



Published in final edited form as:

*Brain Res.* 2011 March 16; 1379: 232–243. doi:10.1016/j.brainres.2010.11.030.

## Perimenopausal Use of Hormone Therapy is Associated with Enhanced Memory and Hippocampal Function Later in Life

Pauline M. Maki, PhD<sup>a</sup>, Lorraine Dennerstein, MBBS PhD DPM<sup>b</sup>, Margaret Clark, PhD<sup>b</sup>, Janet Guthrie, MSc, PhD<sup>b</sup>, Pamela LaMontagne<sup>a</sup>, Deanne Fornelli<sup>a</sup>, Deborah Little, PhD<sup>c</sup>, Victor W. Henderson, MD MS<sup>c</sup>, and Susan M. Resnick, PhD<sup>d</sup>

Pauline M. Maki: pmaki@psych.uic.edu; Lorraine Dennerstein: lorrden@aol.com; Margaret Clark: msclark@unimelb.edu.au; Janet Guthrie: janetrg@unimelb.edu.au; Pamela LaMontagne: pperschler@gmail.com; Deanne Fornelli: deanne.fornelli@gmail.com; Deborah Little: little@uic.edu; Victor W. Henderson: vhenderson@stanford.edu; Susan M. Resnick: resnicks@grc.nia.nih.gov

<sup>a</sup>Departments of Psychiatry and Psychology, Neuropsychiatric Institute MC 913, University of Illinois at Chicago, 912 S Wood St, Chicago IL 60612 USA

<sup>b</sup>Office for Gender & Health and National Ageing Research Institute, Department of Psychiatry, The University of Melbourne, , Victoria, 3050 Australia

<sup>c</sup>Department of Neurology & Rehabilitation (MC 796), 912 S. Wood St Suite 855 – N, Chicago IL 60612, USA

<sup>d</sup>Departments of Health Research & Policy, Neurology & Neurological Sciences, Stanford University 300 Pasteur Dr, A301 MC 5325 Stanford, CA 94305, USA

<sup>e</sup>National Institute on Aging, Laboratory of Personality and Cognition, Biomedical Research Center, room 04B317, 251 Bayview Boulevard, Suite 100, Baltimore, MD 21224-6825

### Abstract

Evidence suggests that initiation of some forms of hormone therapy (HT) early in the perimenopausal or postmenopausal stage might confer benefit to verbal memory and the neural systems underlying memory, whereas late-life initiation confers no benefit or harm. This “critical window hypothesis” remains a topic of debate. Using functional magnetic resonance imaging (fMRI), we examined the long-term impact of perimenopausal HT use on brain function during performance of verbal and figural memory tasks. Participants were 34 postmenopausal women (mean age 60 years) from the Melbourne Women’s Midlife Health Project and included 17 early (perimenopausal) and continuous users of HT and 17 never users matched on age, education, and verbal knowledge. Continuous HT use from the perimenopausal stage versus no use was validated with prospective daily diary records and study visit data. The primary outcome was patterns of brain activation in an *a priori* region of interest in the medial temporal lobe during verbal encoding and recognition of words. Results indicated that perimenopausal HT users performed better than nonusers on the imaging verbal memory task ( $p < .05$ ). During verbal recognition, perimenopausal HT users showed increased activation in the left hippocampus and decreased activation in the parahippocampal gyrus bilaterally compared with never users. Each of these patterns of activation was associated with better memory performance on the imaging memory

© 2010 Elsevier B.V. All rights reserved.

Address all correspondence and requests for reprints to: Pauline M. Maki, PhD Neuropsychiatric Institute MC 913, University of Illinois at Chicago, 912 S Wood St, Chicago IL 60612. Phone: (312)996-6941; fax (312)413-7856.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal Disclaimers that apply to the journal pertain.

task. These results suggest that perimenopausal use of HT might confer long-term benefits to verbal memory and the brain systems underlying verbal memory. More generally, the results support the critical window hypothesis.

## Keywords

hormone therapy; estrogen; memory; fMRI; hippocampus; menopause

---

## 1. Introduction

Before the publication of findings from the Women's Health Initiative (WHI) in 2002 (Rossouw et al., 2002), 17% of women on hormone therapy (HT) in the United States began HT before reaching the final menstrual period, and 65% of women began HT within two years of the final menstrual period (Brett and Chong, 2001). The observational studies of HT prior to the WHI results therefore primarily involved women who initiated treatment in the perimenopausal or early postmenopausal stage. Some evidence suggests that certain forms of HT may confer cognitive benefit when initiated in close temporal proximity to the menopausal transition, whereas later initiation may be neutral or detrimental to cognitive function (see (Henderson, 2009; Maki and Sundermann, 2009) for a review). This hypothesis implies a "critical window" for beneficial action of HT (Resnick and Henderson, 2002).

Basic science studies provide support for this hypothesis, particularly as it applies to memory performance and the structure and function of medial temporal lobe structures subserving memory performance (Gibbs, 2000; Silva et al., 2003). Parallel human studies investigating the effects of early HT use on verbal episodic memory and medial temporal lobe function later in life are lacking. A recent neuroimaging study found that early HT use was associated with enhanced medial temporal lobe function during a working memory task, and that enhanced hippocampal and medial temporal activation predicted better working memory (Berent-Spillson et al.). The critical period hypothesis remains a topic of debate, in part because it is not feasible to conduct randomized clinical trials to test the effects of early HT use on cognitive function many years later, when memory problems may appear.

The aim of the present study was to use fMRI to determine the effects of perimenopausal initiation of HT (i.e., during the perimenopausal stage) on hippocampal activation during performance of memory tasks. Participants were a subsample of women from the Melbourne Women's Midlife Health Project, a prospective, population-based study of the natural history of the menopausal transition (Dennerstein et al., 1993). Samples of women who initiated HT before the final menstrual period and healthy controls who never received HT were selected on the basis of prospective daily diaries and study visit data, providing a valid means of verifying the timing of HT initiation in relation to the final menstrual period. Primary outcomes focused on hippocampal activation during performance of a verbal memory test, and secondary analyses focused on nonverbal (figural) memory. The verbal and figural imaging memory tasks were structured to resemble standardized memory tests commonly administered outside the scanner, and included an initial encoding phase followed by a 20-minute delay and then a yes/no recognition test for the words and figures presented earlier in the encoding phase. We hypothesized that women who initiated HT during the perimenopausal stage and continued using HT would show enhanced verbal memory and hippocampal/parahippocampal function in comparison to those who never initiated HT.

