Using predictors of hormone therapy use to model the healthy user bias: How does healthy user status influence cognitive effects of hormone therapy?

Carey E. Gleason, PhD, MS[1,2,3], **N. Maritza Dowling, PhD**[3,4], **Elliot Friedman, PhD**[5], **Whitney Wharton, PhD**[1,2,3], and **Sanjay Asthana, MD, FRCP(C)**[1,2,3]

[1]Division of Geriatrics and Gerontology, Department of Medicine, University of Wisconsin-Madison, School of Medicine and Public Health, Madison, WI

[2]Geriatric Research, Education, and Clinical Center, William S. Middleton Memorial Veterans Hospital, Madison, WI

[3]Wisconsin Alzheimer's Disease Research Center, University of Wisconsin-Madison, School of Medicine and Public Health, Madison, WI

[4]Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, WI

[5]Institute on Aging, University of Wisconsin-Madison, Madison, WI

Abstract

Objective—This study investigated the phenomenon known as the healthy user bias, by equating hormone therapy (HT) use (past or current) with healthy user status.

Methods—Data from the Survey of Midlife in the United States (MIDUS) were used to indentify predictors of HT use. The unique MIDUS data include psychological, demographic, health-related, and behavioral variables as well as history of HT use. Predictors of HT use were combined to derive propensity scores, describing the likelihood that a woman was an HT user, based on her psychological, demographic, physical, and behavioral profile (i.e., likelihood of being a healthy user) as opposed to her actual use of HT. Finally, cognitive performance on an executive function test was examined in women stratified by propensity score.

Results—Using a multiple logistic regression model, nine variables emerged as predictors of HT use. The nine variables were used to estimate the propensity or conditional probability of using HT for each subject; resultant propensity scores were ranked and divided into tertiles. Women in the highest tertile demonstrated shorter median response latencies on a test of executive function than women who did not use HT.

Conclusions—From an array of psychological, medical and behavioral variables, nine emerged as predictors of HT use. If validated, these features may serve as a means to estimate the phenomenon known as the healthy user bias. Moreover, these data suggest that the degree to which a woman fits a model of a healthy user, may influence cognitive response to HT.

CORRESPONDING AUTHOR (reprints will be available): Carey E. Gleason, PhD, MS Address: Madison VA GRECC (D4211) 2500 Overlook Terrace Madison, WI 53705 Phone: 608-280-7000 Fax: 608-280-7291 ceg@medicine.wisc.edu.

Keywords

Hormone Therapy; Healthy User; Cognition; Menopause; Selection Bias; Executive Functioning

INTRODUCTION

Prior to the publication of findings from two Women's Health Initiative (WHI) ancillary studies, the WHI Memory Study[1–4] (WHIMS) and the WHI Study of Cognitive Aging[5] (WHISCA), it was widely believed that hormone therapy (HT) use could reduce a woman's risk for dementia by 50%.[6–8] This belief was predicated in large part on observational data from longitudinal case-control and cohort studies. In contrast, data from the large, randomized, controlled WHIMS and WHISCA trials revealed adverse cognitive effects and increased risk of dementia for women age 65 and older who were administered a widely-used form of oral HT, conjugated equine estrogen (CEE). The phenomenon referred to as the healthy-user bias was frequently cited as an explanation for the dramatic reversal in findings (e.g., Barrett-Conner et al.,[9] Prentice et al.[10]). Specifically, the healthy-user bias purports that women who use HT during the menopausal transition are systematically different (i.e., healthier) than women who do not use HT; moreover, it is the fundamental differences between HT-users and non-users, and not the exposure to HT, that results in HT's seemingly beneficial health effects, such as fewer cardiovascular risk factors and a reduced risk for dementia found in observational studies.

While the healthy user bias is thought to be a serious threat to study validity, contributing to past misconceptions about HT, it is only loosely and inconsistently operationalized as a research construct. The difficulty in defining the healthy user bias may lie in the fact that sources of bias are multifactorial. For example, behavioral features, such as exercising, not smoking, and regularly consulting with a medical provider; and psychological features such as optimism, sense of control, and proactive behaviors are all alleged to be characteristics of a "healthy user". Another source of selection bias pertaining to healthy users are differences in access to HT, such that only healthy women who regularly visit a physician would be prescribed HT; whereas women in poor health, those already on several medications, or those with a shortened life expectancy may be less likely to be prescribed a potentially protective medication. These features along with the demographic characteristics of being well-educated and in higher socioeconomic strata culminate in a multifaceted picture of the healthy user, who is allegedly more engaged, physically fit, articulate, adherent to medical advice, and psychologically well than her non-user counterpart.

Data from the Survey of Midlife in the United States (MIDUS) include psychological, behavioral and demographic variables, as well as data on HT use from a large population-based sample. In combination, the data provide a unique opportunity to examine the multiple components theoretically associated with the healthy user construct. Using this comprehensive dataset, we characterized the healthy user bias as it applies to women opting to use HT. In other words, the use of HT was equated with the healthy user phenomenon. In doing so, we attempted to define the demographic, personality, and behavioral characteristics of the healthy-user, by optimally identifying relevant predictors of HT use from a large pool of potential predictors. Variables predicting HT use i.e., healthy user status were used in a subsequent analysis to develop a propensity score. The propensity score would then represent the likelihood of being a healthy user, regardless of whether a women was actually using HT or not. We examined the cognitive performance of women after stratifying for healthy-user status. Overall, we hypothesized that the characteristics indentified as predictors of HT use would coincide with the characteristics traditionally portrayed as features of healthy users. Moreover, we predicted that even after accounting for

the healthier user status, cognitive effects related to use versus nonuse of HT would be evident.

