ORIGINAL CONTRIBUTIONS

Combined Postmenopausal Hormone Therapy and Cardiovascular Disease: Toward Resolving the Discrepancy between Observational Studies and the Women's Health Initiative Clinical Trial

Ross L. Prentice[1], Robert Langer[2], Marcia L. Stefanick[3], Barbara V. Howard[4], Mary Pettinger[1], Garnet Anderson[1], David Barad[5], J. David Curb[6], Jane Kotchen[7], Lewis Kuller[8], Marian Limacher[9], and Jean Wactawski-Wende[10] for the Women's Health Initiative Investigators

[1] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA.
[2] Department of Family and Preventive Medicine, University of California, San Diego, CA.
[3] Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA.
[4] Medstar Research Institute, Hyattsville, MD.
[5] Department of Obstetrics and Gynecology, Albert Einstein School of Medicine, Bronx, NY.

[6] John A. Burns School of Medicine, University of Hawaii, Honolulu, HI.
[7] Division of Epidemiology, Health Policy Institute, Medical College of Wisconsin, Milwaukee, WI.
[8] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA.
[9] Division of Cardiovascular Medicine, University of Florida College of Medicine, Gainesville, FL.
[10] Department of Social and Preventive Medicine and Gynecology-Obstetrics, University at Buffalo, Buffalo, NY.

Observational research on postmenopausal hormone therapy suggests a 40–50% reduction in coronary heart disease incidence among women using these preparations. In contrast, the Women's Health Initiative clinical trial of estrogen plus progestin found an elevated incidence over a 5.6-year intervention period through July 7, 2002. Toward explaining this discrepancy, the authors analyzed data from this trial, which included 16,608 postmenopausal women aged 50–79 years, and corresponding data from 53,054 women in the Women's Health Initiative observational study, 33% of whom were estrogen-plus-progestin users at baseline. Estrogen-plus-progestin hazard ratio estimates for coronary heart disease, stroke, and venous thromboembolism in the observational study were 39–48% lower than those in the clinical trial following age adjustment. However, hazard ratios tended to decrease with increasing time from initiation of estrogen-plus-progestin use, and observational study hazard ratio estimates are heavily weighted by longer-term use while clinical trial hazard ratio estimates reflect shorter-term use. Following control for time from estrogen-plus-progestin initiation and confounding, hazard ratio estimates were rather similar for the two cohorts, although there was evidence of some remaining difference for stroke. These analyses have implications for both the design and the analysis of observational studies.

cardiovascular diseases; clinical trials; cohort studies; estrogens; hormone replacement therapy; postmenopause; progestins

Abbreviations: CEE, conjugated equine estrogen; CI, confidence interval; MPA, medroxyprogesterone acetate; WHI, Women's Health Initiative.

Editor's note: An invited commentary on this article appears on page 415, and the authors' response appears on page 419.

Few research reports have stimulated as much public response (1, 2) or have engendered as sustained a discussion among medical practitioners and researchers as the results

of the Women's Health Initiative (WHI) randomized, placebo controlled trial of daily use of 0.625 mg of conjugated equine estrogen (CEE) and 2.5 mg of medroxyprogesterone acetate (MPA) (Prempro; Wyeth Ayerst, Madison, New Jersey). While a major reduction in coronary heart disease incidence had been hypothesized on the basis of a substantial body of observational research (3–5), the WHI estrogen-plus-progestin trial found an elevation in coronary heart disease risk and assessed that overall health risks exceeded benefits over an average 5.6-year follow-up period (6, 7).

The lack of explanation for this departure from expectation has prompted some clinicians and researchers to hypothesize flaws in the WHI trial (8, 9). Others have argued lack of relevance of trial results to important groups of combined hormone therapy users. For example, a recent contribution noted that the WHI was not designed to provide a powerful test of cardioprotective effects among women aged 50–54 years in menopausal transition, and it concluded that observational studies provide "the only applicable clinical guide to this issue" (10, p. 1498).

Other authors have speculated on reasons for a discrepancy between WHI trial results and related observational research, citing confounding in observational studies, the limited ability of observational studies to assess short-term effects, differences among combined hormone therapy preparations, and differences among populations of women studied as possible reasons (11–13). Along these lines, a review (14) noted that evidence for coronary heart disease benefit from hormone therapy is not apparent among studies that control for socioeconomic and other confounding factors. The April 2004 issue of the *International Journal of Epidemiology* includes a review (15) and several commentaries (16–21) on this topic that illustrate the continuing diversity of opinion on the sources of the discrepancy and on the clinical implications of the available evidence.

The implications of WHI trial results for the study designs needed to obtain reliable therapeutic or public health information have also been debated. Perspectives have ranged from the statement that "many people suspended ordinary standards of evidence concerning medical interventions and concluded that hormone therapy was the right thing to prevent heart disease in millions of postmenopausal women despite the absences of any large-scale clinical trials quantifying its overall risk-benefit ratio" (22, p. 519) to the assertion that "the good agreement between the observational studies and the [WHI] trial on endpoints other than CHD [coronary heart disease] confirms the utility and validity of observational studies as monitors of new preventive agents" (23, p. 9).

The WHI, with its multifaceted clinical trial among 68,133 women, including 16,608 in the estrogen-plus-progestin trial, and its observational study among 93,676 women, provides an excellent setting to understand and resolve these discrepancies. Specifically, women were recruited to the clinical trial and observational study from the same underlying populations, and over essentially the same time period, at the 40 WHI clinical centers. Many elements of the protocol and procedures were common to the two WHI components, including the baseline questionnaire and interview data collection, as well as major elements of outcome ascertainment. The clinical trial has more information concerning the effects of hormone therapy during the first few years of use, while the observational study mostly provides information on the effects of longer term use. Both clinical trial and observational study women were personally interviewed at baseline concerning prior hormone therapy use, and they were periodically queried to ascertain hormone therapy use during WHI follow-up. Hence, the WHI provides a context for quantitative assessment of the discrepancy between clinical trial and observational study results. To test assertions (11, 23) that the WHI estrogen-plus-progestin results agree closely with those of observational studies for outcomes other than coronary heart disease, we chose to include stroke and venous thromboembolism, in addition to coronary heart disease, in the comparisons presented here.

