harcourt, Inc; 1996. WMS-III

Welsh KA, Breitner J, Magruder-Habib KM. Detection of dementia in the elderly using the telephone interview for cognitive status. Neuropsychiatry, Neuropsychology, & Behavioral Neurology. 1993; 6:103–110.

Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. Controlled Clinical Trials. 1998; 19:61–109. [PubMed: 9492970]

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. Journal of the American Medical Association. 2002; 288:321–33. [PubMed: 12117397]

Yesavage JA. Geriatric Depression Scale. Psychopharmacology Bulletin. 1988; 24:709–71. [PubMed: 3249773]

Table 1

Distribution of risk factors for cognitive impairment at the time of WHI enrollment for women ages 50-54 who enrolled in WHI HT and later enrolled in WHIMS-Y compared to women who did not enroll in WHIMS-Y.

Characteristics at Time of WHI Enrollment	WHIMS-Y Enrollees WHI Assignment Mean, SD or N, %		WHIMS-Y HT vs. Placebo P-value	Non-WHIMS-Y enrollees WHI Assignment Mean, SD or N, %		Non-WHIMS-Y HT vs. Placebo P-value	WHIMS-Y vs. Non-WHIMS-Y P-value	WHIMS-Y Enrollees WHI Assignment Mean, SD or N, %		WHIMS-Y Enrollees CEE vs. CEE+MPA P-value
	HT N=696	Placebo N=630		HT N=350	Placebo N=346			CEE N=264	CEE+MPA N=432	
Age, yrs.	52.1, 1.3	52.0, 1.3	0.37	52.1, 1.3	52.1, 1.3	0.75	0.56	51.9, 1.4	52.2, 1.3	0.002
Age at last MP	45.0, 6.2	45.2, 6.1	0.76	44.7, 6.4	44.1, 6.6	0.23	0.04	39.1, 6.0	48.2, 3.4	<0.001
Education HS or less At least some college	112, 16.3 577, 83.7	97, 15.5 530, 84.5	0.70	86, 24.9 259, 75.1	86, 25.2 256, 74.9	0.95	<0.001	58, 22.1 204, 77.9	54, 12.7 373, 87.4	0.001
Race/Ethnicity, N % African-American Hispanic Non-Hispanic White Other/Multiple	85, 12.3 29, 4.2 557, 80.3 23, 3.3	80, 12.7 28, 4.5 509, 80.9 12, 1.9	0.46	64, 18.3 36, 10.3 241, 68.9 9, 2.6	76, 22.0 33, 9.5 230, 66.5 7, 2.0	0.65	<0.001	46, 17.4 10, 3.8 197, 74.6 11, 4.2	39, 9.1 19, 4.4 360, 83.7 12, 2.8	0.007
Smoking Status, N % Never Former Current	337, 48.6 263, 37.9 94, 13.5	299, 47.7 226, 36.0 102, 16.3	0.37	147, 42.2 140, 40.2 61, 17.5	171, 50.2 123, 36.1 47, 13.8	0.10	0.69	134, 51.2 92, 35.1 36, 13.7	203, 47.0 171, 39.6 58, 13.4	0.48
Alcohol intake None <1 per day ≥1 per day	158, 22.8 471, 68.1 63, 9.1	155, 24.8 401, 64.1 70, 11.2	0.26	107, 30.9 206, 59.5 33, 9.5	100, 29.2 207, 60.4 36, 10.5	0.84	0.008	72, 27.4 169, 64.3 22, 8.4	86, 20.1 302, 70.4 41, 9.6	0.08

