Title

Dual-Outcome Intention-to-Treat Analyses in the Women's Health Initiative Randomized Controlled Hormone Therapy Trials.

Journal

American journal of epidemiology, 189(9)

ISSN

0002-9262

Authors

Prentice, Ross L
Aragaki, Aaron K
Chlebowski, Rowan T
et al.

Publication Date

2020-09-01

DOI

10.1093/aje/kwaa033

Peer reviewed

American Journal of Epidemiology

Vol. 189, No. 9
DOI: 10.1093/aje/kwaa033
Advance Access publication:
April 21, 2020

Practice of Epidemiology

Dual-Outcome Intention-to-Treat Analyses in the Women's Health Initiative Randomized Controlled Hormone Therapy Trials

Ross L. Prentice*, Aaron K. Aragaki, Rowan T. Chlebowski, Shanshan Zhao, Garnet L. Anderson, Jacques E. Rossouw, Robert Wallace, Hailey Banack, Aladdin H. Shadyab, Lihong Qi, Beverly M. Snively, Margery Gass, and JoAnn E. Manson

* Correspondence to Dr. Ross L. Prentice, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, P.O. Box 19024, Seattle, WA 98109-1024 (e-mail: rprentic@whi.org).

Initially submitted July 15, 2019; accepted for publication January 16, 2020.

Dual-outcome intention-to-treat hazard rate analyses have potential to complement single-outcome analyses for the evaluation of treatments or exposures in relation to multivariate time-to-response outcomes. Here we consider pairs formed from important clinical outcomes to obtain further insight into influences of menopausal hormone therapy on chronic disease. As part of the Women's Health Initiative, randomized, placebo-controlled hormone therapy trials of conjugated equine estrogens (CEE) among posthysterectomy participants and of these same estrogens plus medroxyprogesterone acetate (MPA) among participants with an intact uterus were carried out at 40 US clinical centers (1993–2016). These data provide the context for analyses covering the trial intervention periods and a nearly 20-year (median) cumulative duration of follow-up. The rates of multiple outcome pairs were significantly influenced by hormone therapy, especially over cumulative follow-up, providing potential clinical and mechanistic insights. For example, among women randomized to either regimen, hazard ratios for pairs defined by fracture during intervention followed by death from any cause were reduced and hazard ratios for pairs defined by gallbladder disease followed by death were increased, though these findings may primarily reflect single-outcome associations. In comparison, hazard ratios for diabetes followed by death were reduced with CEE but not with CEE + MPA, and those for hypertension followed by death were increased with CEE + MPA but not with CEE.

cancer; cardiovascular disease; Cox model; diabetes; dual outcomes; fractures; hazard ratio; menopausal hormone therapy

Abbreviations: CEE, conjugated equine estrogens; CHD, coronary heart disease; CI, confidence interval; MPA, medroxyprogesterone acetate; HR, hazard ratio; WHI, Women's Health Initiative.

Editor's note: *An invited commentary on this article appears on page 982, and the authors' response appears on page 985.*

Results from the randomized, double-blind, placebo-controlled hormone therapy trials carried out as part of the Women's Health Initiative (WHI) led to a substantial reduction in the use of menopausal hormone therapy in the United States (1–3) and elsewhere, with reductions in rates of both initiation and continuation (4). Reduction in the use of conjugated equine estrogens (CEE) in conjunction with medroxyprogesterone acetate (MPA) has been linked to subsequent reductions in breast cancer incidence (5) and medical care costs (6). The 2 trials tested the most commonly used hormone therapy preparations at the time, consisting of CEE (0.625 mg daily) among women who were posthysterectomy and CEE + MPA (2.5 mg daily) among women with an intact uterus. A complex pattern of risks and benefits emerged from comprehensive univariate analyses of health outcomes through September 30, 2010, with a median follow-up period of 13 years (7). Corresponding comparisons for all-cause and cause-specific mortality through December 31, 2014, with cumulative follow-up of

18 years, were not significantly different from the null for either preparation, though some results appeared to vary by age group (8).

Multivariate outcome data provide information on the relationship between treatments or exposures and 2 or more outcomes that occur jointly. With outcomes that are rare within study follow-up periods, most multivariate outcome treatment information resides in dual-outcome analyses. The bivariate survival function for such outcome pairs is fully determined by its single- and dual-outcome hazard functions, implying that dual-outcome hazard rate analyses will complement univariate hazard rate analyses rather comprehensively, in many applications of cohort or clinical trial data.

We recently proposed novel methods for dual-outcome hazard rate analysis (9, 10). Here we apply these methods to major clinical outcomes in the WHI Hormone Therapy Trials, to obtain further insight into the health benefits and risks of these important formulations.