METHODS

Participants

Data were obtained from the second wave of MIDUS data collection (MIDUS 2). Participants in this large population-based survey with metropolitan over-samples of non-institutionalized are English-speaking adults, now aged 35 to 86. The intent of the MIDUS project is to investigate patterns, predictors, and consequences of midlife development using a comprehensive array of physical, psychological, and social indicators. The second wave of data collection, MIDUS 2, occurred between January 2004 and January 2006 and included a telephone interview, two self-administered questionnaires mailed, and cognitive testing by telephone. Details of methods employed to ensure maximum follow-up are provided in an online manual.[11] From the original MIDUS sample (N=7,108), 4,963 participants completed the telephone interview, resulting in a longitudinal retention rate of 70% (75% response rate when adjusted for mortality). Of telephone respondents, 81% (N=4,032), returned the two self-administered questionnaires and 85% (N=4205) completed the 45 minute telephone cognitive testing. Altogether, a variety of data were collected, including information on demographics, psychological and physical health status, cognition, lifestyle choices, attitudes, and health-related behaviors. Further information on the MIDUS 2 sample and study methods are available at study's data website.[12]

Groups: Report of Menopausal Status and HT Exposure

The study analyses presented here included only postmenopausal women with complete menopausal status and HT exposure data. While the MIDUS 2 study was not specifically designed to assess women's history of HT use, responses on self-administered questionnaires were used to determine women's menopausal status and exposure to HT. If a respondent indicated that she had stopped menstruating in the last year, she was asked to select a reason why from a list of choices (e.g., menopause, hysterectomy, etc.) and to provide the approximate year of her last menstrual period. To clarify perimenopausal status, women were asked about the regularity of their menstrual periods and if they experienced menopausal symptoms. Women next reported their current or past use of HT. If they had used HT, they indicated their reasons for taking the medication, using a list of options, and provided the dates of initiation and (if applicable) discontinuation of the medications. These data were systematically interrogated to ensure that women were accurately categorized, such that any inconsistencies in the participants' reports were identified and clarified through a review of their responses in total.

Based on their responses, postmenopausal women were categorized into three main groups: Never Users (n=782), Past Users (n=506), or Current Users (n=315). Figure 1 describes the subject selection process. The exclusion of study participants with missing data (described in Figure 1) resulted in a final analytical sample size of 1,189 for the multivariate logistic analyses.

Measures

Variables were selected for possible inclusion in the model predicting HT use if previous published analyses 1) included a variable as a confounder or covariate in analyses (e.g., BMI, smoking, vitamin use, diabetes;[13]), or 2) revealed the variable to be predictive of HT use (e.g., level of exercise,[14] socioeconomic status[15]). Additionally, some variables were included based on a well-justified but hypothesized relationship to the healthy user phenomenon (e.g. psychological well being). We intended the first stage of variable

selection to be highly inclusive. All variables considered for inclusion are listed below, and described in detail in Supplementary Table A (**See Table, Supplemental Digital Content 1**, http://links.lww.com/MENO/A14). Of note, for the present study, our primary interest was to identify variables predicting HT use, not to obtain estimates of the population values for these variables. Therefore, our analyses did not use case weights to correct for differential selection probabilities and non-response.

Demographic and Anthropometric Variables—Demographic data included age, marital status, highest level of education completed, self-identify ethnicity (Hispanic v. Non-Hispanic), and race. For analyses, White and Asian participants were combined and compared to Black/African American, Native American and self-described "Other" racial category. Socioeconomic status was estimated using a log transformed reported household annual income (log income). Anthropometric data included Body Mass Index (BMI) and waist-to-hip ratios. Further details regarding these variables are provided in Supplementary Table A (**See Table, Supplemental Digital Content 1**, http://links.lww.com/MENO/A14).

Psychological and Personality Variables—A range of variables measuring psychological and personality characteristics, either empirically or theoretically linked to the healthy user bias were selected for inclusion in the model predicting HT use. Each is described in Supplementary Table A (**See Table, Supplemental Digital Content 1**, http://links.lww.com/MENO/A14). Briefly, these variables evaluated a participant's perceived control over her health (health locus of control; HLOC-Self and HLOC-Others);[16] her sense of agency or perceived control;[17] self-described level of conscientiousness;[17] optimism;[18] and self esteem.[19] Additionally, six variables from the Psychological Well-Being (PWB) Scale[20] were included: Autonomy, Environmental Mastery, Personal Growth, Positive Relations, Purpose in Life, and Self Acceptance. While there is correlation between the six PWB scales, factor analyses demonstrate that they measure distinct constructs.[20, 21] Moreover, the scales differentially associate with biological processes.[22]

Mood, Level of Satisfaction and Self-Evaluation of Health—Mood and outlook characteristics of participants were considered in the model. Described in Supplementary Table A (**See Table, Supplemental Digital Content 1**, http://links.lww.com/MENO/A14), specific variables included positive and negative affect subscales from the Positive and Negative Affect Schedule[23] (PANAS); depression and anxiety scores from the Composite International Diagnostic Interview – Short Form;[24] a rating of overall life satisfaction;[25] and a single item assessing self-evaluated physical health that has been shown to be a robust predictor of health outcomes.[26]

Health Related Behaviors—Several variables assessing health-related behaviors were selected to evaluate the degree to which women engaged in health-maintenance activities. These included a participant's use of vitamins and supplements; her access to, or utilization of routine medical, dental and optical care; her physical and mental activity; and her use of cigarettes and alcohol. Methods describing how each variable was derived are provided in Supplementary Table A (**See Table, Supplemental Digital Content 1**, http://links.lww.com/MENO/A14).

Physical Health and Medical History—Indicators of current physical and past medical history were investigated as possible predictors of HT use, i.e., the healthy use bias. Details on how these were derived are provided in Supplementary Table A (**See Table, Supplemental Digital Content 1**, http://links.lww.com/MENO/A14). Specifically, women's answers to questions about two common chronic conditions (high blood pressure and diabetes) were included in the model; along with number hospitalizations in the last year;

number of medications; an estimate of physical disability due to health; and past history of stroke, heart attack, head injury or cancer.

Cognitive Variables—The Brief Test of Adult Cognition by Telephone (BTACT) was administered to MIDUS 2 participants, having been found to demonstrate sound psychometric properties and to offer data equivalent to those obtained using face-to-face data collection.[27] The battery consists of tests from five cognitive domains sensitive to age and illness-associated changes in cognition, including measures of verbal memory, working memory, executive function, reasoning, and processing speed. While combined summary scores were available, we selected the only speeded executive function task to examine the potential effects of HT on cognition. As reported previously, evidence suggests that the cognitive effects of HT are not global, but rather domain specific.[28]

Thus, the most challenging and only timed task administered was selected for analyses. The Stop and Go Switch Task (SGST) assesses the ability of participants to switch between two conditional tasks. Under the Normal response condition, participants were to say "go" in response to an examiner reading the word "green," and "stop" in response to the word "red." Conversely, under the Reverse condition, they were instructed to provide the opposite response: say "stop" to the word "green" and "go" to the word "red". Two single-task sets were administered initially; the first being a Normal response condition, the second a Reverse response condition. A third set was administered with the conditions switching between Normal and Reverse at random intervals, referred to as a Mixed-Task set.

