American Journal of Epidemiology

Vol. 144, No. 11
Printed in U.S.A.

Estrogen Replacement Therapy and Cognitive Functioning in the Atherosclerosis Risk in Communities (ARIC) Study

Moyses Szklo,[1] James Cerhan,[2] Ana V. Diez-Roux,[1] Lloyd Chambless,[3] Lawton Cooper,[4] Aaron R. Folsom,[5] Linda P. Fried,[6] David Knopman,[7] and F. Javier Nieto[1]

The association of estrogen replacement therapy (ERT) with cognitive functioning was assessed in 6,110 women aged 48–67 years participating in the Atherosclerosis Risk in Communities (ARIC) study, a multicenter longitudinal investigation. ERT was evaluated in relation to results of three cognitive tests (the Delayed Word Recall (DWR) Test, the Digit Symbol Subtest of the Wechsler Adult Intelligence Scale-Revised (DSS/WAIS-R), and the Word Fluency (WF) Test) using data from the first follow-up visit of the cohort (1990–1992). No consistent associations were seen between ERT and either the DWR test or the DSS/WAIS-R after adjusting for age, education, and additional covariates previously found to be associated with cognitive function scores. Among surgically menopausal women aged 48–57 years, adjusted mean WF scores were slightly greater in ERT current users (mean WF 35.9) than in never users (mean WF 33.5) ($p < 0.02$); and within current users, adjusted WF scores increased with duration of ERT use. However, the finding that ERT was associated with a slightly higher level of performance on only one of three measures offers little support for the hypothesis that ERT has a major protective effect on cognitive function in women less than 68 years of age. The generalizability of these findings to older women who are more likely to experience cognitive decline and who may be using ERT for longer periods of time is limited by the relatively young age of the cohort. *Am J Epidemiol* 1996;144: 1048–57.

cognition; cognition disorders; estrogen replacement therapy; menopause; women's health

Age-related loss in cognitive function is likely to be the result of complex relations between biologic and environmental factors, involving multiple risk factors and mechanisms (1–5). It has been suggested, however, that an important mechanism may be a deficit in central cholinergic transmitter activity (2, 6–10). Estrogens are known to affect the synthesis of acetylcholine through an increase in the activity of choline acetyltransferase (8). Estrogen may also be important in maintaining neuronal interconnections in localized areas of the brain such as the basal forebrain, hippocampus, and cerebral cortex, which are all important areas for cognitive functioning (11–13).

Assessment of the relation of estrogen replacement therapy (ERT) to cognitive functioning in humans has been based on studies of small samples (14–17) or studies of older women (18), including those with Alzheimer's disease (19, 20). Unlike research evaluating the benefits of estrogens on clinical cardiovascular outcomes (21), these studies have yielded inconsistent results (16, 18–20). Thus, there is a need to continue to examine the relation of ERT to cognitive functioning in large samples of women. Data from the Atherosclerosis Risk in Communities (ARIC) study, on which the present analyses were based, offered an opportunity to examine the hypothesis of a protective effect of ERT on cognitive test performance in a large crosssectional sample of women, many of whom had become postmenopausal relatively recently.

Received for publication January 4, 1996, and accepted for publication July 9, 1996.

Abbreviations: ARIC, Atherosclerosis Risk in Communities (Study); DSS/WAIS-R, Digit Symbol Subtest of the Wechsler Adult Intelligence Scale-Revised; DWR, Delayed Word Recall (Test); ERT, estrogen replacement therapy; WF, Word Fluency (Test).

[1] Department of Epidemiology, School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, MD.
[2] Department of Preventive Medicine and Environmental Health, The University of Iowa, Iowa City, IA.
[3] Collaborative Studies Coordinating Center, Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC.
[4] National Heart, Lung, and Blood Institute, National Institutes of Health, Division of Epidemiology and Clinical Applications, Bethesda, MD.
[5] Division of Epidemiology, School of Public Health, University of Minnesota, Minneapolis, MN.
[6] The Welch Center for Prevention, Epidemiology, and Clinical Research, The Johns Hopkins Health Institutions, Baltimore, MD.
[7] Department of Neurology, University of Minnesota, Minneapolis, MN.
Reprint requests to Dr. Moyses Szklo, Department of Epidemiology, School of Hygiene and Public Health, The Johns Hopkins University, 615 N. Wolfe St., Baltimore, MD 21205.

MATERIALS AND METHODS

Study population

The ARIC study is a prospective investigation of clinical and subclinical atherosclerosis in four US communities: Forsyth County, North Carolina; Jackson, Mississippi; selected Minneapolis suburbs, Minnesota; and Washington County, Maryland. The study included population samples totaling 15,792 individuals aged 45–64 years at the time of the baseline examination of which 8,685 (55 percent) were women. About 14 percent in Forsyth County were African-American. Participants in Jackson were exclusively African-American. The Minneapolis and Washington County cohorts were predominantly white. An overview of the study design and procedures has been published (22).

A baseline examination of the total cohort was carried out in 1987–1989 (visit 1) and a follow-up examination, 3 years later, in 1990–1992 (visit 2) when the cohort was aged 48–67 years. Of the 8,481 African-American and white women who were alive at the time of the first follow-up visit (visit 2), 7,921 (93.4 percent) completed this visit, which included cognitive function tests. Women were excluded from our analyses if information was missing on any of the cognitive test scores (n = 84), if they had primary amenorrhea or their menopausal status could not be precisely determined (n = 821), or if they were postmenopausal women who lacked information on current or past use of ERT (n = 238). Women with a history of stroke or transient ischemic attacks (n = 465) or those taking antipsychotic medications (n = 197) were also excluded. In addition, six women with missing information on education were excluded. The final study population on which the present analyses are based was comprised of 6,110 women.