MATERIALS AND METHODS

Recruitment and screening

Detailed WHI eligibility criteria and recruitment methods have been published (6, 24). Eligible women were 50–79 years of age at initial screening, were postmenopausal, had no medical condition associated with a predicted survival of less than 3 years, and were likely to be residing in the same geographic area for at least 3 years. For the hormone therapy trials, additional exclusion criteria involved safety, adherence, and retention concerns. Women ineligible for, or not interested in, the hormone therapy trials or in an overlapping trial of a low-fat eating pattern were given the opportunity to enroll in the observational study, which was intended to provide new risk factor information on major causes of morbidity and mortality among postmenopausal women. All women provided written, informed consent for their respective WHI activities, and they provided a baseline fasting blood specimen, a medications and dietary supplements inventory, and common core questionnaires that included information on medical history, reproductive history, family history, personal habits, psychosocial attributes, and food frequency (25, 26).

Baseline exogenous hormones and clinical trial hormone regimen

Information on lifetime hormone use was obtained on clinical trial and observational study women at baseline by a trained interviewer, assisted by a structured questionnaire and chart displaying colored photographs of various hormone preparations. For postmenopausal hormone therapy, detailed information was obtained on the preparation, estrogen-and-progestin dose, schedule, and route of administration. The age at starting and stopping each preparation was recorded.

Women interested in the hormone therapy trials who were using postmenopausal hormone therapy at initial screening were required to undergo a 3-month washout period. Women with a uterus were potentially eligible for the combined hormone trial of 0.625 mg of CEE and 2.5 mg of MPA in a single daily tablet or for a matching placebo, while women with a prior hysterectomy were potentially eligible

for a companion trial of unopposed estrogen, which involved randomization to a daily dose of 0.625 mg of CEE or a matching placebo. A group of 331 women (each with a uterus) were initially randomized to estrogen alone. Following the release of the Postmenopausal Estrogen/Progestin Interventions (PEPI) study results (27), these women were unblinded and reassigned to estrogen plus progestin, and they are included in the clinical trial combined hormone group in this analysis.

Study population

This report is based on information from the 16,608 women randomized to the combined hormone trial, 8,506 (51.2 percent) of whom were assigned to active estrogen plus progestin, and from the 53,054 women enrolled in the observational study who were with uterus and not using unopposed estrogen at the time of WHI enrollment. Among these 53,054 women, 17,503 (33.0 percent) were current users of combined estrogen-plus-progestin preparations at baseline.

Follow-up and outcome ascertainment

Follow-up and outcome ascertainment procedures in the clinical trial (6, 7, 28–30) involved semiannual contacts and in-clinic annual visits for the collection of standardized information on safety concerns, adherence to study medications, and structured initial reporting of clinical outcome events. Annual mailed follow-up forms in the observational study updated information on the use of hormone therapy, updated selected other risk factor information, and employed the same structured initial reporting of clinical events.

Disease events were initially self-reported for all three clinical outcomes. Ascertainment of information on coronary heart disease, comprising myocardial infarction and death due to coronary heart disease, involved physician adjudication based on the review of pertinent documents at each clinical center. In the clinical trial and in a fraction in the observational study, coronary heart disease and related outcomes were further adjudicated by a central committee with agreement rates of 90 percent for myocardial infarction and 97 percent for death due to coronary heart disease. Similarly, cases of hospitalized stroke were based (29) on rapid neurologic deficit attributable to obstruction or rupture of the arterial system or on a demonstrable lesion compatible with acute stroke. Central neurologists reviewed all stroke cases, as well as transient ischemia attacks and self-reports of stroke in the clinical trial, along with a fraction of such cases in the observational study. Of locally adjudicated strokes in the clinical trial, 94.5 percent were confirmed on central review, while 93.8 percent of centrally adjudicated strokes had been classified as strokes by local adjudicators. Venous thromboembolism comprised (30) hospitalized deep vein thrombosis and pulmonary embolism. The confirmation rates in the clinical trial for locally adjudicated venous thromboembolism events in central review were 96 percent for deep vein thrombosis and 98 percent for pulmonary embolism. In the observational study, only self-reports of (hospitalized) deep vein thrombosis or pulmonary embolism were routinely obtained. The confirmation rate in the clinical trial for self-reported venous thromboembolism events on central adjudication was 80.4 percent.

Statistical analysis

Primary analyses used time-to-event methods based on the Cox regression procedure (31), with time from randomization in the clinical trial and time from enrollment in the observational study as the basic "time" variable. Disease incidence rates during follow-up were stratified on baseline age in 5-year categories and on the WHI component (clinical trial or observational study). Hence, hazard ratio estimates derive from comparisons among women in the same 5-year age interval and the same WHI component who are at the same length of time from enrollment in the WHI.

Disease events in the estrogen-plus-progestin trial were included through July 7, 2002, when women stopped taking study pills. This gives an average 5.6 years of follow-up and a maximum of 8.6 years of follow-up. Follow-up in the observational study subsample was included through February 28, 2003, giving comparable average (5.5 years) and maximum (8.4 years) durations. We used the best available outcome data, comprising all centrally adjudicated coronary heart disease, stroke, and venous thromboembolism events in the estrogen-plus-progestin trial and all locally adjudicated coronary heart disease, stroke, and self-reported venous thromboembolism events in the observational study.