Four variables from the SGST test were selected for analyses: 1) median reaction time under the Normal condition (Normal RT), 2) median reaction time under the Reverse condition (Reverse RT), 3) median reaction time for Mixed-Tasks items wherein a condition switch occurred (Normal to Reverse or Reverse to Normal; referred to here as Mixed-Switch RT), and 4) median reaction time from the Mixed-Task trial wherein a condition switch did not occur (Mixed-Nonswitch RT). For each of the last two variables, Mixed-Switch RT and Mixed-Nonswitch RT, additional variables characterizing aspects of performance within each task were available for follow-up analyses. These included: 1) the median reaction time from Normal and Reverse condition trials occurring immediately after a condition switch occurred (Mixed-Switch Normal Switch RT and Mixed-Switch Reverse Switch RT), and 2) median reaction times from Normal and Reverse condition occurring without a switch in conditions (Mixed-Switch Normal Nonswitch RT and Mixed-Switch Reverse Nonswitch RT).

Statistical Methods

Based on previous publications investigating the association between HT use and cognitive outcomes, forty-eight variables were selected as possible predictors of HT use, i.e., potential qualities of a healthy-user. As a first step in reducing the number of variables included in analyses (**Step 1**), each of the 48 potential predictors was examined individually for its univariate association with HT use (Never v. Past or Current Use). Chi-squared tests were used to determine the statistical significance of the association between categorical variables and HT use. Univariate logistic regression models predicting HT use were selected to examine the statistical association between continuous predictors and HT.

In the next stage of analysis, the P-values associated with each univariate test were ranked and variables with a significance level of $P \leq 0.10$ were selected for inclusion in multivariate logistic regression models to predict HT use (**Step 2**), the goal being to reduce the number of variables without overly limiting the number of potential predictors. Given the potential effect of multicollinearity on parameter estimation, we examined the strength of the associations among all pairs of candidate predictors using Pearson's product-moment

correlations. The magnitude of the correlations ranged from 0.002 to 0.764. To minimize multicollinearity effects and efficiently manage the large number of predictors in the model, we used SAS PROC SCADGLIM, Version 1.1 (http://methodology.psu.edu) as a variable selection method for the multiple logistic (or generalized-linear) model.[29] This automated variable selection procedure employs a penalized likelihood approach known as smoothly clipped absolute deviation (SCAD) estimator to adjust the likelihood function. The SCAD penalty is the preferred selection process in the presence of collinearity and highly dimensional data compared to other penalized regression approaches such as the least absolute shrinkage and selection operator (LASSO) and ridge regression.[30] The Bayesian Information Criterion (BIC) was used as a criterion for final model selection.[31] We also examined model sensitivity, specificity, and receiver operating characteristic (ROC) curves to interpret results.

Finally, using the prediction model derived in Step 2, we estimated the propensity or conditional probability of using HT for each subject given a vector of relevant covariates. The resultant propensity scores (i.e., probabilities were ranging from 0 to 1) were ranked and sub-divided into tertiles. In an exploratory analysis, the cognitive performance for women in each of the tertiles was examined by regressing cognitive scores on HT use status while covarying for age and education level (**Step 3**).

RESULTS

Of the 1603 postmenopausal or perimenopausal women included in Step 1 analysis, 1023 (63.8%) had undergone or were currently experiencing natural menopause. Conversely, 740 (46.2%) women reported undergoing surgical menopause. The mean age of the women included in the Step 1 analysis was 60.7 (SD = 10.3) years. All but 8.7% of the participants completed at least a High School education. In general, participants exhibited low levels of depression and anxiety symptoms, rated their health as good to excellent and viewed themselves as conscientious and optimistic. Table 1 provides a summary of the samples characteristics in HT Users, including the groups Past and Current Users, and Never users.

Univariate analyses conducted with 48 candidate variables suggested that 26 characteristics of the sample were predictive of HT use when examined alone (**Step 1**). Table 1 summarizes each of the 26 variables, as well as the age and educational status of participants, including variable sample sizes, prevalence (for categorical variables), means (for continuous variables), and the P-values obtained in the univariate tests. (Supplementary Table B, **Supplemental Digital Content 2**, http://links.lww.com/MENO/A15 lists the same information for the 22 variables not selected as predictors of HT use, i.e., P-value of univariate analyses > 0.10). Notably, the age of Users (Current and Past) differed from Nonusers. While the difference was statistically significant, this variable was not included in predictor models, as it was thought to reflect a cohort difference. Specifically, older women were more likely to be Past Users, as opposed to Current or Never Users, probably because women well-past menopause were encouraged to discontinue HT after the WHI findings were published.

When the 26 candidate predictors were entered into a multiple logistic regression model predicting HT use, employing the SCAD variable selection approach, the best model (BIC=1262) produced nine statically significant predictors of group membership (Never User v. Past or Current User). Table 2 lists the variables surviving Step 2 analysis. Overall the Area Under the Curve (AUC) obtained when the nine variables listed in Table 2 were used to predict HT Use (Never v. Past or Current) was 0.73.

Finally, cognitive variables measuring speeded executive function were examined in combination with HT propensity scores. Women's likelihood of using HT was characterized as either a low, medium or high propensity to be an HT user i.e., a healthy user. Women in the low and middle ranges demonstrated no significant difference in their performance on the SGST variables. However, women in the highest tertile (i.e., women embodying the healthy user characteristics) demonstrated shorter median response latencies on the SGST Mixed-Switch trials, than women who did not use HT. Table 3 lists the median response times and p-values for cognitive task comparison between HT users and never users in the highest stratum of propensity scores. A follow up comparison of the type of switch task that occurred, revealed that past or current users of HT were faster than never users on tasks requiring them to switch from the Reverse condition to the Normal condition. Their performance when switching from the Normal condition to the Reverse condition was not different based on HT use.