Study variables

For the present report, data on menopausal status, use of ERT, and cognitive scores were obtained in the first follow-up visit (visit 2). Information on education, self-reported health status, fibrinogen, and sport index was available only for the baseline visit (visit 1). All other variables are based on visit 2.

Three neuropsychological tests were applied to the cohort in the follow-up visit: the Delayed Word Recall (DWR) Test (23), the Digit Symbol Subtest of the Wechsler Adult Intelligence Scale-Revised (DSS/WAIS-R) (24), and the Word Fluency (WF) (or Controlled Oral Word Association) Test (25) of the Multilingual Aphasia Examination (26). The tests were administered in study clinics located in each community during one session in a quiet room. The DWR and WF tests were administered by trained interviewers. The DSS/WAIS-R test is self-administered after standardized instruction and is a timed test.

Interviewer performance in ARIC is monitored by tape recording and reviewing by the study coordinator of a random sample of taped interviews. No systematic departures from the protocol were detected by listening to the tapes. In addition, mean scores of these neuropsychological tests obtained by different interviewers were found to be similar.

The DWR test is a test of verbal learning and recent memory. It requires the respondent to recall 10 common nouns after a 5-minute interval during which another test is given. To standardize the elaborative processing of the words to be recalled, individuals are required to compose sentences incorporating the nouns as presented. Test scores range from zero to 10 words recalled. This test has been shown to have a high 6-month test-retest reliability in a study including 26 normal elderly persons (Pearson's correlation coefficient (r) = 0.75) (23).

The DSS/WAIS-R is a paper and pencil test requiring timed translation of numbers 1–9 to symbols using a key. The test measures psychomotor performance and is relatively unaffected by intellectual ability, memory, or learning for most adults (24). It appears to be a sensitive and reliable marker of brain damage (26). The Digit Symbol Subtest was scored as the number of numbers correctly translated to symbols within 90 seconds; the maximum possible is 93. Short-term test-retest reliability has been found to be high in middle-aged individuals (r = 0.82) (24).

The WF test requires the participant to generate as many words as possible in 60 seconds beginning with a letter from the alphabet. Three trials using the letters F, A, and S were conducted, and the WF score was the total number of words generated over the three trials. The test is particularly sensitive to linguistic impairment (25, 27) and early mental decline in older persons (28). It is also a sensitive indicator of damage to the left lateral frontal lobe (25, 27). The immediate test-retest correlation coefficient based on an alternate test form has been found to be 0.82 (29).

Answers to questions on menopause and current use of ERT were ascertained by a trained interviewer. Premenopausal women were those who reported having menstruated in the 2 years before the ARIC examination and who labeled themselves as premenopausal. Women who reported that they had menstruated in the 2 years before the examination but who labeled themselves as postmenopausal or as uncertain menopausal status were categorized as perimenopausal. Postmenopausal women were those who had not menstruated in the last 2 years. Postmeno-

pausal women were further classified into two groups according to type of menopause: surgical menopause, if they had had a bilateral oophorectomy, or natural menopause. The natural menopause group also included nonmenstruating women 55 years of age or older who had had a hysterectomy and had at least one intact ovary. The menopausal status of women less than 55 years of age who had had a hysterectomy without bilateral oophorectomy could not be determined, and they were not included in the final study sample. Postmenopausal women were subclassified as current users (alone or in combination with progestin), former users, and never users of ERT.

ERT included the use of estrogen or estrogen and progestin preparations. Current users included 938 users of estrogen alone, of whom 84 percent took conjugated estrogens, and 246 were users of estrogen and progestin preparations. Of the latter group, 83 percent were users of conjugated estrogens plus medroxyprogesterone acetate. Former users included 588 past users of estrogen and 131 past users of estrogen and progestin preparations.

Information on education was self-reported by study participants. Diastolic and systolic blood pressure levels were calculated as the average of the second and third of three consecutive measurements with a random zero sphygmomanometer. Hypertensives were individuals who had systolic blood pressure of $\geq$140 mmHg, or diastolic blood pressure of $\geq$90 mmHg, or were taking antihypertensive medication. Women were classified as diabetic if they self-reported diabetes, were taking medication for diabetes, had a fasting plasma glucose level $\geq$140 mg/100 ml, or had a nonfasting glucose level $\geq$200 mg/100 ml. Standardized interviews were conducted to determine self-reported physician-diagnosed history of stroke or transient ischemic attack. Physical activity during sport was assessed using a modified version of the Baecke et al. Questionnaire (30) and summarized in a sport index. A depression symptoms score (0–26 range from low to high) was defined using 13 depression-related items from the Maastricht Questionnaire (31). Body mass index was calculated as weight(kg)/height(m^2). Plasma fibrinogen was assessed as previously described (22). Marital status, self-reported health status, and history of smoking and alcohol intake were ascertained by means of interviews.

Statistical analyses

Unadjusted mean and percentile values of cognitive test scores were examined in relation to menopausal status and use of ERT. The distribution of potential confounders by use of ERT in menopausal women was examined by comparing means and proportions across groups.