## 2. Results

### 2.1 Behavioral outcomes

Table 1 shows the demographic and behavioral outcome variables for perimenopausal HT users and never users and shows the results of t-tests comparing the two groups on these variables. T-tests were conducted on the neuropsychological tests administered out of the scanner for the full sample ( $n = 33$ ) of early versus never users. T-tests were conducted on the imaging memory task on the subsample ( $n = 25$ ) with valid imaging date (see Neuroimaging Outcomes below). Groups were well-matched on demographic variables and were similar to the full sample of 257 women from which they were drawn. The full sample mean age was 59.8 years, National Adult Reading Test (NART) scores were 34, and Center for Epidemiologic Studies Depression Scale (CES-D) scores were 6.8. On the full test battery, perimenopausal users tended to perform better than the never users on the immediate and delayed trials of the East Boston Memory Test (EBMT) ( $p < .10$ ), with medium to large effect sizes (Cohen's  $d = .62$  and  $.59$ , respectively). No other differences on the neuropsychological test battery approached statistical significance. On the imaging memory task, perimenopausal HT users showed a higher percent correct on the in-scanner verbal recognition task compared to never users ( $p < .05$ , Cohen's  $d = .70$ ). Perimenopausal users also tended to also show higher percent correct on the verbal matching task ( $p < .06$ , Cohen's  $d = .70$ ), reflecting a very small standard deviation. These differences were not observed on the figural task.

### 2.2 Neuroimaging outcomes

Of the 34 participants, 25 (13 perimenopausal users and 12 never users) had valid fMRI images. The other 8 participants had invalid imaging data because of excessive movement in the scanner ( $n = 3$ ), corrupted data ( $n = 2$ ), missing data ( $n = 2$ ), unusually large ventricles that affected normalization ( $n = 1$ ), and below-chance performance on the verbal recognition test only ( $n = 1$ ; a nonuser who was also  $> 3$  sd below the mean on the verbal recognition task). Table 2 shows regions in the medial temporal lobe where perimenopausal HT users and never users showed significant differences in activation for the six individual conditions (i.e., verbal encoding, verbal recognition, verbal match, figural encoding, figural recognition, and figural match), as well as differences in the primary contrasts of interest (i.e., encoding versus match; recognition versus match). Talairach coordinates in the sagittal (x), coronal (y), and axial (z) planes, anatomical labels, and Brodmann areas (BA) are presented for the voxels at the peak signal difference. The overall pattern of effects when contrasting early versus never users indicated that there were significant group differences in medial temporal activations during individual recognition and match conditions, but not during the individual encoding condition, and that there were more differences during the verbal task than the figural task. Therefore, differences between early and never users were most pronounced during recognition and match conditions of the verbal memory task.

Figure 1 shows the patterns of activation both on “glass brain” images and on a structural anatomical template for three of the regions where HT effects were significant. Compared with never users, perimenopausal HT users showed less activation in right and left parahippocampal gyrus in the perirhinal area (BA 35/36) (Rows 1 and 2, Figure 1) and greater activation in the left hippocampus (Rows 3 and 4) during verbal recognition. The same effect was observed during verbal match, the intended control condition, in the same area (Rows 5 and 6). There was greater activation in never users compared with perimenopausal users in the left hippocampus for the subtraction of verbal encoding minus verbal match. This difference was attributable to differences in the match (control) condition rather than the encoding condition, as no group differences were observed during encoding and perimenopausal users showed greater activation in the match condition.

Group differences in medial temporal areas were also observed during the figural task. Perimenopausal users showed greater activation in the left parahippocampal gyrus (BA 27/35) compared to never users. This area was superior and slightly more medial to the cluster in which they showed lower left parahippocampal activation during verbal recognition. In a comparison of figural recognition versus match, perimenopausal users showed significantly greater activation in the right parahippocampal gyrus (BA 35). Post-hoc analyses at a lower significance level ( $p = .10$ ) indicated that this difference was due to increased activation in the right parahippocampal gyrus during figural recognition in the perimenopausal users compared to the never users, rather than to decreased activation in never users compared to perimenopausal users during figural match.

Table 3 shows group activation maps in medial temporal regions across the two groups combined to characterize the typical pattern of activation on the cognitive tasks. The overall pattern of results revealed significant activations but no significant deactivations. Figure 2 shows the typical patterns of activation in the overall sample at voxels where hormone effects were highest (note comparison with Figure 1). Regions typically activated during task performance overlapped with regions showing significant HT effects for: verbal recognition in the left perirhinal parahippocampal gyrus BA 36 ( $-22, -36, -7$ ), the right perirhinal parahippocampal gyrus ( $20, -30, -12$ ), and the left hippocampus ( $-26, -22, -7$ ); and during verbal match in the left hippocampus ( $-24, -27, -4$ ), though the hormone effects were lateral to this area. There was no overlap in activations between the overall group and the HT group comparisons in any of the figural conditions or in any contrast between conditions, though the left parahippocampus was significantly active during figural recognition at a more liberal threshold of  $k = 20$  and  $p < .01$ . Finally, in a regression analysis, we examined the relationship between scores on the imaging verbal memory task and patterns of brain activation in the medial temporal regions of interest during performance of the verbal recognition task. This analysis addressed whether there was any relationship between the performance and the brain regions showing HT effects. We excluded one woman from the analysis who was an outlier on the verbal memory task to ensure that she was not driving any findings. Higher activation in the left hippocampus ( $-24, -18, -9$ ) was positively associated with verbal memory performance while lower activation in the left parahippocampal gyrus ( $-18, -28, -12$ ) was associated with better verbal memory performance ( $p < .05$ ). These activations overlapped with regions showing HT effects in a direction favoring perimenopausal HT users. There were no significant correlations between activation in the left hippocampus and match performance, nor did the magnitude of activation observed in the left hippocampus during match correlate with memory performance.