DISCUSSION

From a list of candidate variables, describing a variety of psychological, behavioral, mood, and health status characteristics, nine participant features emerged as predictors of HT use. Specifically, HT users were more likely to be White or Asian; had had higher incomes; had lower waist-to-hip ratios, received regular physical exams (1–2 per year), were likely to use calcium supplements, and were less likely to have diabetes than women who never used HT. HT users reported slightly more use of prescription medications than non-users (mean of 2 to 3 medications compared to 1.7 in non-users). Additionally, HT users in our study exhibited higher self esteem and reported a small but significantly lower level of autonomy than non-users.

HT use, more than the use of other medications, has been proposed as a marker for the phenotype referred to as the healthy-user (e.g., Barrett-Conner et al.,[9] Prentice et al.[10]). For this reason, being an HT user was equated in this study with being a healthy user. These analyses were motivated by concerns that any association between HT use and cognitive function in postmenopausal women may be artifacts of the healthy user bias (i.e., HT users are reported to be psychologically and physically healthier, to have better access to preventative care, are more likely to engage in healthy behaviors than nonusers); as such, the superior cognitive performance of HT users was attributed to selection bias rather than direct effects of HT use. In contrast, our results suggest that women possessing the characteristics of a "healthy user" who used HT demonstrated an advantage on an executive functioning task over similar women who never used HT. If confirmed, these data suggest that HT confers beneficial cognitive effects, but only for women in optimal health.

A number of previous studies have examined how HT users differ from non-users. For example, women who used HT in the Nurse's Health Study (NHS) were at lower risk for cardiovascular events (i.e, myocardial infarction and death due to coronary heart disease). However, they were also thinner, ate healthier diets, engaged in more exercise, and used less alcohol than women who did not use HT.[32] Other studies specifically examined how HT users and non-users differed, consistently reporting that women prescribed HT at menopause were less obese, more physically active, and better educated, and in contrast to the NHS findings, were more likely to use alcohol.[33, 34] Also influencing the likelihood of being an HT user was age, type of menopause (surgical v. natural) and the diagnosis of osteoporosis, with users being younger, more likely to have experienced surgical menopause and to carry a diagnosis of osteoporosis.[33] Similar relationships between HT use, and demographic and behavioral characteristics emerge when European cohorts were examined. While less commonly prescribed in European countries, HT use was positively associated with education, socioeconomic status, nulliparity or lower birth rates, non-immigrant status,

lower Body Mass Index, better diet, regular exercise, surgical menopause and osteoporosis.[35–37] Interestingly, the characteristics of user and non-users did not appear to change markedly following the publication of data suggesting HT was potentially harmful for cardiovascular health.[38]

The emerging story that women who were prescribed HT were systematically different than never users, led Matthews et al.[39] to ask if users were indeed healthier than non-users. In other words, a selection bias occurs in HT prescription practices, such that women using HT fit a profile of a healthy user. Unlike other studies describing the HT user, Matthews et al. were able to include both health and psychological factors in their investigation, and found that HT users exhibited healthier cardiovascular risk profiles prior to HT use, described themselves as possessing Type A personality features, and were more likely to be aware of and to express their emotions, motives, and symptoms of stress.

While a number of studies examined characteristics of HT uses, none had access to the wide array of psychological, personality, demographic, and health status variables available in the MIDUS data set. Given the richness of the MIDUS data we may be able to more fully characterize the healthy user bias by examining a greater variety of potential characteristics of HT users than past studies examining a more limited set of predictors.

Overall, the findings of the present analyses portray a similar profile of HT users as that of Matthews et al.[39] For example, compared to non-users, women in the MIDUS study who used HT had higher incomes, were more likely White or Asian, had lower waist-to-hip ratios, and reported higher self esteem. Consistent with them being healthy, HT users in our study received regular physical exams (1–2 per year), were likely to use calcium supplements, and were less likely to have diabetes than women who never used HT. Perhaps because of their regular medical care, HT users reported more use of prescription medications than non-users. Another finding contrasted earlier reports; specifically, HT users in our study reported a small but significantly lower level of autonomy than non-users. These data were collected shortly after the publication of the WHI findings;[40, 41] and the time periods of HT use or non-use reported by respondents straddle the dates of the WHI publications. The finding that women prescribed HT describe themselves as less autonomous than non-users may reflect a reaction to shifts in public perception of HT, such that women describing themselves as autonomous addressed their menopausal transition by opting not to use HT.

Another aspect of the current study involved using the nine variables predictive of HT use to develop healthy user propensity scores for each subject. In doing so, we were able to characterize the propensity of individual women to be HT users or healthy users, regardless of whether they used or did not use HT. Per our theory, women with high propensity scores possess more of the qualities of a healthy user than do women with low propensity scores. An exploratory analysis, comparing cognitive performance on a demanding executive function task suggested that only women possessing more qualities of the healthier user may demonstrate cognitive differences based on their HT exposure. Specifically, the difference in performance was noted on the Stop and Go Switch Task. Women were instructed that they were to provide either concordant (normal) or discordant (reverse) responses to the stimuli "Stop" and "Go," saying either "red" to stop or "green" to stop depending on the instructions (normal v. reverse). In sets that required women to switch between the normal and reverse conditions (mixed-task), women who were past or current users of HT were responded faster than non users when a switch occurred. The advantage in speed appeared to occur when the switch involved moving from the more difficult condition (reverse switch) or to the easier condition (normal switch).

The current study's limitations need to be acknowledged. First it may be inaccurate to equate the healthy user with an HT user. While related, the two concepts likely diverge in some regards. For example, HT use would describe only female healthy users; and it may not apply to non-US and non-European populations or to selected sub-groups within these primarily White, affluent cultures, where medical care is relatively easily accessible. Importantly, this cross-sectional analysis does not imply a causal relationship between HT use and healthy user variables, only an association. Another limitation was the large number of participants who were missing data for one or more of the twenty-six variables included in the Propensity models (Step 2 of analyses; See Figure 1). Over 400 women were excluded for this reason, most of whom were never-users. This would also limit the generalizability of our models, and may have reduced power to detect the relationships between the variables and HT use. There may be biases introduced by selection of only a subset of the MIDUS sample (i.e., women for whom menopausal status and HT use were available.) Finally, the analysis of cognitive data was considered exploratory and p-values were not corrected for multiple comparisons. This increases the risk for Type 1 error, such that the suggestion of cognitive benefits with HT may be inaccurate, reflecting spurious or random findings.