METHODS

Study design and conduct

Details on the design and conduct of the WHI Hormone Therapy Trials have been published elsewhere (11–13). Briefly, 27,347 postmenopausal women aged 50–79 years were recruited at 40 US clinical centers during 1993–1998. Of these, 10,739 women with prior hysterectomy were randomized to receive oral CEE at 0.625 mg/day ($n = 5,310$) or placebo ($n = 5,429$), and 16,608 women with a uterus were randomized to receive CEE + MPA at 2.5 mg/day ($n = 8,506$) or placebo ($n = 8,102$). The primary efficacy and safety outcomes in each trial were coronary heart disease (CHD) and invasive breast cancer, respectively. These outcomes, along with stroke, pulmonary embolism, colorectal cancer, endometrial cancer, hip fracture, and death from any other cause, defined a univariate, time-to-first-outcome "global index" used in trial monitoring. Detailed intervention–versus–placebo-group contrasts for broader cardiovascular disease, cancer, and fracture outcomes, as well as for incidence of diabetes, gallbladder disease, and hypertension, have also shown a relationship with one or both of the hormone therapy formulations (7, 8, 14). These analyses have focused on randomization group in relation to clinical outcomes individually.

All WHI Hormone Therapy Trial participants provided written informed consent for participation during the intervention period, from enrollment in 1993–1998 through March 31, 2005. The CEE + MPA treatment was stopped early in July 2002 following 5.6 years (median) of intervention on the basis of an observed elevation in breast cancer incidence and health risks that exceeded health benefits (12, 15), while the CEE treatment was also stopped early in February 2004 following 7.2 years (median) of intervention, partly on the basis of an elevation in stroke risk of a magnitude similar to that observed for CEE + MPA (13, 15). Additional consent was obtained for postintervention follow-up through September 30, 2010, and over an open-ended subsequent period, with over 80% of surviving participants consenting on each occasion. Nonfatal and fatal outcomes (the latter including periodic National Death Index matching) occurring through December 31, 2016, are included here, resulting in a median of 19.4 years of cumulative follow-up in each trial. Cumulative follow-up is defined here as time from randomization to the end of the follow-up period for each participant. After the trial interventions were stopped and treatment assignments were unmasked, fewer than 4% of participants reported personal postintervention use of systemic menopausal hormones.

Statistical analysis

Multivariate time-to-event data have potential for additional elucidation of effects of treatment on clinical outcomes through the patterns of occurrence for 2 or more outcomes jointly. With infrequent outcome events, most such potential resides in paired-outcome analyses. For example, one can use the bivariate response data to ask questions such as whether participants experience an altered risk of developing hypertension and experiencing CHD during trial follow-up, if they are assigned to active hormone treatment, or whether the risk of invasive breast cancer followed by death from any cause is associated with the hormone therapy preparations studied. Furthermore, comparisons of dual-outcome hazard rates according to which of 2 outcomes occurred first, or according to the dual-outcome hazard ratio dependence on time to first event compared with subsequent time to second event, may provide mechanistic insight.

Intervention-phase data analyses included all randomized participants with censoring at the end of the intervention periods ending on July 7, 2002, and February 29, 2004, for the CEE + MPA trial and CEE-alone trial, respectively, or at earlier loss to follow-up or death. Cumulative-phase analyses also included all participants with censoring on December 31, 2016, or earlier at the end of the woman's consent period (for outcomes that include nonfatal events), or at earlier loss to follow-up or death. According to usual convention, outcome events were assumed to precede censoring times if there were tied times.

Dual-outcome intention-to-treat hazard ratios were estimated using a multivariate generalization of the Cox regression method that allows simultaneous analysis of randomization assignment in relation to univariate (single-outcome) and bivariate (dual-outcome) hazard rates (9, 10). A brief account of the dual-outcome hazard rate estimation methodology is given in Web Appendix 1 (available at https://academic.oup.com/aje).

The new methods allow dual-outcome "baseline" hazard rate stratification that may be time-dependent. The same stratification as in our previous univariate outcome trial analyses (7, 8) was used here for dual-outcome hazard rates. Specifically, hazard rates were stratified on baseline age (50–54, 55–59, 60–69, or 70–79 years), prior disease (if applicable), and randomization status in the WHI Dietary Modification Trial (intervention, comparison, not randomized). Analyses were conducted with participants contributing follow-up time for a specific outcome until the end of the study period under consideration, the date of the first relevant