### 3. Discussion

The aim of this study was to investigate the influence of continuous use of HT from the perimenopausal stage on patterns of hippocampal activation during performance of a verbal memory test. To achieve this aim, we studied healthy midlife women whose timing of exposure to HT was prospectively validated. Seventeen women who initiated HT during the perimenopausal stage were compared with 17 controls who were similar in age, education, depressive symptoms, and estimated verbal intelligence, but who never used HT. Based on our previous neuroimaging studies in older long-term HT users (Maki and Resnick, 2000; Resnick et al., 1998), we hypothesized that HT would be associated with enhanced verbal memory performance and with differences in activation in both the parahippocampal gyrus and the hippocampus proper. Each of these hypotheses was supported.

Perimenopausal HT users compared to never users showed deactivation in the parahippocampal region. This deactivation contrasted with the typical pattern of increased

activation in the parahippocampal region observed across the combined sample of perimenopausal and never users. Perimenopausal HT users also showed increased activation in the left hippocampus. This activation was an exaggeration of a typical pattern of increased activation observed in the hippocampus across the combined sample. Notably, lower activation in the left parahippocampal gyrus and higher activation in the left hippocampus were each associated with better performance on the verbal memory task, whereas activation during the match task was unrelated to verbal recognition or match performance. The association between the activation pattern during verbal recognition in perimenopausal hormone initiators and better memory performance suggests that the alterations in brain function associated with perimenopausal initiation of HT are favorable. Overall, these results support the hypothesis that early and continued HT use confers a benefit to verbal memory performance and modulates hippocampal-parahippocampal function underlying verbal memory performance.

Significant performance differences on verbal memory tests in favor of the perimenopausal users were evident on the imaging verbal recognition test ( $p < .05$ , two-tail). A trend ( $p < .10$ ) was observed on both the immediate and delayed trials of the EBMT, a story recall test that was performed outside the scanner. Notably, the effect size for the immediate and delayed trials of the EBMT were .59 and 0.62 standard deviations, respectively, which according to standard classifications are substantive medium (0.50) to large (0.80) effect sizes. The lack of statistical significance for these medium to large effects stemmed from a small sample size. Our findings on the EBMT are consistent with previous placebo-controlled trials in surgically menopausal women showing that estradiol valerate (with or without testosterone enanthate) improves story recall (Sherwin, 1988); (Phillips and Sherwin, 1992). The importance of menopausal stage and hormone use on story recall is further supported by new findings from the Study of Women's Health Across the Nation (SWAN) based on 2362 women who completed the EBMT annually over a 6-year period as they transitioned from premenopause to postmenopause (Greendale et al., 2009). As in the present study, prior HT users in SWAN showed improved performance on the first administration of the story recall test compared to never users. Furthermore, perimenopausal women in SWAN did not show the same magnitude of improvement over time on the EBMT that was observed in premenopausal women ( $p < .10$ ), suggesting a decrease in story recall during the perimenopausal stage. Our data indicate that intervening during the perimenopausal stage with HT is associated with improvements in story recall later in life. Some of the benefit of HT on story recall may be associated with improvement in hot flashes. In a pilot study involving 29 early postmenopausal women with moderate to severe hot flashes, we found that lower performance on story recall was associated with an increase in physiological hot flashes, as measured by ambulatory skin conductance monitors, but not with self-reported hot flashes (Maki et al., 2008).

In contrast to the effect of HT on story recall, the effect of HT on memory for word lists was very small (i.e., range 0.5 to .20 standard deviations) in the present study. Results of a prior study from the Melbourne Women's Midlife Health Project ( $n = 326$ ) (Henderson et al., 2003a) suggest that larger samples may be needed to see an effect of perimenopausal HT use on word list memory. In that prior study, there was no overall effect of HT on word list memory (story recall was not assessed), but in post-hoc analyses women who initiated HT before the final menstrual period (i.e., during the perimenopausal stage) performed better on that memory test compared to women who initiated HT after the final menstrual period (Henderson et al., 2003a). Furthermore, word list recall was positively associated with years of HT. The Nurses Health Study found no effect of postmenopausal estrogen use within the three years following the final menstrual period on 2-year change in performance on a telephone version of the EBMT in a sample of older women (mean age = 74) (Kang et al., 2004). Together, these findings raise the possibility that HT intervention during the

perimenopausal stage rather than during the early postmenopausal stage may be key to enhancement in story recall later in life.

Based on results from our prior PET studies in elderly long-term HT users (Maki and Resnick, 2000; Resnick et al., 1998) and a wealth of basic science studies showing estrogen effects on the hippocampus (Gibbs, 2000), the present fMRI analyses focused on medial temporal lobe areas, specifically the parahippocampal gyrus and hippocampus proper. For the verbal task, group comparisons revealed significant increases in activation in perimenopausal HT users compared to controls in the left hippocampus. This increase was evident during verbal recognition and match conditions but not during verbal encoding. The left hippocampal regions subserving retrieval and matching of abstract words overlapped substantially, suggesting an overlap in cognitive processes mediated by the left hippocampus during the two conditions. The finding in HT users of increased hippocampal activation during the match task suggests that early initiators may have engaged in memory processing during the control task, particularly because the match task always followed recognition and hippocampal activation is commonly observed in verbal memory tasks (Cabeza and Nyberg, 2000; Schacter and Wagner, 1999). Analyses across nonusers and users combined showed that the left hippocampus is typically active during this verbal recognition memory task. In light of the perimenopausal user's significant advantage in the recognition condition and near-significant advantage in the match condition ( $p < .10$ , effect size = .70 sd), these results suggest that this augmentation of the typical patterns of left hippocampal activation is advantageous for perimenopausal users. Regression analyses support this interpretation as well, because higher activation in the left hippocampus was associated with better memory. More generally, the pattern of results suggests that beneficial effects of estrogen on medial temporal regions subserving verbal memory are observed in regions that subserve retrieval of studied words rather than encoding of to-be-learned words. There were no group differences in activation during verbal encoding alone, and a group difference observed when verbal encoding was compared to match was due to differences in the match condition.