Adjustment for selected variables was carried out using linear regression methods. Selection of potential confounding variables was based on examination of associations with ERT and cognitive function. Mean scores were adjusted for age, race, education, marital status, self-reported health status, depression score, smoking status, drinking status, hypertension, diabetes, plasma fibrinogen, body mass index, and sport index. Among postmenopausal women, scores were also adjusted for time since menopause. Age, depression score, fibrinogen, sport index, body mass index, and time since menopause were included as continuous variables. All other variables were included as dummy variables. Education was categorized as follows: 8th grade or less, 9–11th grade, high school completion, vocational school, incomplete college, college completion, and graduate or professional school. Self-reported health status was based on four categories: excellent, good, fair, and poor. Smoking and drinking status was categorized as current, former, and never. Marital status was categorized as married, widowed, separated/divorced, and never married.

The adjusted analyses of mean scores were done for two age strata in years: 48–57, which included pre-, peri-, and postmenopausal women; and 58–67, including only postmenopausal women (these age groups correspond approximately to the age groups 45–54 and 55–64 at visit 1 commonly used in reports of ARIC baseline findings). To explore the relation between duration of ERT use and cognitive scores, current and former users of ERT were categorized into three groups based on the tertiles of the distribution of the duration of use of estrogens. Adjusted mean differences for each duration of use category with respect to never users were estimated for current and former users separately.

RESULTS

A total of 6,110 women were available for analysis, of whom 4,549 were white (74.5 percent). As seen in table 1, the mean age of these women was about 57 years at visit 2, when the cognitive tests were conducted. Thirty-four percent had at least some college education.

At the time of the first follow-up visit, 78.0 percent of the study participants were postmenopausal. Bilateral oophorectomy (surgical menopause) had occurred in 17.8 percent of all women and accounted for nearly one fourth of the postmenopausal women. As expected, the proportion of ERT use, particularly current use, was appreciably greater in surgically menopausal than in naturally menopausal women (table 2). Mean

TABLE 1. Characteristics of 6,110 women in the study sample, the Atherosclerosis Risk in Communities Study

Characteristic	%	Mean (SD*)
Age at visit 2		57 (5.6)
Education		
Incomplete high school	20.6	
Complete high school or vocational school	45.6	
College or more	33.8	
Menopausal status and use of ERT* at visit 2		
Premenopausal	12.0	
Perimenopausal	10.0	
Postmenopausal	78.0	
Cognitive scores at visit 2		
DWR*		6.9 (1.5)
DSS/WAIS-R*		47.0 (14.5)
WF*		34.1 (12.3)

* SD, standard deviation; DWR, Delayed Word Recall Test; DSS/WAIS-R, Digit Symbol Subtest of the Wechsler Adult Intelligence Scale-Revised; WF, Word Fluency Test.

TABLE 2. Percentages of ERT* use by type of menopause for 6,110 women in the study sample, the Atherosclerosis Risk in Communities Study

Type of menopause	%
Natural	
ERT never	69.3
ERT former	12.9
ERT current	17.8
Surgical	
ERT never	28.9
ERT former	22.5
ERT current	48.6

* ERT, use of estrogen or estrogen + progestin.

cognitive scores for the total study population were 6.9, 47.0, and 34.1 for the DWR, DSS/WAIS-R, and the WF tests, respectively. The coefficients of variation for these tests were, respectively, 22, 31, and 36 percent, indicating reasonable variability around the mean values.

In general, current users tended to have used hormones longer than former users, and older women had used hormones longer than younger women (data not shown in a table). Among women aged 48–57 years, the median duration of estrogen use was 6 years (interquartile range 3–10) among current users and 1 year (interquartile range 0–4 years) among former users. Among women aged 58–67, the median duration of use was 9 years among current users (interquartile range 3–16 years) and 2 years among former users (interquartile range 0–6).

Mean scores for all three tests were inversely related to age and directly related to educational level (not shown in a table). For example, for those with a college education, mean score differences between younger (48- to 57-year-old) and older (58- to 67-year-old) women were 0.5 for the DWR test, 5.6 for the DSS/WAIS-R, and 1.8 for the WF test. In younger subjects, the differences between those with a graduate education and those with incomplete high school were 0.9 for the DWR test, 19.4 for the DSS/WAIS-R, and 18.1 for the WF test.

The distributions of potential confounding variables by ERT use in postmenopausal women are shown in table 3. Current ERT users were younger, better educated, more likely to be white, and more often married than never users. Current users were also more likely than never users to perceive themselves in excellent or good health, were less frequently hypertensive or diabetic, had a lower body mass index, were less often smokers but more often drinkers, and had lower mean depression scores and a lower mean fibrinogen level. Mean values for the sport index were similar in current, former, and never users of ERT. With the exception of depression score and sport index, former users occupied an intermediate position between never and current users in risk factor profiles. Mean age at menopause decreased slightly from never, to former, to current users. Time since menopause was slightly greater in women who reported being former users than in the other two groups.

The unadjusted means of cognitive test scores by menopausal status and use of ERT are not meaningful because they are likely to be heavily confounded by age and education. Thus, for all tests, mean scores were found to be lower for post- than for premenopausal women; among the postmenopausal, they were lower for ERT never users than for current users, with former ERT users showing intermediate values (data not shown in a table). For example, for the WF test, the mean scores were 37.2 for premenopausal and 33.2 for postmenopausal women. For the natural menopausal group, the mean WF test scores were 32.5 for never users of ERT, 34.6 for former users, and 35.9 for current users; for the surgical menopausal group, they were 28.9, 33.1, and 34.9, respectively. A similar pattern was found for the DWR test and the DSS/WAIS-R.