Still, the present study may help clarify the contradictory findings regarding HT use, cognitive function, and risk for dementia. As noted, it was once widely believed that hormone therapy (HT) use could reduce a woman's risk for dementia by up to 50%.[6–8] In sharp contrast, findings from WHIMS,[1–4] suggested cognitive harm and increased risk for dementia. To clarify the contradiction, some proposed a theoretical "critical window" during which HT needed to be applied in order to be protective (e.g. Resnick & Henderson;[42] Sherwin[43]). Specifically, in observational studies, most use of HT occurred at or shortly after menopause, as opposed to initiating use after a decade or more of no exposure to hormones (endogenous or exogenous), as was the case for many WHI participants. Per the critical window hypothesis, use outside the critical time frame around menopause, could have no effect or could in fact be harmful. The data from the present study suggest an additional explanation. Specifically, the beneficial effects of HT may be sufficiently small, such that illnesses like diabetes, and life long exposure to stress and poverty overwhelm any potential benefit of HT medications. On the other hand, for women in good health, whose socioeconomic status and psychological health are optimized, HT may confer some cognitive benefits.

There is less equivocation about the biological effects of endogenous hormones compared to exogenous HT. For example, cardiac event rates are low in premenopausal women; however, the rates rise dramatically after menopause.[44] Similarly, an earlier age of menopause is associated with increase risk for dementia in women with Down's syndrome.[45] The confusion regarding exogenous hormones administered as HTs highlights the pressing need to reconcile the contradictory science regarding the neurobiological and cognitive effects of menopausal HT. In the meantime, dramatic reversals followed by continued equivocation in scientific opinions have left women confused and mistrustful about any reports on the neurobiological effects of HT. The present study offers insight as to how the healthy-user bias might be defined, and how these intrinsic differences between HT users and non-users could influence the potential cognitive benefits of HT. Continued exploration of the healthy user bias and other key differences between observational and clinical trial data (e.g., when in relation to menopause did HT exposure occur) is needed in order to address and overcome the inconsistencies in findings.

CONCLUSIONS

The present study offers insights into the nature of the healthy user bias and how this influential factor may have contributed to past misconceptions about HT. By equating the

concepts of HT user and healthy user, this study offers a means to estimate a person's propensity to be a healthy user. These data suggest that the degree to which a person fits a model of healthier user status, may influence cognitive response to HT.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Shumaker SA, Legault C, Kuller L, et al. Conjugated equine estrogens and incidence of probable dementia and mild cognitive impairment in postmenopausal women: Women's Health Initiative Memory Study. Jama. Jun 23; 2004 291(24):2947–2958. [PubMed: 15213206]

2. Shumaker SA, Legault C, Rapp SR, et al. 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. Jama. May 28; 2003 289(20):2651–2662. [PubMed: 12771112]

3. Espeland MA, Rapp SR, Shumaker SA, et al. Conjugated equine estrogens and global cognitive function in postmenopausal women: Women's Health Initiative Memory Study. Jama. Jun 23; 2004 291(24):2959–2968. [PubMed: 15213207]

4. Rapp SR, Espeland MA, Shumaker SA, et al. Effect of estrogen plus progestin on global cognitive function in postmenopausal women: the Women's Health Initiative Memory Study: a randomized controlled trial. Jama. May 28; 2003 289(20):2663–2672. [PubMed: 12771113]

5. Resnick SM, Maki PM, Rapp SR, et al. Effects of combination estrogen plus progestin hormone treatment on cognition and affect. J Clin Endocrinol Metab. May; 2006 91(5):1802–1810. [PubMed: 16522699]

6. Miller MM, Monjan AA, Buckholtz NS. Estrogen replacement therapy for the potential treatment or prevention of Alzheimer's disease. Ann N Y Acad Sci. Dec.2001 949:223–234. [PubMed: 11795357]

7. Hogervorst E, Williams J, Budge M, Riedel W, Jolles J. The nature of the effect of female gonadal hormone replacement therapy on cognitive function in post-menopausal women: a meta-analysis. Neuroscience. 2000; 101(3):485–512. [PubMed: 11113299]

8. Fillit HM. The role of hormone replacement therapy in the prevention of Alzheimer disease. Arch Intern Med. Sep 23; 2002 162(17):1934–1942. [PubMed: 12230415]

9. Barrett-Connor E, Grady D, Stefanick ML. The rise and fall of menopausal hormone therapy. Annu Rev Public Health. 2005; 26:115–140. [PubMed: 15760283]

10. Prentice RL, Langer R, Stefanick ML, et al. Combined postmenopausal hormone therapy and cardiovascular disease: toward resolving the discrepancy between observational studies and the Women's Health Initiative clinical trial. Am J Epidemiol. Sep 1; 2005 162(5):404–414. [PubMed: 16033876]

11. Ryff, C.; Almeida, D.; Ayanian, J., et al. National Survey of Midlife. Development in the United States. (MIDUS II), 2004–2006. ICPSR 4652. Field Report for MIDUS 2 Longitudinal Sample. http://www.icpsr.umich.edu/icpsrweb/NACDA/studies/28683/detail

12. Midlife Development in the United States (MIDUS II), 2004–2006.
http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4652?archive=ICPSR&q=midus+2. Updated Last Updated Date

13. Kang JH, Weuve J, Grodstein F. Postmenopausal hormone therapy and risk of cognitive decline in community-dwelling aging women. Neurology. Jul 13; 2004 63(1):101–107. [PubMed: 15249618]

14. Isaacs AJ, Britton AR, McPherson K. Why do women doctors in the UK take hormone replacement therapy? J Epidemiol Community Health. Aug; 1997 51(4):373–377. [PubMed: 9328541]

15. Petitti DB, Freedman DA. Invited commentary: how far can epidemiologists get with statistical adjustment? Am J Epidemiol. Sep 1; 2005 162(5):415–418. discussion 419–420. [PubMed: 16014781]

16. Wallston BS, Wallston KA, Kaplan GD, Maides SA. Development and validation of the health locus of control (HLC) scale. J Consult Clin Psychol. Aug; 1976 44(4):580–585. [PubMed: 939841]

17. Lachman, ME.; Weaver, SL. The Midlife Development Inventory (MIDI) Personality Scales: Sacle construciton and score: Technical Report. 1997.