Group differences in patterns of brain activation were also observed in the parahippocampal gyrus during verbal tasks. Specifically, perimenopausal HT users showed less activation compared to controls in the right and left parahippocampal gyri during verbal recognition. That result replicates our previous PET findings of enhanced verbal memory and decreased parahippocampal gyrus activation in HT users compared to nonusers during verbal recognition (Resnick et al., 1998), albeit in a more anterior and medial parahippocampal region corresponding to perirhinal cortex. The consistency of these findings across samples and imaging methods suggests that decreased parahippocampal activation may be a critical neural substrate underlying the association between use of HT and enhanced verbal memory. The typical pattern of activation in the parahippocampal gyrus is one of activation, as suggested by the analysis across the combined groups of perimenopausal users and nonusers shown in Figure 2. Insights into the role of the parahippocampal gyrus during word recognition tasks can be gained from a previous fMRI study which used a mixed blocked and event-related experimental design in order to dissociate item components associated with transient recovery of word-specific information from state components associated with a sustained retrieval mode across trials to meet ongoing goals (Donaldson et al., 2001). That study found sustained decreases in activation in bilateral regions of parahippocampal cortex, suggesting that the parahippocampal gyrus might become deactivated when participants entered into a sustained state of retrieval during recognition tasks. In the present study, regression analyses suggested that deactivation in the left parahippocampal gyrus relates to better verbal memory performance, and this effect was specific to the verbal recognition condition. Together, the findings of enhanced hippocampal activation and decreased parahippocampal activation in perimenopausal HT users suggest that perimenopausal HT

enhances both state-dependent and recollective processes contributing to verbal recognition performance.

In secondary analyses we examined the effects of perimenopausal HT on figural encoding and retrieval. In an earlier study we found greater parahippocampal deactivation and better figural memory among HT users compared to nonusers (Resnick et al., 1998). In contrast, in this study perimenopausal HT users showed an increase in parahippocampal gyrus activation and no advantage over nonusers in figural memory. It is possible that our earlier findings reflected an enhanced state-dependent retrieval mode during figural tasks, though this particular role of the parahippocampal gyrus has been demonstrated only during verbal memory tasks (Donaldson et al., 2001). A modest figural memory benefit ( $p < .01$ ) with combined conjugated equine estrogen plus medroxyprogesterone acetate (CEE+MPA) but not CEE alone was observed in WHISCA, a randomized clinical trial in women over age 65 (Resnick et al., 2004; Resnick et al., 2006). This suggests a possible influence of MPA on figural memory in older women. Note, however, that no benefits of CEE+MPA on figural memory were observed in a clinical trial of 180 younger postmenopausal women (mean age = 52 years) (Maki et al., 2007). These studies suggest that figural memory benefits may be most evident in samples of elderly women receiving combination HT rather than in midlife women.

Lastly, it is worth noting that the evidence of a functional enhancement in the hippocampus with perimenopausal hormone initiation contrasts with evidence of detrimental effects of HT on hippocampal structure from the Women's Health Initiative Memory Study (WHIMS) MRI substudy (Resnick et al., 2009). WHIMS-MRI examined brain volumes in 1,403 women aged 71–89 years who underwent brain scans on average 3.0 years post-trial for the CEE + MPA trial and 1.4 years post-trial for the CEE-alone trial. Total brain volumes, hippocampal volumes, and frontal lobe volumes were significantly lower among women randomized to receive HT compared to those randomized to receive placebo. Notably, the clinical significance of these volumetric differences was evident in the demonstration of a relationship between cognitive impairment (i.e., mild cognitive impairment or dementia) and loss of hippocampal and total brain volume among older women randomized to receive CEE-based therapies but not among older women randomized to receive placebo (Espeland et al., 2009).

There were several limitations to our study. Although we were careful to take specific steps to minimize biases associated with HT, we cannot rule them out entirely. Our fMRI study of the effects of early and continued HT use on brain function later in life was observational by design and was conducted almost 10 years after menopause, on average. To ensure validity of timing and duration of HT exposure, we examined prospective daily diary and medical records for women participating in the Melbourne Women's Midlife Health Project. We drew specifically from the pool of women who initiated HT before the final menstrual period and continued HT so that we could test the extremes of the critical window hypothesis, specifically to study whether early and continued use of HT confers benefits later in life. The users and nonusers of HT were carefully matched with respect to age, education, depression and estimated verbal intelligence. Valid ascertainment of HT exposure is critical because recall bias is common in self reports of both medication use and age at menopause (den Tonkelaar, 1997; West et al., 1995). In this observational study, the specific form of HT was not controlled and therefore the results reflect the most commonly used preparations, particularly estradiol (60%) and conjugated equine estrogens (35%). These results then reflect common effects of these two estrogen preparations, rather than effects particular to a certain preparation. Finally, our data do not address potential effects of perimenopausal HT on prefrontal cortex, a region shown to be functionally enhanced during an fMRI verbal memory task in randomized trials of HT (Joffe et al., 2006; Persad et al.,

2009) and in our previous PET studies (Maki and Resnick, 2000). Our image acquisition parameters were optimized to detect effects in medial temporal lobe structures so the number of superior slices to address frontal lobe effects was minimal. On the other hand, our functional data complement new structural neuroimaging data demonstrating larger hippocampal volumes among women who used HT at the time of menopause compared with women who had never used HT (Erickson et al.).

In summary, the present study examined the effects of continuous use of HT from the perimenopausal stage to later in life (i.e., mean age = 60 years) on verbal memory and the medial temporal structures subserving verbal memory. Our design was observational by design, given the substantive logistical and financial challenges necessary to examine this issue in a randomized controlled trial. One notable strength of our design is that exposure to HT was validated through prospectively collected research records. Additionally, our two comparison groups were well matched on potential confounding variables. Our behavioral and neuroimaging findings support the critical window hypothesis and suggest that perimenopausal HT is associated with enhanced verbal memory and enhanced hippocampal and parahippocampal function. We did not find similar support for the critical window hypothesis as applied to figural memory. The present finding raises the possibility that the critical window may begin in the perimenopausal stage. Future studies should directly contrast the effects of perimenopausal versus early postmenopausal HT use on verbal memory and the functional circuitry underlying verbal memory, including the hippocampus and prefrontal cortex.