Examination of multivariable-adjusted mean cognitive scores in pre-, peri-, and postmenopausal women was limited to ages 48–57. For ages 58–67, only postmenopausal women were included. The age distributions were fairly homogenous across the age range in both age categories.

Among women aged 48–57 years (table 4), mean WF test scores were slightly higher for ERT former and current users than for never users in both naturally and surgically menopausal women. These differences

TABLE 3. Distribution of potential confounders by use of estrogen replacement therapy (ERT) among postmenopausal women, the Atherosclerosis Risk in Communities Study

	ERT never users		ERT former users		ERT current users	
	Mean ± SE*	%	Mean ± SE	%	Mean ± SE	%
Age (years)	59.3 ± 0.09		59.2 ± 0.19		56.8 ± 0.15	
Incomplete high school		27.1		21.7		14.8
White		69.7		75.4		78.5
Married		68.5		73.0		78.3
Excellent or good health		80.6		81.3		88.0
Hypertensive†		40.6		38.8		35.5
Diabetic†		14.5		12.4		6.4
Body mass index	28.7 ± 0.12		27.9 ± 0.22		26.9 ± 0.15	
Current smokers		21.4		21.1		19.8
Current drinkers		45.5		48.8		55.6
Depression score	7.6 ± 0.11		7.9 ± 0.22		7.0 ± 0.17	
Fibrinogen (mg/dl)	316.9 ± 1.23		308.4 ± 2.31		293.2 ± 1.80	
Sport index	2.3 ± 0.01		2.4 ± 0.03		2.4 ± 0.02	
Age at menopause	46.2 ± 0.10		44.7 ± 0.20		44.2 ± 0.20	
Years since menopause	13.2 ± 0.10		14.5 ± 0.30		12.6 ± 0.20	

* SE, standard error.
† As defined in text.

were statistically significant at the alpha = 0.05 level for current users only among surgically menopausal women. Patterns for the other cognitive scores were inconsistent and not supportive of the hypothesis of a protective effect of ERT.

For 58- to 67-year-old women (table 3), former and current users had slightly higher mean WF test scores than never users among both naturally and surgically menopausal women, but differences were not statistically significant at the alpha = 0.05 level. No consistent patterns were observed for the other cognitive scores.

In both age groups, the patterns observed in postmenopausal women remained similar after additional adjustment for time since menopause.

The analyses above were repeated using the odds of having a cognitive score at or below the 20th percentile relative to having a score equal to or above the median value as the outcome. Results (not shown) were generally similar to those obtained using cognitive scores as continuous variables.

Mean differences in cognitive scores by duration of ERT use in current users are shown in table 5. In postmenopausal women aged 48–57, mean WF test scores increased slightly with increasing duration of ERT use in both naturally and surgically menopausal women.

Using linear regression, this trend was statistically significant when entering the median values for each tertile ($p = 0.004$) in surgically menopausal women. No consistent or statistically significant patterns were observed for the other cognitive scores.

In postmenopausal women aged 58–67 years, mean WF test scores increased slightly in current users who were in the two upper duration of use categories, as compared with never users; however, differences were not statistically significant at the 0.05 level (table 5). Findings for the lowest third of duration of ERT use were not consistent with a protective effect of ERT. No clear patterns were documented for the other cognitive scores. Among former users, no clear associations were observed between duration of ERT use and cognitive scores (data not shown).

Because improvement of depressive mood has been postulated as an explanation for the possible link between ERT and cognitive functioning (14, 18, 32–36), adjusted means were recalculated after removing depression score from the linear regression function; however, results remained virtually unchanged.

Analyses replacing ERT and menopause data obtained in the follow-up visit conducted in 1990–1992 with data collected during the baseline visit (1987–1989) yielded very similar results (data not shown).

DISCUSSION

Current therapies aimed at improving cognitive functioning are based on correcting the deficit in central cholinergic transmitter activity that may explain the age changes in cognitive performance (7, 8). For example, administration of estradiol to oophorectomized female rats is associated with an increased activity of choline acetyltransferase in the brain (8, 9). Estrogen may also be important in maintaining neuronal interconnections. Growth-promoting effects of es-

TABLE 4. Adjusted mean cognitive scores and standard errors (SEs) by menopausal status and use of estrogen replacement therapy (ERT) stratified by age, the Atherosclerosis Risk in Communities Study†,‡

	DWR§	DSS/WAIS-R§	WF§
	Mean ± SE*	Mean ± SE	Mean ± SE
Women aged 48–57 years (n = 3,208)			
Premenopausal (n = 692)	7.2 ± 0.05	50.4 ± 0.41	35.4 ± 0.44
Perimenopausal (n = 564)	7.2 ± 0.06	50.1 ± 0.43	36.0 ± 0.46
Natural menopause			
ERT never (n = 878)	7.1 ± 0.05	49.7 ± 0.36	34.8 ± 0.38
ERT former (n = 150)	7.2 ± 0.11	50.9 ± 0.82	35.1 ± 0.87
ERT current (n = 316)	7.0 ± 0.08	49.3 ± 0.57	35.4 ± 0.61
Surgical menopause			
ERT never (n = 139)	7.0 ± 0.11	50.8 ± 0.86	33.5 ± 0.92
ERT former (n = 117)	7.0 ± 0.12	49.0 ± 0.93	35.7 ± 0.98
ERT current (n = 352)	7.1 ± 0.07	51.0 ± 0.54	35.9 ± 0.57*
Postmenopausal women aged 58–67 years (n = 2,694)			
Natural menopause			
ERT never (n = 1,607)	6.6 ± 0.04	43.9 ± 0.23	32.7 ± 0.26
ERT former (n = 315)	6.8 ± 0.08*	44.0 ± 0.52	33.0 ± 0.58
ERT current (n = 331)	6.6 ± 0.08	43.4 ± 0.52	33.0 ± 0.58
Surgical menopause			
ERT never (n = 159)	6.7 ± 0.11	44.3 ± 0.75	32.7 ± 0.83
ERT former (n = 121)	6.6 ± 0.13	44.8 ± 0.85	33.0 ± 0.95
ERT current (n = 161)	6.6 ± 0.11	43.7 ± 0.74	33.1 ± 0.82