Characteristics at Time of WHI Enrollment	WHIMS-Y Enrollees WHI Assignment Mean, SD or N, %		WHIMS-Y HT vs. Placebo P-value	Non-WHIMS-Y enrollees WHI Assignment Mean, SD or N, %		Non-WHIMS-Y HT vs. Placebo P-value	WHIMS-Y vs. Non-WHIMS-Y P-value	WHIMS-Y Enrollees WHI Assignment Mean, SD or N, %		WHIMS-Y Enrollees CEE vs. CEE+MPA P-value
	HT N=696	Placebo N=630		HT N=350	Placebo N=346			CEE N=264	CEE+MPA N=432	
Body Mass Index—kg/m^2, N %	16, 2.3	14, 2.2	0.62	5, 1.4	7, 2.0	0.21	0.06	2, 0.8	14, 3.3	<0.001
<20	208, 30.0	170, 27.1		89, 25.6	72, 20.9			65, 24.8	143, 33.2	
20-25	217, 31.3	192, 30.6		98, 28.2	120, 34.9			74, 28.2	143, 33.2	
25-29	145, 20.9	137, 21.9		89, 25.6	74, 21.5			73, 27.9	72, 16.7	
30-34	107, 15.4	114, 18.2		67, 19.3	71, 20.6			48, 18.3	59, 13.7	
≥35										
Hypertension Status, N %	547, 78.9	497, 79.1	0.93	270, 77.4	268, 77.7	0.92	0.43	193, 73.1	354, 82.5	0.003
No	146, 21.1	131, 20.9		79, 22.6	77, 22.3			71, 26.9	75, 17.5	
Yes										
Prior CVD, N %	542, 92.5	510, 92.9	0.79	273, 90.1	282, 91.3	0.62	0.14	200, 91.3	342, 93.2	0.41
No	44, 7.5	39, 7.1		30, 9.9	27, 8.7			19, 8.7	25, 6.8	
History CVD										
Hysterectomy, N %	432, 62.1	382, 60.6	0.59	220, 62.9	193, 55.8	0.06	0.37	0	432, 100.0	--
No	264, 37.9	248, 39.4		130, 37.1	153, 44.2			264, 100.0	0	
Yes										
Years since last regular menstrual period	12.7, 6.1	12.5, 6.0	0.76 0.51	13.6, 5.7	13.6, 5.9	0.99 0.42	0.05 0.46	12.7, 6.1	-- 4.0, 3.4	-- --
Prior hysterectomy	4.0, 3.4	3.9, 3.1		4.0, 3.1	4.3, 3.7			--		
No prior hysterectomy										
Years since last regular menstrual period	287, 54.8	270, 54.4	0.92	150, 53.0	141, 50.5	0.45	0.003	25, 13.7	262, 76.6	<0.001
0 to 5 years	110, 21.0	109, 22.0		51, 18.0	44, 15.8			51, 28.0	59, 17.3	
6 to 10 years	127, 24.2	117, 23.6		82, 29.0	94, 33.7			106, 58.2	21, 6.1	
11+ years										
Prior HT at WHI screening	336, 52.1	279, 46.8	0.09	165, 49.4	178, 53.1	0.63	0.001	110, 45.1	226, 56.4	0.01
Never	203, 31.5	222, 37.3		134, 40.1	124, 37.0			83, 34.0	120, 29.9	
Past	106, 16.4	95, 15.9		35, 10.5	33, 9.9			51, 20.9	55, 13.7	
Current										

Characteristics at Time of WHI Enrollment	WHIMS-Y Enrollees WHI Assignment Mean, SD or N, %		WHIMS-Y HT vs. Placebo P-value	Non-WHIMS-Y enrollees WHI Assignment Mean, SD or N, %		Non-WHIMS-Y HT vs. Placebo P-value	WHIMS-Y vs. Non-WHIMS-Y P-value	WHIMS-Y Enrollees WHI Assignment Mean, SD or N, %		WHIMS-Y Enrollees CEE vs. CEE+MPA P-value
	HT N=696	Placebo N=630		HT N=350	Placebo N=346			CEE N=264	CEE+MPA N=432	
Adherence Exposure: Adherence × Years	0.81, 0.21 5.51, 2.45	0.82, 0.20 5.34, 2.51	0.43 0.24	0.79, 0.22 5.15, 2.48	0.78, 0.23 5.00, 2.60	0.34 0.43	0.003 0.002	0.82, 0.20 6.12, 2.69	0.81, 0.22 5.11, 2.21	0.70 <0.001

[1] Other CVD defined as MI, angina, PCTA, stroke, or CABG

Table 2

Descriptive statistics for the dependent measures used in the confirmatory factor analysis of Cognitive Function (N = 1326).

Measure of Cognitive Function	Mean	SD	Skewness	Kurtosis
EBMT	10.08	1.68	−0.88	1.27
VFA	20.18	5.79	0.50	0.80
TRAILSA	8.57	2.21	2.67	16.98
TRAILSB	37.69	29.73	5.49	41.79
DIGITSF	8.64	2.55	0.20	−0.70
DIGITSB	7.13	2.58	0.65	0.07
ERECALL	9.68	1.93	−1.25	3.37

Note. EBMT = East Boston Memory Test Total Score; VFA = Verbal Fluency Animals word count; TRAILA = Trails Part A time in secs.; TRAILSB = Trails Part B time in secs.; DIGITF = Digits Forward score; DIGITB = Digits Backward score; ERECALL = East Boston Memory Test recall score.

Table 3

Fit indices and factor loadings for the confirmatory factor analysis model of Cognitive Function in WHIMS-Y participants at first visit ($N = 1326$).

Model	df	X^2	RMSEA	SRMR	CFI
a. One factor Cognitive Function	11	17.36, $p = .10$	.02 (90% CI = .000 −.039)	.02	.99
Factor and item	Factor loading (p value)				
CF → EBMT	.28 (.000)				
CF → VFA	.68 (.000)				
CF → TRAILSA	−.21 (.000)				
CF → TRAILSB	−.72 (.000)				
CF →DIGITSF	.27 (.000)				
CF → DIGITSB	.26 (.000)				
CF → ERECALL	.31 (.000)				
Correlations between specified measures in the model					
TRAILSB ↔ VFA	.61 (.000)				
DIGITSB ↔ DIGITSF	.58 (.000)				
ERECALL ↔ EBMT	.68 (.000)				

Note. EBMT = East Boston Memory Test Total Score; VFA = Verbal Fluency Animals word count; TRAILA = Trails Part A time in secs.; TRAILSB = Trails Part B time in secs.; DIGITF = Digits Forward score; DIGITB = Digits Backward score; ERECALL = East Boston Memory Test recall score.