The possibility that baseline characteristics confound the relation of estrogen-plus-progestin use to cardiovascular disease risk was examined by carrying out regression analysis of the clinical trial and observational study data that included selected baseline risk factors. The dependence of the hazard ratio on time from initiation of the current episode of estrogen-plus-progestin use was examined in Cox regression analyses by estimating separate hazard ratios for less than 2, 2–5, and more than 5 years, with proportional hazards within these time periods. The regression variable for these estrogen-plus-progestin hazard ratios is time dependent as women move from one time from initiation period to another during WHI follow-up. At a specific follow-up time in the WHI, the time from initiation of the current episode for the estrogen-plus-progestin group in the clinical trial was defined as the time from randomization, and for the estrogen-plus-progestin group in the observational study, it was defined by summing the time that a woman had used estrogen plus progestin at baseline plus the time from observational study enrollment. The time that a woman had used estrogen plus progestin at baseline was determined by going back in time from observational study enrollment until a gap in estrogen-plus-progestin usage was encountered, with a usage gap of 1 year or longer defining the starting point for the episode. Combined hormone use in the observational study was classified several ways, including estrogen-plus-progestin preparation, estrogen preparation and dose, and progestin preparation and dose.

TABLE 1.   Cardiovascular disease incidence rates in the Women's Health Initiative clinical trial (1994–2002) and observational study (1994–2003) of estrogen plus progestin

	Clinical trial			Observational study		
	Placebo	Estrogen plus progestin	Ratio	Control	Estrogen plus progestin	Ratio
No. of women	8,102	8,506		35,551	17,503	
Coronary heart disease						
No. of events	147	188		615	158	
Annualized incidence (%)	0.33	0.39	1.18	0.32	0.16	0.50
Age-adjusted* annualized incidence (%)	0.33	0.40	1.21	0.28	0.20	0.71
Stroke						
No. of events	107	151		490	123	
Annualized incidence (%)	0.24	0.31	1.29	0.25	0.13	0.52
Age-adjusted* annualized incidence (%)	0.24	0.32	1.33	0.22	0.17	0.77
Venous thromboembolism						
No. of events	76	167		336	153	
Annualized incidence (%)	0.17	0.35	2.10	0.17	0.16	0.94
Age-adjusted* annualized incidence (%)	0.17	0.35	2.10	0.16	0.17	1.06

* Age adjusted to the 5-year age distribution in the clinical trial cohort.

The sensitivity of hazard ratio estimates to lack of adherence to estrogen-plus-progestin group designation was examined by restricting the follow-up period for each clinical trial or observational study woman to the time period when she remained adherent to her estrogen-plus-progestin or control group designation. Specifically, the follow-up period for each woman was censored 6 months after she stopped taking combined hormones in the estrogen-plus-progestin groups or initiated hormone therapy use in the control groups, after which hazard ratio estimates were recalculated. The 6-month period was included to accommodate hormone therapy changes resulting from diagnostic workup. Nominal 95 percent confidence intervals and two-sided significance tests (*p* values) are presented.

RESULTS

Annualized incidence rates comparing estrogen-plus-progestin groups with corresponding control groups are unfavorable for estrogen plus progestin for each of coronary heart disease, stroke, and venous thromboembolism in the clinical trial (table 1). In contrast, the annualized incidence rate for the estrogen-plus-progestin group in the observational study is only about one-half that in the control group for coronary heart disease and stroke, and it is slightly less than the rate in the control group for venous thromboembolism. However, the average age in the estrogen-plus-progestin and placebo groups was 63.3 years and 63.2 years, respectively, in the clinical trial, compared with 60.8 years in the estrogen-plus-progestin group and 64.7 years in the control group in the observational study. Hence, table 1 also provides annualized incidence rates, adjusted to the 5-year age distribution in the clinical trial cohort. Cox model analyses of these data, with stratification on cohort and baseline

5-year age categories, give an estimated ratio of estrogen-plus-progestin hazard ratio in the observational study to that in the clinical trial of 0.61 (95 percent confidence interval (CI): 0.46, 0.81) for coronary heart disease, 0.58 (95 percent CI: 0.42, 0.82) for stroke, and 0.52 (95 percent CI: 0.37, 0.73) for venous thromboembolism. Hence, these analyses indicate a 39–48 percent lower estimated hazard ratio in the observational study compared with that in the clinical trial (each *p* < 0.01) following age adjustment.

Confounding in the observational study might explain some of this discrepancy. Estrogen-plus-progestin users in the observational study cohort have somewhat more favorable cardiovascular disease profiles for a number of risk factors, compared with the observational study control group (table 2). Details on the lifetime use of hormone therapy preparations for these women at the time of their screening for potential WHI participation show that estrogen-plus-progestin users in the observational study were more likely to have used hormone therapy for extended periods of time prior to WHI enrollment (table 3).

Confounding and estrogen-plus-progestin hazard ratios

Various factors were considered as potential confounding factors in Cox regression analyses of the clinical trial and observational study data. These included established cardiovascular disease risk factors, as well as various behavioral, dietary, and physical activity baseline measures. Factors that related strongly to one or more of coronary heart disease, stroke, or venous thromboembolism incidence were included. Table 4 shows estrogen-plus-progestin hazard ratio estimates in the clinical trial and observational study, while controlling for the listed potential confounding factors for each clinical

TABLE 2. Age-adjusted* baseline risk factor distribution in the Women's Health Initiative clinical trial (1994–2002) and observational study (1994–2003)