## 4. Experimental Procedures

### 4.1 Participants

The sample was drawn from a cognitive follow-up study of women enrolled in the Melbourne Women's Midlife Health Project, a longitudinal population-based study of women's health across the menopausal transition (Dennerstein et al., 1993). Original entry criteria into the longitudinal study included menstruation during the prior 3 months, age 45 to 55, and no history of estrogen-containing HT. Recruitment began in 1991 and was based on random telephone dialing in metropolitan Melbourne. Of 2001 women interviewed, 438 women who fulfilled the selection criteria agreed to participate in the longitudinal study. Medical information was updated at annual study visits and included face-to-face interviews regarding menopausal status, medications, and operations, including hysterectomy. Prospective daily diaries provided information about menstrual cycle characteristics and reproductive phase. The retention rate over 11 years was 85%. The cognitive study involved 372 women aged 56–67 (mean age 60) invited in 2002 to participate in a phase I cognitive study (see (Clark et al., 2004) for details).

Of the 257 women in the phase I study, 65 women agreed to participate in an ancillary neuroimaging study aimed at understanding how HT affects cognitive function. Participants gave written consent in accordance with the University of Melbourne Human Ethics Committee and received 50AUD compensation. Of these, 51 women met the entry criteria for this study, based on prospective diaries and medical and medication records. Additional inclusion criteria included postmenopausal status at the time of cognitive testing, and exclusion criteria included current serious medical illness, surgically-induced menopause, current use of drugs that affect central nervous system function (including non-HT hormonal preparations), diabetes mellitus, history of head injury with loss of consciousness for greater than 1 hour, current Axis I psychiatric disorder, and evidence of depression (shortened version of the Center for Epidemiological Studies Depression Scale score >16)(Andresen et al., 1994) and alcohol abuse (Michigan Alcohol Screening Test score > 5). Additional exclusionary criteria for the imaging study included weight greater than 300 lb. (due to the



dimensions of the MRI scanner) and prohibited implanted metallic device. Of the 51 eligible women, 17 women had never used HT and 17 had initiated HT before the final menstrual period and continued using HT at the time of this assessment. Enrollment fell short of the goal of 20 women per group. Additional data collected from women with other histories of HT use were not analyzed because of small sample sizes (4 current and 5 past users who initiated HT at least one year after the final menstrual period, and 8 past users who initiated before the final menstrual period). The present analysis therefore involved 34 women, 17 early, continuous users and 17 never users. Among the 17 HT users, 10 used estradiol preparations, six used conjugated equine estrogen (CEE) preparations, one used tibolone; 13 of the 17 also received a form of progestogen (7 medroxyprogesterone acetate, 6 norethisterone). One never user did not complete the neuropsychological test battery, so neuropsychological tests were compared on 16 early users and 17 never users.

## 4.2 Procedures

Each participant underwent a 1-hour session of structural and functional magnetic resonance imaging (fMRI). Blood oxygen level-dependent (BOLD) imaging was performed on a 3 T GE Horizon LX Echospeed scanner at the Brain Research Institute at the Austin Hospital in Melbourne Australia. Twenty images were acquired through the cerebral hemispheres in an oblique axial plane parallel to the AC-PC plane. T1-weighted images were acquired for anatomic localization, and a gradient ECHO EPI sequence was employed for fMRI studies (flip angle of 60°, echo time of 40 ms, 64 by 64 matrix, 24-cm FOV, TR = 1.5 s, 20 slices, slice thickness = 3.5mm, skip= 0.5mm). An event-related design was used with a 10 s interstimulus interval for activation tasks. The activation tasks were separate verbal and figural recognition memory tests and were similar to the tasks used in previous PET studies (Maki and Resnick, 2000; Resnick et al., 1998). Each of the memory tests consisted of three separate tasks – encoding, yes/no recognition, and yes/no matching (control). Task order was: verbal encoding, figural encoding, 20-minute delay concurrent with structural image acquisition, verbal recognition, figural recognition, verbal matching and figural matching. Test items were abstract words and figures. In each of the two encoding tasks, participants viewed 20 targets individually and pressed the ‘yes’ button on a hand-held unit each time an item appeared. In each of the two recognition tasks, 40 items were shown individually, and participants were instructed to discriminate studied from unstudied items by responding ‘yes’ or ‘no’. In each of the two matching tasks, 40 pairs of items were shown, and participants were instructed to discriminate identical from non-identical pairs.

## 4.3 Neuroimaging analysis

We preprocessed and analyzed the data using Statistical Parametric Mapping (SPM99; Wellcome Department of Cognitive Neurology, London, UK). Functional MR EPI images were realigned to the first volume of each task to correct for interscan movement and co-registered to the individual’s T1 structural image. The structural image was spatially normalized to the SPM99 T1 MNI/ICBM (Montreal Neurologic Institute/International Consortium for Brain Mapping) template, and those parameters were applied to the spatial normalization of functional EPI images. A smoothing filter was applied at 12-mm (full width half maximum) Gaussian kernel. Events of interest for each task were those occurring from stimulus onset to just before the onset of the next stimulus. Individual analyses were performed convolving for the standard SPM canonical hemodynamic response function consisting of the sum of two gamma functions. Data were analyzed using a two-level mixed effects model (random effect analysis). The WFU Pickatlas (Maldjian et al., 2003) and the AAL template (Tzourio-Mazoyer et al., 2002) were used to define *a priori* regions of interest (ROIs), including the hippocampus and the parahippocampal gyrus. Linear contrasts of the parameter estimates for the effects of interest were obtained for each subject and then entered into a between-subjects analysis that produced a statistical parametric map of the *t*

statistic at every voxel within the ROIs. Results within these ROIs for the primary analysis of HT effects are presented with a threshold at  $P < 0.01$  (uncorrected) and a minimum cluster size of 10 contiguous voxels. In addition to between-subjects contrasts of each of the six conditions, two different subtractions were calculated for each memory task (i.e., verbal and figural recognition): encoding versus match and recognition versus match. We also carried out one-sample t-tests for each of the six conditions and each of the subtractions in the ROIs order to characterize the typical pattern of activation and deactivation on the tasks in medial temporal regions across the two groups combined. For this analysis, we used a more conservative threshold of  $P < .001$  and a minimum cluster size of 30 voxels. Anatomical localization of significant activations of interest was determined by converting MNI coordinates to standard Talairach coordinates using a non-linear transformation (<http://www.mrc-cbu.cam.ac.uk/Imaging/Common/downloads/MNI2tal/mni2tal.m>) and through use of standard atlases (Talairach and Tournoux, 1988) and the Talairach Daemon (Lancaster et al., 2000).