18. Scheier MF, Carver CS. Optimism, coping, and health: assessment and implications of generalized outcome expectancies. Health Psychol. 1985; 4(3):219–247. [PubMed: 4029106]

19. Rosenberg, M. Society and the adolescent self-image. Princeton University Press; Princeton, N.J.: 1965.

20. Ryff CD, Keyes CL. The structure of psychological well-being revisited. J Pers Soc Psychol. Oct; 1995 69(4):719–727. [PubMed: 7473027]

21. Ryff CD. In the eye of the beholder: views of psychological well-being among middle-aged and older adults. Psychol Aging. Jun; 1989 4(2):195–201. [PubMed: 2789747]

22. Ryff CD, Dienberg Love G, Urry HL, et al. Psychological well-being and ill-being: do they have distinct or mirrored biological correlates? Psychother Psychosom. 2006; 75(2):85–95. [PubMed: 16508343]

23. Mroczek DK, Kolarz CM. The effect of age on positive and negative affect: a developmental perspective on happiness. J Pers Soc Psychol. Nov; 1998 75(5):1333–1349. [PubMed: 9866191]

24. Kessler RC, Andrews G, Mroczek D, Ustun B, Wittchen HU. The World Health Organization Composite International Diagnostic Interview Short-Form (CIDI-SF). International Journal of Methods in Psychiatric Research. 1998; 7:171–185.

25. Prenda KM, Lachman ME. Planning for the future: a life management strategy for increasing control and life satisfaction in adulthood. Psychol Aging. Jun; 2001 16(2):206–216. [PubMed: 11405309]

26. Myint PK, Luben RN, Surtees PG, et al. Relation between self-reported physical functional health and chronic disease mortality in men and women in the European Prospective Investigation into Cancer (EPIC-Norfolk): a prospective population study. Ann Epidemiol. Jun; 2006 16(6):492–500. [PubMed: 16005244]

27. Tun PA, Lachman ME. Telephone assessment of cognitive function in adulthood: the Brief Test of Adult Cognition by Telephone. Age Ageing. Nov; 2006 35(6):629–632. [PubMed: 16943264]

28. Asthana S. Estrogen and cognition: a true relationship. JAGS. 2004; 52:316–318.

29. Dziak, J.; Lemmon, D.; Li, R.; Huang, L. PROC SCADGLIM User's Guide (Version 1.1). The Methodology Center, Penn State; University Park: 2010.

30. Fan J, Li R. Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association. 2001; 96(456):1348–1360.

31. Schwarz G. Estimating the dimension of a model. Annals of Statistics. 1978; 6:461–464.

32. Stampfer MJ, Willett WC, Colditz GA, Rosner B, Speizer FE, Hennekens CH. A prospective study of postmenopausal estrogen therapy and coronary heart disease. N Engl J Med. Oct 24; 1985 313(17):1044–1049. [PubMed: 4047106]

33. Cauley JA, Cummings SR, Black DM, Mascioli SR, Seeley DG. Prevalence and determinants of estrogen replacement therapy in elderly women. Am J Obstet Gynecol. Nov; 1990 163(5 Pt 1):1438–1444. [PubMed: 2240084]

34. Egeland GM, Kuller LH, Matthews KA, Kelsey SF, Cauley J, Guzick D. Premenopausal determinants of menopausal estrogen use. Prev Med. May; 1991 20(3):343–349. [PubMed: 1862056]

35. Persson I, Bergkvist L, Lindgren C, Yuen J. Hormone replacement therapy and major risk factors for reproductive cancers, osteoporosis, and cardiovascular diseases: evidence of confounding by exposure characteristics. J Clin Epidemiol. May; 1997 50(5):611–618. [PubMed: 9180654]

36. Ringa V, Ledesert B, Gueguen R, Schiele F, Breart G. Determinants of hormonal replacement therapy in recently postmenopausal women. Eur J Obstet Gynecol Reprod Biol. Jul 24; 1992 45(3):193–200. [PubMed: 1511767]

37. Parazzini F, La Vecchia C, Negri E, Bianchi C, Fedele L. Determinants of estrogen replacement therapy use in northern Italy. Rev Epidemiol Sante Publique. 1993; 41(1):53–58. [PubMed: 8465065]

38. Parazzini F. Trends of determinants of hormone therapy use in Italian women attending menopause clinics, 1997–2003. Menopause. Jan-Feb;2008 15(1):164–170. [PubMed: 17581495]

39. Matthews KA, Kuller LH, Wing RR, Meilahn EN, Plantinga P. Prior to use of estrogen replacement therapy, are users healthier than nonusers? Am J Epidemiol. May 15; 1996 143(10): 971–978. [PubMed: 8629615]

40. Rossouw JE, Anderson GL, Prentice RL, 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. Jul 17; 2002 288(3):321–333. [PubMed: 12117397]

41. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA. Jul 17; 2002 288(3):321–333. [PubMed: 12117397]

42. Resnick SM, Henderson VW. Hormone therapy and risk of Alzheimer disease: a critical time. JAMA. Nov 6; 2002 288(17):2170–2172. [PubMed: 12413378]

43. Sherwin BB. Estrogen and memory in women: how can we reconcile the findings? Horm Behav. Mar; 2005 47(3):371–375. [PubMed: 15708768]

44. Hu FB, Stampfer MJ, Manson JE, et al. Trends in the incidence of coronary heart disease and changes in diet and lifestyle in women. N Engl J Med. Aug 24; 2000 343(8):530–537. [PubMed: 10954760]