* p value for difference vis-a-vis never users in each group < 0.05.

† Adjusted for age, race, education, marital status, self-reported health status, depression score, smoking status, drinking status, hypertension, diabetes, serum fibrinogen, body mass index, and sport index, as defined in text. All scores rounded to nearest tenth.

‡ Of the 6,110 women in the study population, 52 were excluded from this table because they were 58 years of age or over but were not postmenopausal. An additional 156 women were excluded because they had missing information on one or more of the covariates.

§ DWR, Delayed Word Recall Test; DSS/WAIS-R, Digit Symbol Subtest of the Wechsler Adult Intelligence Scale-Revised; WF, Word Fluency Test.

trogen on neurons have been demonstrated in organotypic explant cultures of adult rat CNS (37, 38) and in studies of bilateral ovariectomized rats (39). The latter studies have shown a significant decrease in apical dendritic spine density in the CA1 pyramidal cells of the hippocampus of ovariectomized rats, which can be blocked with the administration of estrogen. Furthermore, estrogen may make neurons of the basal forebrain (12), as well as the hippocampus and cerebral cortex (13), more sensitive to neurotrophins (e.g., nerve growth factor); neurotrophins have an important role in the growth and maintenance of dendrites and axons. In addition, it has also been hypothesized that in humans, ERT could indirectly affect cognitive functioning through an improvement of the depressive mood that seems to occur with menopause (14, 18, 32–36).

Broad cognitive domains including psychomotor speed and efficiency, language, and memory are sampled by the three neuropsychological instruments used in the present study, which have been shown to be sensitive to brain dysfunction. The WF test—for which the only somewhat consistent associations were documented—is sensitive to linguistic impairment and has been used in a dementia screening battery (40) and to detect early mental decline in older persons (27, 28). The WF test is also sensitive to frontal lobe dysfunction, particularly that of the left frontal lobe (41). However, the WF test is not specific for frontal lobe dysfunction, and scores can be influenced by damage to a variety of areas of the brain (27). Thus, no conclusions can be drawn from lower performance on the WF test and decreased function of a specific area of the brain. However, the difference in effects of WF compared with DSS/WAIS-R scores could be consistent with prior claims that the effects of estrogen are more potent on verbal processes (18, 42). Perhaps the effect on verbal performance is at the level of retrieval rather than encoding, to account for the lack of effect on the DWR results.

The current study, the largest to date, provides weak (if any) support for the hypothesis that ERT is inde-

TABLE 5. Adjusted mean differences and standard errors (SEs) in cognitive scores by duration of use of estrogen replacement therapy (ERT) in current users stratified by age, the Atherosclerosis Risk in Communities Study†

	DWR‡	DSS/WAIS-R‡	WF‡
	Mean ± SE	Mean ± SE	Mean ± SE
Postmenopausal women aged 48–57 years			
Natural menopause (n = 1,118)			
ERT never users	Reference	Reference	Reference
Lowest third (0–3 years)	−0.1 ± 0.12	−1.2 ± 0.94	0.0 ± 1.00
Middle third (4–8 years)	0.0 ± 0.14	−1.5 ± 1.02	0.2 ± 1.08
Upper third (9–44 years)	−0.3 ± 0.19	0.5 ± 1.45	1.4 ± 1.54
Surgical menopause (n = 480)			
ERT never users	Reference	Reference	Reference
Lowest third (0–3 years)	0.4 ± 0.21	0.9 ± 1.60	0.4 ± 1.76
Middle third (4–8 years)	0.2 ± 0.19	0.3 ± 1.43	3.1 ± 1.57
Upper third (9–44 years)	0.1 ± 0.16	−1.2 ± 1.25	3.7 ± 1.38*
Postmenopausal women aged 58–67 years			
Natural menopause (n = 1,874)			
ERT never users	Reference	Reference	Reference
Lowest third (0–5 years)	0.0 ± 0.13	−0.8 ± 0.84	−1.4 ± 0.94
Middle third (6–13 years)	0.0 ± 0.15	0.0 ± 0.97	1.8 ± 1.08
Upper third (14–46 years)	−0.1 ± 0.17	−0.1 ± 1.13	1.9 ± 1.26
Surgical menopause (n = 318)			
ERT never users	Reference	Reference	Reference
Lowest third (0–5 years)	−0.1 ± 0.31	−1.9 ± 1.97	−1.0 ± 2.16
Middle third (6–13 years)	−0.1 ± 0.27	1.3 ± 1.69	1.2 ± 1.85
Upper third (14–46 years)	−0.2 ± 0.22	−1.1 ± 1.38	1.6 ± 1.51

* p value for linear trend using median values for each tertile 0.004.