#### 4.4 Neuropsychological testing

Neuropsychological assessments were conducted by psychologists at the Office for Gender and Health at the Royal Melbourne Hospital between the hours of 8:30 am and 11:15 am and lasted for approximately 2 hours (see (Clark et al., 2004) for details on the cognitive study). Three standardized tests of verbal memory and one test of spatial memory were administered and examined in relation to HT including: 1) *modified California Verbal Learning Test -II (CVLT - II)* (Delis et al., 2000), a measure of verbal learning and memory involving 3 oral presentations of a 16-item word list comprised of 4 words from each of 4 semantic categories (max score = 48) and a 20- to 30-minute delayed recall trial (max = 16); 2) *East Boston Memory Test (EBMT)* (Scherr et al., 1988), a story recall test involving the oral presentation of a 3-sentence story and immediate (max = 12) and 3-minute delayed trials (max = 12); 3) *Unrelated Word List*, a 10-item test of memory for unrelated words that was originally used as part of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Welsh et al., 1994), was validated in an Australian sample (Collie et al., 1999), and was administered three years before this investigation (Henderson et al., 2003b). The task includes three learning trials, each varying in word order (max score = 30), and a 5-minute delayed trial (max = 10); 4) *Wechsler Memory Scale-III Faces subtest* (Corporation, 1997) a nonverbal memory task involving the presentation of 24 to-be-remembered faces, an immediate yes/no face recognition task involving the 24 studied faces and 24 distractors (max = 48), and a 30-minute delayed recognition task (max = 48). Two tests of verbal knowledge were also administered to evaluate baseline cognitive ability. One was the *modified Primary Mental Abilities (PMA) Vocabulary Test* (Thurstone and Thurstone, 1962), a timed, multiple-choice test in which participants are given 3 minutes to complete up to 50 items requiring them to choose the one word among four alternatives that is synonymous with the target word (max = 50). The other was the *New Adult Reading Test (NART)* (Nelson and O'Connell, 1978), a reading test requiring the pronunciation of a 61-item list of irregularly spelled words that increases in difficulty (max = 61).

#### Acknowledgments

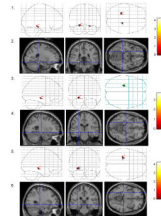
This work was supported by the Intramural Research Program of the National Institute on Aging, NIH and by an unrestricted grant from Wyeth Pharmaceuticals. P. Maki received research funding from Wyeth Pharmaceuticals for this investigator-initiated study. Lorraine Dennerstein is currently a consultant/advisory board member for Boehringer-Ingelheim and Bayer Schering Women's Healthcare. She has previously been a consultant/ advisory board member for Wyeth. M. Clark, J. Guthrie, V. Henderson, P. LaMontagne, D. Fornelli, D. Little, and S.M. Resnick have no conflicts of interest.

We also wish to acknowledge the contribution of the Austin Hospital Brain Research Unit.

## References

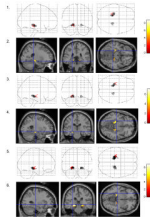
- Andresen EM, et al. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *Am J Prev Med* 1994;10:77–84. [PubMed: 8037935]
- Berent-Spillon A, et al. Early menopausal hormone use influences brain regions used for visual working memory. *Menopause* 17:692–699. [PubMed: 20300040]
- Brett, K.; Chong, Y. National Center for Health Statistics. Maryland: Hyattsville; 2001. *Hormone Replacement Therapy: Knowledge and Use in the United States*.
- Cabeza R, Nyberg L. Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci* 2000;12:1–47. [PubMed: 10769304]
- Clark MS, et al. Normative verbal and non-verbal memory test scores for Australian women aged 56–67. *Aust N Z J Psychiatry* 2004;38:532–540. [PubMed: 15255826]
- Collie A, et al. Norms and the effects of demographic variables on a neuropsychological battery for use in healthy ageing Australian populations. *Aust N Z J Psychiatry* 1999;33:568–575. [PubMed: 10483853]
- Corporation, TP. The WAIS-III-WMS-III technical manual. San Antonio, TX: Harcourt-Brace; 1997.
- Delis, D., et al. Adult Version Manual, second edition, vol. San Antonio, TX: Psychological Corporation, Harcourt-Brace; 2000. *California Verbal Learning Test*.
- den Tonkelaar I. Validity and reproducibility of self-reported age at menopause in women participating in the DOM-project. *Maturitas* 1997;27:117–123. [PubMed: 9255746]
- Dennerstein L, et al. Menopausal symptoms in Australian women. *Med J Aust* 1993;159:232–236. [PubMed: 8412889]
- Donaldson DI, et al. Dissociating state and item components of recognition memory using fMRI. *Neuroimage* 2001;13:129–142. [PubMed: 11133316]
- Erickson KI, et al. A cross-sectional study of hormone treatment and hippocampal volume in postmenopausal women: evidence for a limited window of opportunity. *Neuropsychology* 24:68–76. [PubMed: 20063947]
- Espeland MA, et al. Brain volumes, cognitive impairment, and conjugated equine estrogens. *J Gerontol A Biol Sci Med Sci* 2009;64:1243–1250. [PubMed: 19729392]
- Gibbs RB. Long-term treatment with estrogen and progesterone enhances acquisition of a spatial memory task by ovariectomized aged rats. *Neurobiol Aging* 2000;21:107–116. [PubMed: 10794855]
- Greendale GA, et al. Effects of the menopause transition and hormone use on cognitive performance in midlife women. *Neurology* 2009;72:1850–1857. [PubMed: 19470968]
- Henderson V, et al. Estrogen exposures and memory at midlife: A population-based study of women. *Neurology* 2003a;60:1369–1371. [PubMed: 12707448]
- Henderson VW. Aging, estrogens, and episodic memory in women. *Cogn Behav Neurol* 2009;22:205–214. [PubMed: 19996872]
- Henderson VW, et al. Estrogen exposures and memory at midlife: a population-based study of women. *Neurology* 2003b;60:1369–1371. [PubMed: 12707448]
- Joffe H, et al. 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;13:411–422. [PubMed: 16735938]
- Kang JH, et al. Postmenopausal hormone therapy and risk of cognitive decline in community-dwelling aging women. *Neurology* 2004;63:101–107. [PubMed: 15249618]
- Lancaster JL, et al. Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp* 2000;10:120–131. [PubMed: 10912591]
- Maki PM, et al. Objective hot flashes are negatively related to verbal memory performance in midlife women. *Menopause* 2008;15:848–856. [PubMed: 18562950]
- Maki PM, et al. Hormone therapy in menopausal women with cognitive complaints: a randomized, double-blind trial. *Neurology* 2007;69:1322–1330. [PubMed: 17893293]