45. Coppus AM, Evenhuis HM, Verberne GJ, et al. Early Age at Menopause is Associated with Increased risk of Dementia and Mortality in Women with Down Syndrome. J Alzheimers Dis. Oct 8.2009
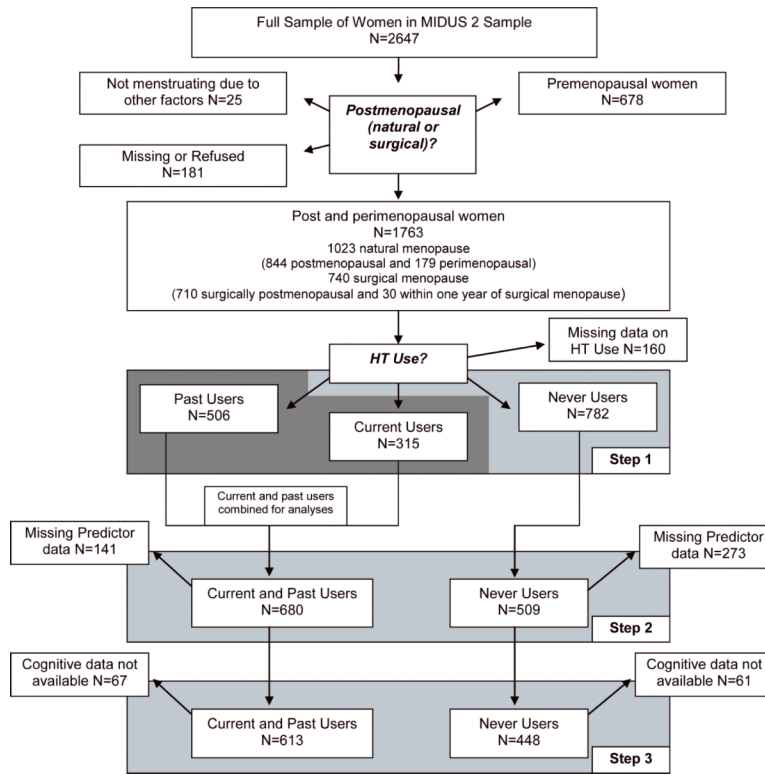
Figure 1.
Selection of participants included in current analyses. Sample is from MIDUS 2 data set, which includes approximately 70% of MIDUS 1 participants. In **Step 1** of analysis, 49 potential predictors were tested individually for univariate association with HT use (N=1603). In **Step 2**, 26 predictors surviving univariate analysis were entered in combination into multiple logistic regression models to predict HT use (N=1,189). Based on resultant model HT use propensity scores, meant to quantify Healthy User traits were derived for every subject. Participants were grouped into tertiles, representing the degree to which participants resembled a HT user (i.e., a `healthy user'). In **Step 3**, cognitive abilities of the three tertiles of women were examined.

Table 1

Table includes twenty-six variables considered potential predictors of HT use. For categorical variables, p-values were derived using regression models predicting HT Use (never user v. past or current user); chi square models were used for categorical variables. Age and Education for the groups are also provided.

Variables	Never User		Past User		Current User		Combined Past and Current Users		Significance: Never Users v Past or Current Users
	N		N		N		N		p-value
Demographic and Anthropometric Variables									
Age in years, mean (SD)	782	59.8 (11.2)	506	62.7 (9.1)	315	59.5 (9.4)	821	61.5 (9.3)	0.01
Education, N(%) High School degree	782	704 90.0%	505	472 (93.5%)	314	285 (90.8%)	819	757 92.4%	0.40
Marital Status, N(%) married	781	455 (58.2%)	506	319 (63.0%)	315	220 (69.8%)	821	539 (65.7%)	0.01
Annual household income, mean K (SD)	782	65.7 (51.5)	506	69.0 (51.7)	314	73.3 (53.5)	820	70.7 (52.4)	0.06
Race, N(%) non-Caucasian	777	77 (9.9%)	505	27 (5.3%)	314	18 (5.7%)	819	45 (5.5%)	<0.01
Body Mass Index, mean (SD)	740	28.1 (6.7)	475	27.7 (6.2)	295	27.1 (5.1)	770	27.4 (5.8)	0.04
Waist to Hip Ratio, mean (SD)	713	0.9 (0.1)	468	0.8 (0.1)	291	0.8 (0.1)	759	0.8 (0.1)	0.01
Psychological and Personality Variables									
Optimism, mean total of responses[a] (SD)	779	22.8 (5.1)	502	23.7 (4.8)	314	23.7 (4.9)	816	23.7 (4.8)	<0.01
Self Esteem, mean total of responses[a] (SD)	779	36.7 (7.7)	501	38.0 (7.4)	315	37.9 (7.5)	816	37.9 (7.4)	<0.01
Autonomy[b], mean total of responses[a] (SD)	781	37.0 (7.3)	505	36.2 (7.2)	315	36.1 (7.1)	816	36.2 7.1	0.03
Environmental Mastery[b], mean total of responses[a] (SD)	781	37.6 (7.8)	505	38.2 (7.6)	315	38.3 (7.9)	816	38.2 (7.7)	0.10
Personal Growth[b], mean total of responses[a] (SD)	781	38.3 (7.3)	505	39.2 (6.6)	315	39.1 (7.3)	816	39.2 (6.9)	0.02
Positive Relations[b], mean total of responses[a] (SD)	781	41.2 (7.1)	505	42.2 (6.4)	315	41.8 (6.9)	816	42.0 (6.6)	0.02
Purpose in Life[b], mean total of responses[a] (SD)	781	37.7 (7.2)	505	38.8 (6.8)	315	38.9 (7.0)	816	38.8 (6.9)	<0.01
Self Acceptance[b], mean total of responses[a] (SD)	781	37.5 (8.8)	505	38.3 (8.2)	315	38.4 (8.2)	816	38.4 (8.2)	0.05
Mood, Level of Satisfaction and Self-Evaluation Variables									
Life Satisfaction, mean response[a] (SD)	782	7.6 (1.7)	506	7.7 (1.6)	315	7.8 (1.6)	821	7.8 (1.6)	0.04
Self rated health, N(%) rating health as Good	782	646 (82.6%)	506	432 (85.4%)	315	270 (85.7%)	821	702 (85.5%)	0.10
Health Related Behaviors									
Reports regular use of Multivitamins, N(%)	782	420 (53.7%)	506	346 (68.4%)	315	219 (69.5%)	821	565 (68.8%)	<0.01