Risk factor	Clinical trial		Observational study	
	Placebo	Estrogen plus progestin	Control	Estrogen plus progestin
Ethnicity (%)				
White	83.9	84.0	82.3	89.2
Black	7.1	6.4	8.5	2.5
Hispanic	5.2	5.5	4.3	2.6
American Indian	0.4	0.3	0.5	0.3
Asian/Pacific Islander	2.1	2.3	2.9	4.4
Unknown	1.3	1.5	1.5	1.1
Body mass index, kg/m^2 (%)				
Underweight (<18.5)	0.7	0.7	1.4	1.5
Normal (18.5–24.9)	30.0	29.8	37.6	51.2
Overweight (25.0–29.9)	35.2	35.3	33.7	31.6
Obesity I (30.0–34.9)	20.5	21.4	16.5	10.9
Obesity II (35.0–39.9)	8.9	9.0	6.6	3.2
Extreme obesity III (≥40)	4.6	3.8	4.2	1.6
Education (%)				
0–8 years	2.2	2.4	2.0	0.7
Some high school	4.5	4.4	3.6	1.6
High school diploma/GED†	20.0	19.1	16.8	11.7
School after high school	38.0	39.7	34.8	32.7
College degree or higher	35.3	34.5	42.9	53.4
Smoking (%)				
Never smoked	50.0	49.6	51.3	48.7
Past smoker	39.5	39.9	41.7	46.6
Current smoker	10.5	10.4	7.0	4.7
Age at menopause, years (mean (SD†))	50.0 (4.7)	50.0 (4.8)	50.2 (4.8)	50.8 (4.9)
Physical functioning, score‡ (mean (SD))	82.9 (18.9)	82.8 (18.6)	82.1 (20.0)	85.1 (17.2)

* To the 5-year age distribution in the clinical trial cohort.
† GED, general equivalency diploma; SD, standard deviation.
‡ Paraphrased from the RAND 36-Item Health Survey (32). The score falls in the range 0–100 (with 100 the best).

outcome. The estimated estrogen-plus-progestin hazard ratios in the observational study are increased following accommodation of these factors but remain 30–38 percent lower than in the clinical trial (each $p < 0.05$). The estimated ratio of estrogen-plus-progestin hazard ratio in the observational study to that in the clinical trial is 0.70 (95 percent CI: 0.52, 0.95) for coronary heart disease, 0.72 (95 percent CI: 0.52, 1.02) for stroke, and 0.62 (95 percent CI: 0.43, 0.88) for venous thromboembolism in these analyses.

Table 4 also shows hazard ratios for the potential confounding factors. To ensure a common interpretation of estrogen-plus-progestin hazard ratios in the two cohorts, the same potential confounding factors and corresponding hazard ratios were used in the clinical trial and observational study. Age within 5-year strata relates strongly to each of the cardiovascular disease outcomes, so that a 5-year age stratification alone may be insufficient to avoid confounding.

Similarly, body mass index shows evidence of linear trend with venous thromboembolism risk after allowing for differences in cardiovascular disease rates among the body mass index categories listed. A higher value for the physical functioning construct was associated with lower disease rates. Education relates strongly to coronary heart disease risk and could be an important source of confounding if overlooked. Additional analyses included family income categories and additional behavioral constructs (32) to assess general health, role limitation due to physical health, pain, vitality, social functioning, role limitation due to emotional health, and mental health. Physical activity and dietary factors were also considered with a regression calibration (33) accommodation of measurement error, using data from a measurement precision study in the observational study (26). None of these inclusions had meaningful further influence on the estrogen-plus-progestin hazard ratios.

TABLE 3. Age-adjusted* lifetime hormone therapy use history, expressed as percentage at baseline, in the Women's Health Initiative clinical trial (1994–2002) and observational study (1994–2003)

	Clinical trial		Observational study	
	Placebo	Estrogen plus progestin	Control	Estrogen plus progestin
Unopposed estrogen duration, years				
Never	89.4	89.4	91.4	86.6
<2	5.1	5.1	4.3	3.4
2–5	3.0	2.8	2.2	3.4
>5	2.5	2.7	2.1	6.7
Recency of estrogen-alone use				
Never user	89.4	89.4	91.4	86.6
Current user	0.6	0.6	0.0	0.0
Within past 1–4 years	1.8	1.8	1.9	3.9
Past use, 5–10 years ago	1.9	1.7	1.6	4.1
Past use, >10 years ago	6.3	6.6	5.1	5.4
Estrogen-plus-progestin duration, years				
Never	82.7	82.2	86.0	0.0
<2	7.4	7.2	6.6	14.7
2–5	5.0	5.1	3.4	23.2
>5	5.0	5.5	4.0	62.1
Recency of estrogen-plus-progestin use				
Never user	82.7	82.2	86.0	0.0
Current user	5.5	5.9	0.0	100.0
Within past 1–4 years	7.1	7.4	8.7	0.0
Past use, 5–10 years ago	2.8	2.9	3.4	0.0
Past use, >10 years ago	1.9	1.6	1.9	0.0

* To the 5-year age distribution in the clinical trial cohort.

Hazard ratio dependence on time from initiation of the current estrogen-plus-progestin episode

An important remaining source of discrepancy between clinical trial and observational study hazard ratios is elucidated by accommodating a dependence of hazard ratios on time from initiation of the current estrogen-plus-progestin episode (refer to Materials and Methods). Table 5 includes the same potential confounding factors (not shown for brevity) as shown in table 4. The left side of table 5 provides separate hazard ratio estimates for the clinical trial and observational study in each of the three estrogen-plus-progestin time-from-initiation periods. The numbers of estrogen-plus-progestin group women experiencing cardiovascular disease events in each time period are also shown. These numbers make clear that the observational study is very sparse concerning the first 2 years from estrogen-plus-progestin initiation, while the clinical trial is comparably sparse after 5 years' duration. Note that hazard ratios within time-from-initiation periods are now more similar for each of the three clinical outcomes, with little evidence of hazard ratio reduction with estrogen plus progestin, except possibly for coronary heart disease beyond 5 years from estrogen-plus-progestin initiation.