† Adjusted for age, race, education, time since menopause, marital status, self-reported health status, depression score, smoking status, drinking status, hypertension, diabetes, serum fibrinogen, body mass index, and sport index, as defined in text. All scores rounded to nearest tenth. Tertiles based on distribution of duration of use in each age group.

‡ DWR, Delayed Word Recall Test; DSS/WAIS-R, Digit Symbol Subtest of the Wechsler Adult Intelligence Scale-Revised; WF, Word Fluency Test.

pendently related to cognitive functioning in postmenopausal women less than 67 years of age. Consistent associations were documented for only one of the tests—the WF test—and were statistically significant only in surgically menopausal women 48–57 years of age. Although a stronger ERT effect in surgically than in naturally menopausal women is consistent with the hypothesis of a protective effect of ERT, the fact that patterns were clearer in younger than in older women is not consistent, inasmuch as one would expect the ERT effect to be stronger in older women who are more likely to be affected by cognitive decline. Chance is a possible explanation for our findings. We performed a total of 24 comparisons: Former and current ERT users were compared with never users in two categories of menopause (natural and surgical) and two age groups (48–57 and 58–67 years) using three cognitive tests (see table 4). Of these comparisons, only one (4 percent) was statistically significant. In addition, even when associations were documented, they tended to be relatively weak. For example, in the younger (48- to 57-year-old) surgically menopausal women, in whom the association with mean WF score was strongest, the difference between current and never users—although statistically significant—translates into an average of only 2.4 words generated per minute (table 4). In postmenopausal women 48–57 years old, increasing duration of use of ERT did appear to be associated with increasing WF scores. However, because of the high correlation of word fluency with education, the WF test scores may be an additional independent proxy for intellect. Thus, any misclassification or residual confounding by education and the lack of a direct measure of intellect could explain the associations observed both for current versus never users and for duration of use. Furthermore, it could be postulated that if ERT exerts a protective effect on word fluency but not recall, then its protective effect is probably not on some type of pre-Alzheimer's process.

Previous studies examining the association of ERT with cognitive function in individuals both with and

without Alzheimer's disease have yielded inconsistent results. In small studies, results have either favored (14) or not favored (15, 17) the hypothesis that ERT protects against menopause-related decline in cognition.

Sherwin (16) conducted a crossover randomized clinical trial including 50 women undergoing surgical menopause who were given a combined estrogen-androgen preparation, estrogen alone, or androgen alone. She found no differences in cognitive test results between the postoperative treatment and the preoperative phases. However, presumably because women who had a hysterectomy but whose ovaries were not removed showed stability in both cognitive performance and circulating sex steroid concentrations, the author concluded that changes in the endocrine milieu after bilateral oophorectomy may have an effect, albeit modest, on cognitive functioning.

In a subsequent study of 19 women undergoing hysterectomy with bilateral oophorectomy for benign disease, Phillips and Sherwin (42) randomly assigned 10 women to receive estrogen and nine women to receive placebo after surgery; the Wechsler Memory Scale was given preoperatively and again 2 months postoperatively. The results suggested that estrogen influenced verbal memory (as measured by immediate and delayed recall of paired associates and immediate recall of paragraphs), but not visual memory (as measured by immediate or delayed recall from the Visual Reproduction Test) or attention (as measured by digit span).

An observational prospective study of the associations of ERT with results of 12 cognitive tests was conducted by Barrett-Connor and Kritz-Silverstein (18) in 800 elderly women (mean age 77 years) living in Rancho Bernardo, California. The authors found that current or past use of ERT was largely unrelated to cognitive scores, but women who used estrogen for at least 20 years scored significantly higher on the category fluency test (a test very similar to word fluency) compared with never users.

Recent case-control studies of the relation of ERT to cognitive functioning have focused on patients with Alzheimer's disease. Paganini-Hill and Henderson (20) carried out a nested case-control study within the Leisure World Cohort study. They identified 138 deceased individuals whose death certificates mentioned Alzheimer's disease, senile dementia, dementia, or senility from a total of approximately 2,500 deaths occurring in the cohort from 1981 through 1992. Four death- and birthdate-matched deceased controls were chosen for each case. The odds of "caseness" was found to be lower in ERT users than in nonusers (odds ratio = 0.69, 95 percent confidence interval 0.46–1.03). Although this study provided evidence that ERT may have an effect on postmenopausal Alzheimer's and related diseases, the inclusion of only dead cases and controls makes it difficult to discern the effect on incidence from that on case fatality.

Brenner et al. (19) identified 107 cases of newly diagnosed Alzheimer's disease and 120 age-matched controls among female members of the Group Health Cooperative of Puget Sound, Seattle, Washington. The authors could not find an association with ERT after adjusting for age and hysterectomy before or after age 55 (odds ratio for users/nonusers = 1.1, 95 percent confidence interval 0.6–1.8).

Our results should be interpreted with caution, in view of the limitations of the study. First, our findings are cross-sectional, raising concerns about possible selection bias. If women who are having memory problems are more likely to be prescribed ERT by their physicians, our failure to detect associations may be due to the selection of women with low cognitive scores into the ERT current users group. On the other hand, because survival may be better both in individuals who perform well in cognitive testing (43–46) and in those using ERT (21), persons using ERT and with preserved cognitive functioning may be more likely to be alive and willing to participate in the ARIC study. This type of bias, however, would lead to an overestimation of the associations supporting the hypothesis of a protective effect. Thus, although possibly explaining some of the WF test results, survival bias is an unlikely explanation for our failure to detect consistent associations with the DWR test and the DSS/WAIS-R results.