Variables	Never User		Past User		Current User		Combined Past and Current Users		Significance: Never Users v Past or Current Users
	N		N		N		N		p-value
Reports regular use of Vitamin C, N(%)	782	195 (24.9%)	506	168 (21.5%)	315	90 (28.6%)	821	258 (31.4%)	<0.01
Reports regular use of Calcium, N(%)	782	315 (40.3%)	506	310 (61.3%)	315	185 (58.7%)	821	495 (60.3%)	<0.01
No. of physical exams/year, N(%) reporting 1–2	584	321 (55.0%)	483	308 (63.8%)	283	179 (63.3%)	766	487 (63.6%)	<0.01
No. of dental exams/year, N(%) reporting 1–2	566	341 (60.2%)	466	332 (71.2%)	276	201 (72.8%)	742	533 (71.8%)	<0.01
No. of eye exams/year, N(%) reporting 1–2	557	301 (54.0%)	459	287 (62.5%)	265	165 (62.3%)	724	452 (62.4%)	<0.01
Frequency of mental activities, mean score[a] (SD)	762	16.6 (5.3)	494	17.5 (5.0)	306	17.2 (5.1)	800	17.4 (5.0)	<0.01
Alcohol use, N(%) reporting light to moderate use	746	593 (79.5%)	492	418 (85.0%)	308	253 (82.1%)	800	671 (83.9%)	0.04
Physical Health and Medical History									
Diabetes, N(%) experience or treated for condition	782	91 (11.6%)	506	50 (9.9%)	315	17 (5.4%)	821	67 (8.2%)	0.02
Physical limitations, mean score[a] (SD)	742	18.2 (8.3)	487	18.7 (8.4)	306	19.1 (8.8)	793	18.9 (8.5)	0.10
N of prescription medications, mean number (SD)	748	1.7 (1.9)	486	2.1 (1.7)	308	2.8 (1.7)	794	2.3 (1.7)	<0.01

[c]PANAS: Positive and Negative Affect Schedule

[d]Score reflects frequency of activity in two settings (work and leisure) and across two time periods (Winter and Summer)

[a]Higher scores indicate higher levels of the variable being measured

[b]Scale from the Psychological Well Being Inventory

Table 2

Twenty-six variables were entered in to multiple logistic regression models predicting HT Use (Never v. Past or Current). Using SAS PROC SCADGLIM, Version 1.1 (http://methodology.psu.edu) as a variable selection method, nine variables were associated with HT use. These variables were used to derive propensity scores.

Variable	B	S.E.	P-value	Odds Ratio
Demographic and Anthropometric Variables				
Race				
Caucasian and Asian groups combined (R)				
African American	−1.26	0.39	<0.01	0.29
Other	0.00	0.34	1.00	1.00
Waist to Hip Ratio	−2.36	0.85	0.01	0.09
Annual household income	0.22	0.07	<0.01	1.24
Psychological and Personality Variables				
Self Esteem	0.03	0.01	<0.01	1.03
Autonomy	−0.04	0.01	<0.01	0.97
Mood, Level of Satisfaction and Self-Evaluation Variables				
Health Related Behaviors				
Reports regular use of Calcium	0.78	0.13	<0.01	2.18
N of physical exams/year				
0 routine visits/year (R)				
1–2 routine visits/year	0.83	0.21	<0.01	2.29
3 or more routine visits/year	0.74	0.23	<0.01	2.09
Physical Health and Medical History				
Diabetes, experience or treated for condition	−1.00	0.22	<0.01	0.37
N of prescription medications	0.36	0.05	<0.01	1.44
Constant	−1.65	1.20	0.17	0.19

(R) - Reference category of the variable

Table 3

Results of the Regressions of Cognitive Task Performance (Stop and Go Switch task; SGST) on HT Use for Upper Tertile of Propensity to be an HT User / Healthy User (i.e., women possessing characteristics of HT Users). Women in the highest tertile were likely to be Caucasian or Asian, have lower waist to hip ratios, higher household income, report higher self-esteem, but lower sense of autonomy, use calcium supplements regularly, take more prescription medications, get regular physical exams, and were less likely to have diabetes compared to women in lower propensity score tertiles. **Table 3.A.** results of regressions on four parameters of SGST performance. **Table 3.B.** describes results for regression on two subsets of trials for the Mixed-Switch conditions.

Table 3.A. Four global performance scores

	B	SE	t	p-value
Median reaction time under the Normal condition (Normal RT)				
(Constant)	0.891	0.332	2.680	0.008
Age	0.008	0.005	1.642	0.101
Education	0.002	0.018	0.110	0.913
HT Use	**−0.181**	**0.102**	**−1.782**	**0.085**
Median reaction time under the Reverse condition (Reverse RT)				
(Constant)	0.856	0.088	9.684	<0.001
Age	0.004	0.001	2.972	0.003
Education	−0.012	0.005	−2.650	0.008
HT Use	**−0.028**	**0.027**	**−1.030**	**0.304**
Median reaction time from the Mixed-Task trial wherein a condition switch did not occur (Mixed-Nonswitch RT)				
(Constant)	0.877	0.080	10.930	<0.001
Age	0.004	0.001	3.372	0.001
Education	−0.012	0.004	−2.833	0.005
HT Use	**−0.026**	**0.025**	**−1.042**	**0.298**
Median reaction time for Mixed-Tasks items wherein a condition switch occurred (Mixed Switch Normal to Reverse or Mixed Switch Reverse to Normal)				
(Constant)	0.917	0.178	5.140	<0.001
Age	0.009	0.003	3.573	<0.001
Education	−0.013	0.009	−1.320	0.188
HT Use	**−0.132**	**0.055**	**−2.416**	**0.016**

Table 3.B. Subsets of performance scores for Mixed-task Switch Tasks

	B	SE	t	p-value
Median reaction time for Mixed-Tasks items wherein the condition switched from Reverse to Normal (Mixed-Switch Normal Switch RT)				
(Constant)	0.812	0.261	3.114	0.002
Age	0.014	0.004	3.650	<0.001
Education	−0.004	0.014	−0.299	0.765
HT Use	**−0.198**	**0.079**	**−2.492**	**0.030**
Median reaction time for Mixed-Tasks items wherein the condition switch from Normal to Reverse (Mixed-Switch Reverse Switch RT)				

Table 3.B. Subsets of performance scores for Mixed-task Switch Tasks

	B	SE	t	p-value
(Constant)	0.661	0.061	10.837	<0.001
Age	0.004	0.001	5.117	<0.001
Education	−0.007	0.003	−2.273	0.024
HT Use	**−0.022**	**0.019**	**−1.161**	**0.246**