For coronary heart disease, a test of equality of the three clinical trial hazard ratios with the corresponding three observational study hazard ratios was far from significant ($p > 0.6$), as was also the case for venous thromboembolism ($p > 0.8$). For stroke, however, there was evidence ($p = 0.01$) of residual difference between observational study and clinical trial hazard ratios after adjusting for time from estrogen-plus-progestin initiation and confounding factors. The estimated ratio of estrogen-plus-progestin hazard ratio in the observational study to that in the clinical trial was 0.93 (95 percent CI: 0.64, 1.36) for coronary heart disease, 0.76 (95 percent CI: 0.49, 1.18) for stroke, and 0.84 (95 percent CI: 0.54, 1.28) for venous thromboembolism, following control for confounding factors and allowing a common hazard ratio dependence on time from estrogen-plus-progestin initiation in the two cohorts.

Hazard ratio in subgroups

The analyses described above were repeated in the following baseline subgroups of the clinical trial and observational study cohorts: the approximately 94 percent of clinical trial and observational study women without a personal

TABLE 4.   Cardiovascular disease hazard ratio estimates for estrogen-plus-progestin use and for potential confounding factors, in analyses of data from the Women's Health Initiative clinical trial (1994–2002) and observational study (1994–2003)

Factor	Coronary heart disease		Stroke		Venous thromboembolism	
	Hazard ratio	95% confidence interval	Hazard ratio	95% confidence interval	Hazard ratio	95% confidence interval
Estrogen plus progestin						
In clinical trial	1.27	1.00, 1.61	1.21	0.93, 1.59	2.13	1.59, 2.85
In observational study	0.87	0.72, 1.05	0.86	0.70, 1.07	1.31	1.07, 1.61
Age in years (linear)	1.10	1.05, 1.15	1.07	1.02, 1.13	1.05	1.00, 1.11
White race (no vs. yes)	0.84	0.69, 1.03	1.16	0.95, 1.42	0.65	0.50, 0.84
Body mass index (kg/m^2)						
25–29 vs. <25	1.11	0.91, 1.35	0.96	0.76, 1.20	1.18	0.93, 1.51
30–34 vs. <25	1.20	0.89, 1.63	1.02	0.71, 1.46	1.84	1.35, 2.51
≥35 vs. <25	1.18	0.72, 1.93	0.78	0.43, 1.43	1.40	0.85, 2.31
Linear	1.01	0.98, 1.04	1.01	0.98, 1.05	1.05	1.02, 1.07
Education						
School after high school (yes vs. no)	0.76	0.65, 0.88	0.97	0.81, 1.17	0.87	0.72, 1.06
College degree (yes vs. no)	0.69	0.59, 0.81	0.98	0.81, 1.18	0.86	0.70, 1.04
Smoking						
Past smoker (yes vs. no)	1.33	1.16, 1.52	1.08	0.93, 1.25	0.88	0.76, 1.04
Current smoker (yes vs. no)	2.35	1.90, 2.90	1.98	1.55, 2.52	1.06	0.77, 1.46
Age at menopause (years/10)	0.88	0.78, 0.99	0.91	0.79, 1.04	0.96	0.82, 1.12
Physical functioning construct/10	0.85	0.83, 0.88	0.87	0.84, 0.90	0.91	0.88, 0.94

history of cardiovascular disease (coronary heart disease, stroke, or venous thromboembolism); the 74 percent of clinical trial women and 85 percent of observational study control group women who had not used hormone therapy prior to WHI enrollment and the 79 percent of observational study estrogen-plus-progestin group women who had not used estrogen plus progestin prior to their "current hormone therapy episode" at baseline screening; women aged less than 60 years; women less than 10 years from menopause; and women (24 percent in the clinical trial and 39 percent in the observational study) having fewer than 5 years from menopause during which they did not use hormone therapy. The estimated estrogen-plus-progestin hazard ratios were fairly similar for the clinical trial and observational study in each of these subgroups. It can also be commented that estrogen-plus-progestin hazard ratios in the two cohorts did not differ substantially between women aged less than 60 years at baseline compared with women aged 60 years or more, between women less than 10 years from menopause compared with women 10 or more years from menopause, or between women having less than 5 years from menopause without hormone therapy compared with women having 5 or more years.

Hazard ratio sensitivity to lack of estrogen-plus-progestin adherence

Comparisons so far have focused on the estrogen-plus-progestin randomization group in the clinical trial and on current use of estrogen plus progestin at baseline in the observational study, since the intention-to-treat clinical trial analyses have both reliability and useful interpretation. However, differential adherence patterns between the two cohorts could affect hazard ratio comparisons. Hence, analyses were also carried out with the follow-up period for each woman restricted to the time period within which she continued to adhere to her estrogen-plus-progestin group designation (refer to Materials and Methods). The estimated ratios of estrogen-plus-progestin hazard ratio in the observational study to those in the clinical trial, following control for confounding and time from estrogen-plus-progestin initiation, were 0.86 (95 percent CI: 0.56, 1.30) for coronary heart disease, 0.82 (95 percent CI: 0.50, 1.34) for stroke, and 0.79 (95 percent CI: 0.50, 1.25) for venous thromboembolism in these analyses, rather similar to those given above.