Another potential problem in the study is that recall bias regarding use of ERT may be more severe in persons scoring lower on cognitive tests. Furthermore, although we attempted to control for multiple confounding variables, our study also suggests that ERT current users and never users differ in a variety of ways, raising the possibility of residual confounding.

One major caveat regarding the degree to which the data in this paper argue against the protective effect of ERT on cognitive function is the relatively young ages of the study participants. The duration of use among ERT users may have been too short for it to have a detectable effect on cognitive function. If the protective effects of ERT appear only with increasing duration of use and increasing time after menopause, they may not yet be apparent in most of our cohort members. In addition, the study population may be too young to experience substantial cognitive decline, making it virtually impossible to detect a protective effect of ERT, even if it in fact exists. The fact that the only statistically significant associations we observed

were in younger rather than older women (contrary to what one would expect) further argues for the possibility that these associations are due to residual confounding rather than to a true effect of ERT in these age groups.

The finding that ERT was not clearly associated with a higher level of performance may on the one hand reflect the fact that a major protective effect on cognitive functioning is seen only in older women. On the other hand, it is possible that the WF test—for which a slightly better cognitive performance was seen—is more sensitive to early cognitive decline than the DWR test in the pre-Alzheimer's process. With additional cohort follow-up, the ARIC investigators will be able to revisit the hypothesis that ERT has a protective effect on cognitive decline using prospective data on an aging cohort. Additional follow-up will allow assessment not only of cognitive functioning at a given visit but also of temporal changes.

ACKNOWLEDGMENTS

The ARIC Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, and N01-HC-55022.

The authors thank the staff at the ARIC centers: *University of North Carolina at Chapel Hill, Chapel Hill, North Carolina*: Carol Summers, Catherine Burke, Deanna Horwitz, Carmen Woody, Debbie Rubin-Williams, Witold Sieradzan, Louis Wijnberg, George Williams; *University of Mississippi Medical Center, Jackson, Mississippi*: Agnes L. Hayes, Roberta Howell, Jane G. Johnson, Patricia F. Martin; *University of Minnesota, Minneapolis, Minnesota*: Carol DeYoung, Jaci Dion, Lowell Hedquist, Ellie Justiniano; *The Johns Hopkins University, Baltimore, Maryland*: Sunny Harrell, Carole Shearer, Pam Grove, Mary A. Cocodrilli; *University of Texas Medical School, Houston, Texas*: Valerie Stinson, Pam Pfile, Hoang Pham, Teri Trevino; *The Methodist Hospital-Houston, Houston, Texas*: Maria L. Messi, Val Creswell, Julita Samoro, Wanda Wright; *Bowman-Gray School of Medicine, Winston-Salem, North Carolina*: Regina deLacy, Delilah Cook, Carolyn Bell, Teresa Crotts, Suzanne Pillsbury.

REFERENCES

1. Amaducci LA, Fratiglione L, Rocca WA, et al. Risk factors for clinically diagnosed Alzheimer's disease: a case-control study of an Italian population. Neurology 1986;36:922–31.
2. Bartus RT, Dean RL III, Beer B, et al. The cholinergic hypothesis of geriatric memory dysfunction. Science 1982; 217:408–17.
3. Broe GA, Henderson AS, Creasey H, et al. A case-control study of Alzheimer's disease in Australia. Neurology 1990; 40:1698–707.
4. Graves AB, White E, Koepsell TD, et al. A case-control study of Alzheimer's disease. Ann Neurol 1990;28:766–74.
5. Heyman A, Wilkinson WE, Stafford JA, et al. Alzheimer's disease: a study of epidemiological aspects. Ann Neurol 1984; 15:335–41.
6. Coyle JT, Price DL, DeLong MR. Alzheimer's disease: a disorder of cortical cholinergic innervation. Science 1983;219: 1184–90.
7. Fillit H, Weinreb H, Cholst I, et al. Observations in a preliminary open trial of estradiol therapy for senile dementia—Alzheimer's type (SDAT). Psychoneuroendocrinology 1986;11:337–45.
8. Luine VN. Estradiol increases choline acetyltransferase activity in specific basal forebrain nuclei and projection areas of female rats. Exp Neurol 1985;89:484–90.
9. Luine VN, McEwen BS. Sex differences in cholinergic enzymes of diagonal band nuclei in the rat preoptic area. Neuroendocrinology 1983;36:475–82.
10. Luine VN, Park D, Joh T, et al. Immunochemical demonstration of increased choline acetyltransferase concentration in rat preoptic area after estradiol administration. Brain Res 1980; 191:273–7.
11. McEwen BS, Coirini H, Danielsson A, et al. Steroid and thyroid hormones modulate a changing brain. J Steroid Biochem Molec Biol 1991;40:1–14.
12. Toran-Allerand CD, Miranda RC, Bentham WDL, et al. Estrogen receptors colocalize with low-affinity nerve growth factor receptors in cholinergic neurons of the basal forebrain. Proc Natl Acad Sci 1992;89:4668–72.
13. Miranda RC, Sohrabji F, Toran-Allerand CD. Interactions of estrogens with the neurotrophins and their receptors during neural development. Horm Behav 1994;28:367–75.
14. Fedor-Freybergh P. The influence of oestrogen on the well-being and mental performance in climacteric and postmenopausal women. Acta Obstet Gynaecol Scand 1977;64:5–69.
15. Rauramo L, Langerspetz K, Engblom P, et al. The effect of castration and peroral estrogen therapy on some psychological functions. Front Horm Res 1975;8:133–51.
16. Sherwin BB. Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. Psychoneuroendocrinology 1988;13:345–57.
17. Vanhulle G, Demol R. A double-blind study into the influence of estriol on a number of psychological tests in postmenopausal women. In: Van Keep PA, Greenblatt RB, Albeaux-Fernet M, eds. Consensus on menopausal research. Baltimore, MD: University Park Press, 1976:94–9.
18. Barrett-Connor E, Kritz-Silverstein D. Estrogen replacement therapy and cognitive function in older women. JAMA 1993; 269:2637–41.
19. Brenner DE, Kukull WA, Stergachia A, et al. Postmenopausal estrogen replacement therapy and the risk of Alzheimer's disease: a population-based case-control study. Am J Epidemiol 1994;140:262–7.
20. Paganini-Hill A, Henderson VW. Estrogen deficiency and risk of Alzheimer's disease in women. Am J Epidemiol 1994;140: 256–61.
21. Ross RK, Paganini-Hill A, Mack TM, et al. Cardiovascular benefits of estrogen replacement therapy. Am J Obstet Gynecol 1989;160:1301–6.
22. The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) study: design and objectives. Am J Epidemiol 1989;129:687–702.
23. Knopman DS, Ryberg S. A verbal memory test with high predictive accuracy for dementia of the Alzheimer type. Arch Neurol 1989;46:141–5.
24. Wechsler D. WAIS-R manual. Cleveland, OH: The Psychological Corporation, 1981.
25. Lezak MD. Neuropsychological assessment. 2nd ed. New York, NY: Oxford University Press, 1983:331–2.