- Maki PM, Resnick SM. Longitudinal effects of estrogen replacement therapy on PET cerebral blood flow and cognition. *Neurobiology of Aging* 2000;21:373–383. [PubMed: 10867223]
- Maki PM, Sundermann E. Hormone therapy and cognitive function. *Hum Reprod.* 2009 Update.
- Maldjian JA, et al. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 2003;19:1233–1239. [PubMed: 12880848]
- Nelson HE, O'Connell A. Dementia: the estimation of premorbid intelligence levels using the New Adult Reading Test. *Cortex* 1978;14:234–244. [PubMed: 679704]
- Persad CC, et al. Enhanced neuroactivation during verbal memory processing in postmenopausal women receiving short-term hormone therapy. *Fertil Steril* 2009;92:197–204. [PubMed: 18692790]
- Phillips SM, Sherwin BB. Effects of estrogen on memory function in surgically menopausal women. *Psychoneuroendocrinology* 1992;17:485–495. [PubMed: 1484915]
- Resnick SM, et al. The Women's Health Initiative Study of Cognitive Aging (WHISCA): a randomized clinical trial of the effects of hormone therapy on age-associated cognitive decline. *Clin Trials* 2004;1:440–450. [PubMed: 16279282]
- Resnick SM, et al. Postmenopausal hormone therapy and regional brain volumes: the WHIMS-MRI Study. *Neurology* 2009;72:135–142. [PubMed: 19139364]
- Resnick SM, Henderson VW. Hormone therapy and risk of Alzheimer disease: a critical time. *JAMA* 2002;288
- Resnick SM, et al. Estrogen effects on PET cerebral blood flow and neuropsychological performance. *Hormones and Behavior* 1998;34:171–184. [PubMed: 9799627]
- Resnick SM, et al. Effects of combination estrogen plus progestin hormone treatment on cognition and affect. *J Clin Endocrinol Metab* 2006;91:1802–1810. [PubMed: 16522699]
- Rossouw JE, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *Jama* 2002;288:321–333. [PubMed: 12117397]
- Schacter DL, Wagner AD. Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus* 1999;9:7–24. [PubMed: 10088896]
- Scherr PA, et al. Correlates of cognitive function in an elderly community population. *Am J Epidemiol* 1988;128:1084–1101. [PubMed: 3189282]
- Sherwin BB. Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. *Psychoneuroendocrinology* 1988;13:345–357. [PubMed: 3067252]
- Silva I, et al. Onset of estrogen replacement has a critical effect on synaptic density of CA1 hippocampus in ovariectomized adult rats. *Menopause* 2003;10:406–411. [PubMed: 14501601]
- Talairach, J.; Tournoux, P. Co-planar stereotaxic atlas of the human brain, vol. New York: Thieme Medical Publishers; 1988.
- Thurstone, L.; Thurstone, T. Primary mental abilities, vol. Chicago: Science Research Associates; 1962.
- Tzourio-Mazoyer N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 2002;15:273–289. [PubMed: 11771995]
- Welsh KA, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology* 1994;44:609–614. [PubMed: 8164812]
- West SL, et al. Recall accuracy for prescription medications: self-report compared with database information. *Am J Epidemiol* 1995;142:1103–1112. [PubMed: 7485055]



**Figure 1. Results of primary region of interest analysis in the medial temporal lobe: increased hippocampal activation and decreased parahippocampal activation in perimenopausal initiators of hormone therapy compared to never users**

Sagittal, coronal, and axial projections of medial temporal regions showing significant ( $p < .01$ ) differences in activation between perimenopausal users of hormone therapy and nonusers. Odd numbered rows show glass brain images of significant findings, and even numbered rows show activations superimposed on a structural anatomical template, with most significant voxel depicted in the cross-hairs. The scale to the right of each pair of figures shows the color scale corresponding to the z-values for that particular analysis. Rows 1 and 2 show significantly decreased activation in perimenopausal estrogen users compared to nonusers in bilateral parahippocampal gyrus. Rows 3 and 4 show significantly increased activation in perimenopausal users compared to never users in the left hippocampus during verbal recognition, and Rows 4 and 5 show a similar effect during verbal match.



**Figure 2. Characteristic activations in the medial temporal lobe during performance of verbal tasks**

Sagittal, coronal, and axial projections of medial temporal regions showing significant ( $p < .01$ ) increases in activation across the two groups combined. There were no significant decreases in activation. Odd numbered rows show glass brain images of significant findings, and even numbered rows show activations superimposed on a structural anatomical template. For ease of presentation and reference to HT effects, the cross-hairs reference the regions showing HT effects in the group comparisons, as depicted in Figure 2. The scale to the right of each pair of figures shows the color scale corresponding to the z-values for that particular analysis. Rows 1 and 2 show significantly increased activation in the bilateral parahippocampal gyrus during verbal recognition across the two groups combined (in contrast to the decreased activation seen in perimenopausal HT users in Figure 1). Rows 3 and 4 show significantly increased activation across the two groups combined in the left hippocampus during verbal recognition (in comparison to the increased activation seen in perimenopausal HT users in Figure 1). Rows 4 and 5 show a similar effect during verbal match.

Demographic characteristics and performance on fMRI and neuropsychological tests for current, early initiators of hormone therapy and never users.