Hazard ratio dependence on preparation and on estrogen and progestin dose

Among the 17,503 baseline estrogen-plus-progestin users in the observational study, 13,565 (78 percent) used CEE, while 1,377 used other estrone sulfate-dominant estrogens, 1,359 used oral estradiol, and 642 used transdermal estradiol. A total of 16,649 (95 percent) of these women used MPA, with 13,065 (75 percent) using a CEE/MPA combination. About 95 percent of the CEE/MPA users were on a daily regimen. Among these women, 11,095 (87 percent) used the standard 0.625-mg/day CEE dose, while 966 women used 0.3 mg/day, and 632 women used a higher

TABLE 5.   Estrogen-plus-progestin hazard ratios as a function of years from initiation of the current episode of estrogen-plus-progestin use in the Women's Health Initiative clinical trial (1994–2002) and observational study (1994–2003)

Time from estrogen-plus-progestin initiation* (years)	Clinical trial			Observational study		
	Cases in estrogen-plus-progestin group (no.)	Hazard ratio	95% confidence interval	Cases in estrogen-plus-progestin group (no.)	Hazard ratio	95% confidence interval
Coronary heart disease						
<2	80	1.68	1.15, 2.45	5	1.12	0.46, 2.74
2–5	80	1.25	0.87, 1.79	27	1.05	0.70, 1.58
>5	28	0.66	0.36, 1.21	126	0.83	0.67, 1.01
Stroke						
<2	43	1.15	0.71, 1.87	7	2.10	0.96, 4.56
2–5	79	1.49	1.02, 2.17	12	0.48	0.24, 0.93
>5	29	0.74	0.39, 1.39	104	0.89	0.71, 1.18
Venous thromboembolism						
<2	73	3.10	1.85, 5.19	7	2.37	1.08, 5.19
2–5	72	1.89	1.24, 2.88	27	1.52	1.01, 2.29
>5	22	1.31	0.64, 2.67	119	1.24	0.99, 1.55

* Time from estrogen-plus-progestin initiation in the clinical trial is time from Women's Heath Initiative enrollment, while time from estrogen-plus-progestin initiation in the observational study is time from Women's Health Initiative enrollment plus the duration of the current episode of estrogen-plus-progestin use at enrollment.

dose. Similarly, 10,188 (80 percent) of these women used 2.5 mg/day of MPA, while 5,440 used a higher daily dose.

To ensure that comparisons of hazard ratios in the clinical trial and observational study were not influenced by the range of estrogen-plus-progestin preparations, dosages, and schedules in the observational study, analyses were carried out restricting the estrogen-plus-progestin group in the observational study to the 12,136 (69 percent of total) women who used (at baseline) the same daily combination of 0.625 mg of CEE and 2.5 mg of MPA studied in the clinical trial. Following this restriction, the estimated ratio of estrogen-plus-progestin hazard ratio in the observational study to that in the clinical trial was 0.85 (95 percent CI: 0.57, 1.28) for coronary heart disease, 0.77 (95 percent CI: 0.48, 1.22) for stroke, and 0.92 (95 percent CI: 0.59, 1.42) for venous thromboembolism, again similar to the estimated ratios previously given.

DISCUSSION

The analyses presented here indicate that adjustment for confounding factors and time from hormone therapy initiation can bring estrogen-plus-progestin hazard ratios from the WHI observational study into fairly close agreement with those from the clinical trial for coronary heart disease and venous thromboembolism and, to a lesser extent, for stroke. The observational study hazard ratios for stroke in these analyses may be subject to some residual confounding, as could also be the case for the other clinical outcomes. For coronary heart disease and venous thromboembolism, hazard ratios decrease with increasing time from estrogen-

plus-progestin initiation (6, 7, 30), so that overall hazard ratios in the observational study, which are heavily weighted by longer term estrogen-plus-progestin use, are less than those in the clinical trial, which are heavily weighted by shorter term use. The principal comparisons are based on intention-to-treat analyses in the clinical trial and baseline estrogen-plus-progestin use in the observational study. However, the extent of agreement between clinical trial and observational study hazard ratios for these clinical outcomes was evidently not much affected by estrogen-plus-progestin group adherence during WHI follow-up, by dependencies of hazard ratios on baseline characteristics or exposures, or by variations in estrogen-plus-progestin preparations, dosages, or schedules in the observational study.

These analyses reinforce early elevations in cardiovascular disease risk among estrogen-plus-progestin users. For coronary heart disease, such early elevation is consistent with the Heart Estrogen/progestin Replacement Study (HERS) secondary prevention trial of the same CEE/MPA regimen (34), which found no overall effect on coronary heart disease risk over an average 4.1-year follow-up period. Other hormone therapy trials of secondary prevention of atherosclerosis progression (35–40) have mostly reported neutral or unfavorable effects over fairly short follow-up periods. For stroke, the Heart Estrogen/progestin Replacement Study (41) reported a nonsignificant elevation in risk, while the secondary prevention Women's Estrogen and Stroke Trial of estradiol (42) found no overall effect but with an indication of elevation in the first months of use. For venous thromboembolism, the Heart Estrogen/progestin Replacement Study (43) found a substantial early elevation

in risk. Moreover, a small secondary prevention trial of estradiol plus norethisterone was stopped early on the basis of an excess of venous thromboembolism events (44).

It is interesting to consider other observational study findings in relation to the WHI clinical trial. For coronary heart disease, several observational studies among healthy postmenopausal women included confounding control efforts and reported results as a function of hormone therapy duration. Some (45–48), but not all (49, 50), provided hints of early coronary heart disease risk elevation, but most had limited precision for estimating early hormone therapy effects. In addition, few reported associations separately for estrogen plus progestin and for estrogen alone, and there is now randomized controlled trial evidence (51) to suggest that the coronary heart disease implications of CEE alone are more favorable than for combined CEE and MPA. Similarly, an early stroke elevation among hormone therapy users has been reported in some (52), but not all (53), recent observational studies.

The Nurses' Health Study followed women over the age interval when postmenopausal hormone therapy would likely be initiated, and it reported coronary heart disease results separately for estrogen and for estrogen plus progestin but did not find an early elevation in risk among healthy postmenopausal women (49). However, analyses of Nurses' Health Study data to date use only a snapshot of current hormone therapy use at biennial contacts, so that an estrogen-plus-progestin user would be classified as a nonuser for her first year of use on average and would be classified as a nonuser permanently if estrogen-plus-progestin use started and stopped within a biennial period prior to hormone therapy status ascertainment. As an exercise, we conducted a simulation study, wherein the estrogen-plus-progestin group assignment in the WHI clinical trial was randomly contaminated in a similar fashion and found that evidence for an early elevation in coronary heart disease risk and evidence for a time trend in the coronary heart disease hazard ratio typically disappeared under these circumstances.