26. Benton AL, Hamsher K. Multilingual aphasia examination. 2nd ed. Iowa City, IA: AJA Associates, 1989.
27. Tranel D. Neuropsychological assessment. Psychiatr Clin North Am 1992;15:283–99.
28. Benton AL, Eslinger PJ, Damasio AR. Normative observations on neuropsychological test performances in old age. J Clin Neuropsychol 1981;3:33–42.
29. Franzen MD. Multilingual aphasia examination. In: Keyser DJ, Sweetland RC, eds. Test critiques. Vol 5. Kansas City, MO: Test Corporation of America, 1986:278–82.
30. Baecke JA, Vurema J, Frijters JER. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 1982;36:936–42.
31. Appels A, Mulder P. Excess fatigue as a precursor of myocardial infarction. Eur Heart J 1988;9:758–64.
32. Furuhjelm M, Fedor-Freybergh P. The influence of estrogens on the psyche in climacteric and post-menopausal women. In: Van Keep PA, Greenblatt RB, Albeaux-Fernet M, eds. Consensus on menopause research. Baltimore, MD: University Park Press, 1976:84–93.
33. Gerdes LC, Sonnendecker EW, Polakow ES. Psychological changes effected by estrogen-progestogen and clonidine treatment in climacteric women. Am J Obstet Gynecol 1982;142: 98–103.
34. Klaiber EL, Broverman DM, Vogel W, et al: Estrogen therapy for severe persistent depressions in women. Arch Gen Psychiatry 1979;36:550–4.
35. Kopera H. Estrogens and psychic functions. Fron Horm Res 1973;2:118–33.
36. Malleson J. An endocrine factor in certain affective disorders. Lancet 1953;2:158–64.
37. Matsumoto A, Arai Y. Neuronal plasticity in the deafferented hypothalamic arcuate nucleus of adult female rats and its enhancement by treatment with estrogen. J Comp Neurol 1981;197:197–206.
38. Frankfurt M, Gould E, Wooley C, et al. Gonadal steroids modify spine density in the ventromedial hypothalamus neurons: a Golgi study. Neuroendocrinology 1990;51:530–5.
39. Gould E, Wooley C, Frankfurt M, et al. Gonadal steroids regulate dendritic spine density in hippocampal pyramidal cells in adulthood. J Neurosci 1990;10:1286–91.
40. Eslinger PJ, Damasio AR, Benton AL, et al. Neuropsychological detection of abnormal mental decline in older persons. JAMA 1985;253:670–4.
41. Tranel D, Anderson SW, Benton AL. Development of the concept of "executive function" and its relationship to the frontal lobes. In: Boller F, Grafman J, eds. Handbook of neuropsychology. Vol 9. Amsterdam: Elsevier, 1988.
42. Phillips SM, Sherwin BB. Effects of estrogen on memory function in surgically menopausal women. Psychoneuroendocrinology 1992;17:485–95.
43. Evans DA, Smith LA, Scherr PA, et al. Risk of death from Alzheimer's disease in a community population of older persons. Am J Epidemiol 1991;134:403–12.
44. Liu IY, LaCroix AZ, White LR, et al. Cognitive impairment and mortality: a study of possible confounders. Am J Epidemiol 1990;132:136–43.
45. Swan GE, Carmelli D, LaRue A. Performance on the digit symbol substitution test and 5-year mortality in the Western Collaborative Group Study. Am J Epidemiol 1995;141:32–40.
46. Deeg DJH, Hofman A, van Zonneveld RJ. The association between change in cognitive function and longevity in Dutch elderly. Am J Epidemiol 1990;132:973–82.