**Table 1**

	<u>Early Users</u> (n = 17)	<u>Never Users</u> (n = 17)		<u>t-test</u> t(31)	<u>Effect</u> <u>Size</u>
<b>Demographics</b>					
Age, y	60.16 (2.88)	60.04 (2.74)			
Education, % > 12 y	52.9	65			
<b>Verbal IQ</b>					
NART	36.88 (4.25)	38.19 (6.32)			
PMA	38.39 (14.07)	39.68 (11.18)			
<b>Mood</b>					
CES-D	6.20 (3.68)	6.35 (3.87)			
			<u>Group</u> <u>Difference</u>	<u>t-test</u> t(31)	<u>Effect</u> <u>Size</u>
<b>Verbal Memory</b>					
CVLTII - Immediate	27.76 (5.13)	26.94 (4.84)	0.82 (4.99)	0.48	0.16
CVLTII - Delay	8.94 (1.48)	9.44 (3.58)	-0.50 (2.53)	-0.53	0.20
CERAD - Immediate	21.65 (3.60)	21.81 (2.76)	-0.16 (3.18)	-0.15	0.05
CERAD - Delay	5.65 (2.26)	6.00 (2.37)	-0.35 (2.32)	0.44	0.15
EBMT-Immediate#	10.76 (1.39)	9.81 (1.68)	0.95 (1.54)	1.78	0.62
EBMT -Delay#	10.76 (1.64)	9.75 (1.77)	1.01 (1.71)	1.71	0.59
<b>Face Memory</b>					
Immediate	34.88 (4.48)	34.56 (4.53)	0.32 (4.51)	0.20	0.07
Delayed	37.35 (3.90)	36.75 (3.92)	0.60 (3.91)	0.44	0.15
			<u>Group</u> <u>Difference</u>	<u>t-test</u> t(23)	<u>Effect</u> <u>Size</u>
<b>fMRI Task</b>					
Verbal Recognition *a	76.15 (10.97)	67.92 (6.97)	10.19 (4.06)	2.22	0.91
Figural Recognition	75.96 (7.40)	76.46 (6.17)	-0.59 (6.48)	-0.28	0.06

Notes: Means and standard deviations unless otherwise indicated;

\* p < .05 for t-test of group differences in outcome measures for full sample of 33;

# p < .10, for full sample of 33, two-tail;

<sup>a</sup> $p < .05$  for percent correct on imaging memory tasks in sample of 25 with complete neuroimaging data, two-tail; NART = New Adult Reading Test; PMA = Primary Mental Abilities Vocabulary Test; CES-D = Center for Epidemiological Studies Depression Scale; CVLT-II = California Verbal Learning Test; EBMT = East Boston Memory Test; effect sizes calculated with Cohen's  $d$ . One never-user completed only the in-scanner memory tests, so the sample size for the out of scanner memory tests was 33.



**Table 2**  
Significant Differences in Medial Temporal Brain Regions Between Early Users of Hormone Therapy and Never Users.

Condition, Direction of Effect	Talairach Coordinate			Brain Region	Z
	x	y	z		
Verbal Recognition					
Never>Early	-22	-35	-10	Left PHG, perirhinal, BA 36	3.48
Early>Never	20	-30	-14	Right PHG, perirhinal, BA 35	3.24
Verbal Encoding	-26	-22	-7	Left Hippocampus	3.28
Verbal Match					
Never>Early				none	
Early>Never				none	
Verbal Encoding - Match					
Never>Early	-30	-27	-4	Left Hippocampus	3.67
Early>Never					
Verbal Recognition - Match					
Never>Early	-26	-24	-6	Left Hippocampus	3.18
Early>Never				none	
Figural Recognition					
Never>Early				none	
Early>Never				none	
Figural Encoding					
Never>Early				none	
Early>Never				none	
Figural Encoding - Match					
Never>Early	-16	-31	-3	Left PHG, perirhinal, BA 27/35	3.14
Early>Never					
Figural Recognition					
Never>Early				none	
Early>Never				none	
Figural Encoding - Match					
Never>Early				none	

Talairach Coordinate				
Condition, Direction of Effect	x	y	z	Brain Region
Early>Never				none
Figural Recognition - Match				
Never>Early				none
Early>Never	20	-33	-7	Right PHG, BA 35
				3.02

Notes: Significant results presented for verbal and figural memory tasks for each condition alone and each condition contrasted with the other. Direction of effects indicates whether early HT users showed greater (>) or less (<) activation compared to never users. BA = Brodmann Area; PHG = parahippocampal gyrus; none = not available (no significant differences).

Significant Activations in Medial Temporal Brain Regions in the Combined Sample of Early Users and Never Users.

**Table 3**

Condition, Direction of Effect	x	y	z	Brain Region	Z score
Verbal Encoding				None	
Verbal Recognition					
	<b>-14</b>	<b>-28</b>	<b>-7</b>	<b>Left PHG, BA 28/35</b>	<b>5.29</b>
	-24	-24	-11	Left PHG, BA 28	4.45
	<b>16</b>	<b>-27</b>	<b>-5</b>	<b>Right PHG, BA 28/35</b>	<b>4.47</b>
	26	-30	-10	Right PHG, BA 35/36	3.29
Verbal Match					
	<b>-16</b>	<b>-26</b>	<b>-5</b>	<b>Left PHG, BA 28, L Hipp</b>	<b>5.48</b>
	<b>16</b>	<b>-26</b>	<b>-7</b>	<b>Right PHG, BA 28, R Hipp</b>	<b>4.66</b>
Figural Encoding					
				None	
Figural Recognition					
				None	
Figural Match					
	<b>-26</b>	<b>-28</b>	<b>-9</b>	<b>Left PHG, BA 35, Left Hipp</b>	<b>4.83</b>
	-18	-27	-4	Left PHG, BA 28, 35	4.45
	<b>24</b>	<b>-26</b>	<b>-5</b>	<b>Left PHG, BA 28</b>	<b>4.11</b>

Notes: Significant activations across the two groups combined are shown for verbal and figural memory tasks for each condition alone and each condition contrasted with the others. No significant deactivations were observed. Bolded activations refer to primary clusters and unbolded text refers to subclusters. BA = Brodmann Area; PHG = parahippocampal gyrus; none = no significant differences.