There are implications of these analyses for the design and analysis of observational studies when the hazard ratio of interest varies with time. For example, a cohort comprising new initiators of the study exposure (12) can be expected to yield meaningful average hazard ratio estimates over various periods of time from exposure initiation, even if a proportional hazards assumption is imposed. However, studies like the WHI observational study that enrolled participants having various exposure durations will need to use nonstandard data analysis methods to assess exposure effects. Specifically, such studies may be able to provide hazard ratio estimates as a function of time from exposure initiation, from which average hazard ratios over certain time from initiation periods can be calculated.

In summary, these analyses aid in our understanding of sources of bias in observational studies, and they indicate that the apparent discrepancy between the clinical trial and observational studies, such as the WHI observational study, may be substantially explained by classical confounding and differences in the distributions of time from estrogen-plus-progestin initiation. The inability of those factors to provide a full explanation for differences between stroke hazard ratios reinforces the importance of randomized controlled trial evidence, especially when public health implications are great.

University of Arizona, Tucson/Phoenix, Arizona: Tamsen Bassford; *University at Buffalo, Buffalo, New York*: Jean Wactawski-Wende; *University of California at Davis, Sacramento, California*: John Robbins; *University of California at Irvine, Orange, California*: Allan Hubbell; *University of California at Los Angeles, Los Angeles, California*: Howard Judd; *University of California at San Diego, LaJolla/Chula Vista, California*: Robert D. Langer; *University of Cincinnati, Cincinnati, Ohio*: Margery Gass; *University of Florida, Gainesville/Jacksonville, Florida*: Marian Limacher; *University of Hawaii, Honolulu, Hawaii*: David Curb; *University of Iowa, Iowa City/Davenport, Iowa*: Robert Wallace; *University of Massachusetts/Fallon Clinic, Worcester, Massachusetts*: Judith Ockene; *University of Medicine and Dentistry of New Jersey, Newark, New Jersey*: Norman Lasser; *University of Miami, Miami, Florida*: Mary Jo O'Sullivan; *University of Minnesota, Minneapolis, Minnesota*: Karen Margolis; *University of Nevada, Reno, Nevada*: Robert Brunner; *University of North Carolina, Chapel Hill, North Carolina*: Gerardo Heiss; *University of Pittsburgh, Pittsburgh, Pennsylvania*: Lewis Kuller; *University of Tennessee, Memphis, Tennessee*: Karen C. Johnson; *University of Texas Health Science Center, San Antonio, Texas*: Robert Brzyski; *University of Wisconsin, Madison, Wisconsin*: Gloria Sarto; *Wake Forest University School of Medicine, Winston-Salem, North Carolina*: Denise Bonds; and *Wayne State University School of Medicine/Hutzel Hospital, Detroit, Michigan*: Susan Hendrix.

Conflict of interest: none declared.

REFERENCES

1. The truth about hormones. (Cover story). Time magazine July 22, 2002.
2. The end of the age of estrogen. (Cover story). Newsweek magazine July 22, 2002.
3. Stampfer M, Colditz G. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. Prev Med 1991;20:47–63.
4. Grady D, Rubin SB, Pettiti DB, et al. Hormone therapy to prevent disease and prolong life in postmenopausal women. Ann Intern Med 1992;117:1016–37.
5. Barrett-Conner E, Grady D. Hormone replacement therapy, heart disease, and other considerations. Annu Rev Public Health 1998;19:55–72.
6. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. Principal results from the Women's Health Initiative randomized controlled trial. Writing Group for the Women's Health Initiative Investigators. JAMA 2002; 288:321–33.
7. Manson JE, Hsia J, Johnson KC, et al. Estrogen plus progestin and the risk of coronary heart disease. N Engl J Med 2003; 349:523–34.
8. Goodman D, Goldzieher J, Ayala C. Critique of the report from the Writing Group of the WHI. Menopausal Med 2003;10:1–4.
9. Creasman WT, Hoel D, DiSaia PJ. WHI: now that the dust has settled: a commentary. Am J Obstet Gynecol 2003;189: 621–6.
10. Naftolin F, Taylor HS, Karas R, et al. The Women's Health Initiative could not have detected cardioprotective effects of starting hormone therapy during the menopausal transition. Fertil Steril 2004;81:1498–501.
11. Grodstein F, Clarkson TB, Manson JE. Understanding the divergent data on postmenopausal hormone therapy. N Engl J Med 2003;348:645–50.
12. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol 2003;158:915–20.
13. Michels KB, Manson JE. Postmenopausal hormone therapy: a reversal of fortune. Circulation 2003;107:1830–3.
14. Humphrey LL, Chan BKS, Sox HC. Postmenopausal hormone replacement therapy and the primary prevention of cardiovascular disease. Ann Intern Med 2002;137:273–84.
15. Stampfer MJ, Colditz GA. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. Int J Epidemiol 2004;33:445–53.
16. Stampfer MJ. Commentary: hormones and heart disease: do trials and observational studies address different questions? Int J Epidemiol 2004;33:454–5.
17. Vandenbroucke JP. Commentary: the HRT story: vindication of old epidemiological theory. Int J Epidemiol 2004;33:456–7.
18. Barrett-Conner E. Commentary: observation versus intervention—what's different? Int J Epidemiol 2004;33:457–9.
19. Kuller LH. Commentary: hazards of studying women: the oestrogen oestrogen/progesterone dilemma. Int J Epidemiol 2004;33:459–60.
20. Petitti D. Commentary: hormone replacement therapy and coronary heart disease: four lessons. Int J Epidemiol 2004; 33:461–3.
21. Lawlor DA, Smith GD, Ebrahim S. Commentary: the hormone replacement–coronary heart disease conundrum: is this the death of observational epidemiology? Int J Epidemiol 2004; 33:464–7.
22. Herrington DM, Howard TD. From presumed benefit to potential harm—hormone therapy and heart disease. N Engl J Med 2003;349:519–21.
23. Whittemore AS, McGuire V. Observational studies and randomized studies of hormone replacement therapy: what can we learn from them? Epidemiology 2003;14:8–10.
24. Hays J, Hunt JR, Hubbell FA, et al. The Women's Health Initiative recruitment methods and results. Ann Epidemiol 2003;13(suppl):S18–77.
25. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials 1998;19:61–109.
26. Langer RD, White E, Lewis CE, et al. The Women's Health Initiative observational study: baseline characteristics of participants and reliability of baseline measures. Ann Epidemiol 2003;13(suppl):S107–21.
27. Effects of hormone replacement therapy on endometrial histology in postmenopausal women: the Postmenopausal Estrogen/Progestin Interventions (PEPI) trial. The Writing Group of the PEPI Trial. JAMA 1996;275:370–5.
28. Curb D, McTiernan A, Heckbert SR, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. Ann Epidemiol 2003;13(suppl):S122–8.
29. Wassertheil-Smoller S, Hendrix S, Limacher M, et al. Effect of estrogen plus progestin on stroke in postmenopausal women: Women's Health Initiative. JAMA 2003;289:2673–84.
30. Cushman M, Kuller LH, Prentice R, et al. Estrogen plus progestin and risk of venous thrombosis. JAMA 2004;292: 1573–80.
31. Cox DR. Regression models and life tables (with discussion). J R Stat Soc (B) 1972;34:187–220.

32. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. Health Econ 1993;2:217–27.
33. Wang CY, Hsu L, Feng ZD, et al. Regression calibration in failure time regression with surrogate variables. Biometrics 1997;53:131–45.
34. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. JAMA 1998;280:605–13.
35. Hodis HN, Mack WJ, Lobo RA, et al. Estrogen in the prevention of atherosclerosis. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2001;135:939–53.
36. Herrington DM, Reboussin DM, Brosnihan KB, et al. Effects of estrogen replacement on the progression of coronary-artery atherosclerosis. N Engl J Med 2000;343:522–9.
37. Clarke SC, Kelleher J, Lloyd-Jones H, et al. A study of hormone replacement therapy in postmenopausal women with ischaemic heart disease: the Papworth HRT atherosclerosis study. BJOG 2002;109:1056–62.
38. Waters DD, Alderman EL, Hsia J, et al. Effects of hormone replacement therapy and antioxidant vitamin supplements on coronary atherosclerosis in postmenopausal women: a randomized controlled trial. JAMA 2002;288:2432–40.
39. Oestrogen therapy for prevention of reinfarction in postmenopausal women: a randomized placebo controlled trial. ESPRIT Team. Lancet 2002;360:2001–8.
40. Hodis HN, Mack WJ, Azen SP, et al. Hormone therapy and the progression of coronary-artery atherosclerosis in postmenopausal women. N Engl J Med 2003;349:535–45.
41. Simon JA, Hsia J, Cauley JA, et al. Postmenopausal hormone therapy and risk of stroke. The Heart and Estrogen/progestin Replacement Study. Circulation 2001;103:638–42.
42. Viscoli CM, Brass LM, Kernan WN, et al. A clinical trial of estrogen-replacement therapy after ischemic stroke. N Engl J Med 2001;345:1243–9.
43. Grady D, Wenger NK, Herrington D, et al. Postmenopausal hormone therapy increases risk of venous thrombolic disease. The Heart and Estrogen/progestin Replacement Study. Ann Intern Med 2000;132:689–96.
44. Hoibraaten E, Qvigstad E, Arnesen H, et al. Increased risk of recurrent venous thromboembolism during hormone replacement therapy. Thromb Haemost 2000;84:962–7.
45. Rosenberg L, Slone D, Shapiro S, et al. Noncontraceptive estrogens and myocardial infarction in young women. JAMA 1980;244:339–42.
46. Hernandez Avila M, Walker AM, Jick H. Use of replacement estrogens and the risk of myocardial infarction. Epidemiology 1990;1:128–33.
47. Criqui MH, Suarez L, Barrett-Conner E, et al. Postmenopausal estrogen use and mortality. Results from a prospective study in a defined, homogeneous community. Am J Epidemiol 1988;128:606–14.
48. Sidney S, Petitti DB, Quesenberry CP Jr. Myocardial infarction and the use of estrogen and estrogen-progestin in postmenopausal women. Ann Intern Med 1997;127:501–8.
49. Grodstein F, Manson JE, Colditz GA, et al. A prospective, observational study of postmenopausal hormone therapy and primary presentation of cardiovascular disease. Ann Intern Med 2000;133:933–41.
50. Heckbert SR, Weiss NS, Koepsell TD, et al. Duration of estrogen replacement therapy in relation to risk of incident myocardial infarction in postmenopausal women. Arch Intern Med 1997;157:1330–6.
51. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. Women's Health Initiative Steering Committee. JAMA 2004;291:1701–12.
52. Lemaitre RN, Heckbert SR, Psaty BM, et al. Hormone replacement therapy and associated risk of stroke in postmenopausal women. Arch Intern Med 2002;162:1954–60.
53. Pettiti DB, Sidney S, Quesenberry CP Jr, et al. Ischemic stroke and use of estrogen and estrogen/progestin as hormone replacement therapy. Stroke 1998;29:23–8.