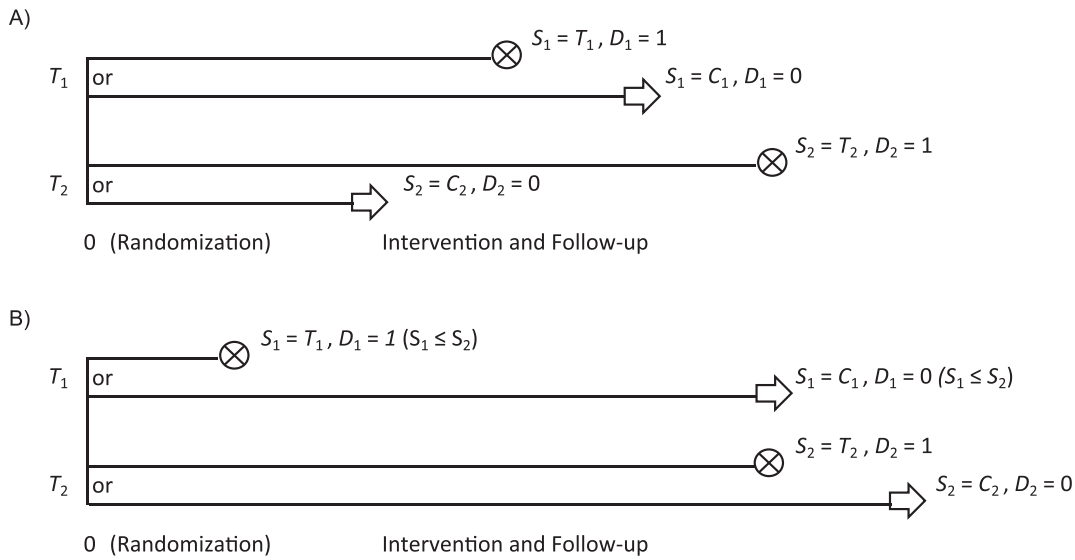

Figure 1. Dual-outcome times to clinical event or censoring in the Women's Health Initiative, 1993–2016. Dual outcomes are shown as T_1 and T_2, with corresponding potential censoring times C_1 and C_2. Observed follow-up times are shown as S_1, the minimum of T_1 and C_1, and S_2, the minimum of T_2 and C_2. Censoring indicator variables D_1 and D_2 take the value 1 if $S_1 = T_1$ and if $S_2 = T_2$, respectively. A) Outcome variables (T_1 and T_2) that each may include a nonfatal component; B) an outcome variable (T_1) that may include a nonfatal component, in conjunction with death from any cause (T_2). Note that dual-outcome events lie on or above the main diagonal ($t_1 = t_2$) in panel B. Dual-outcome hazard rates at follow-up times (t_1, t_2) implicitly assume continued participant survival at $T_1 = t_1$ and $T_2 = t_2$ for any (t_1, t_2).

clinical event, death, or loss to follow-up/withdrawal from active follow-up, whichever came first. Figure 1 presents a schematic outline of dual-outcome event and censoring times.

Since intervention influences may be less plausible for some outcomes at times well beyond the end of the trial intervention periods, cumulative-phase analyses of clinical outcomes in conjunction with all-cause mortality were also conducted with times to clinical outcomes additionally censored at the end of the WHI program intervention phase (March 31, 2005).

For choice of clinical outcomes, we considered the primary, global index, secondary, and self-reported outcomes listed in Figure 2 of the paper by Manson et al. (7), adding hypertension incidence, which was subsequently reported (14). As in previous reports (7, 14), participants with self-reported prior outcomes at baseline were excluded from analyses that included the particular self-reported outcome. Since the dual-outcome analyses have efficiency that depends primarily on the number of observed dual-outcome occurrences during follow-up, we restricted the set of clinical outcomes to those having at least 150 participants with incident events during the intervention phase of one or both of the hormone therapy trials. This restriction yielded 10 major clinical outcome categories: CHD (nonfatal myocardial infarction plus coronary death), coronary artery bypass graft/percutaneous coronary intervention, stroke (ischemic plus hemorrhagic), venous thromboembolic event, invasive breast cancer, other cancer (excluding nonmelanoma skin cancer), total fractures, (treated) diabetes, gallbladder disease, and (treated) hypertension. These outcomes were also considered in combination with death from any cause. Fracture outcomes were adjudicated only for fractures of the hip beyond the trial intervention period, so fracture outcomes over the cumulative course of follow-up were restricted to hip fractures.

Dual-outcome hazard ratios and 95% confidence intervals were calculated for each outcome pair. Additionally, by exercising the time-dependent regression features for modeled log hazard ratios, we estimated separate hazard ratios for dual outcomes having nonfatal components according to which of the 2 events occurred first. All such analyses retain an intention-to-treat interpretation.

Dual-outcome hazard ratios are displayed only if 20 or more women experienced the dual events within the designated follow-up periods. Simulation studies (9) with a binary covariate (probability 0.5 for each value) under a model with proportional single- and dual-outcome hazard ratios gave estimation results that did not exhibit hazard ratio bias, and corresponding confidence interval coverage rates were close to nominal levels, even when the expected number of dual outcomes was as few as 20. Of course, asymptotic distributional approximations can be expected to be more accurate when numbers of dual outcomes are still larger.

All statistical tests were 2-sided, and nominal P values of 0.05 or less were regarded as significant. P values should be interpreted cautiously in view of the multiple tests conducted.

RESULTS

Baseline demographic and disease characteristics were well balanced between randomization groups in each trial. Compared with women in the CEE + MPA trial, women

in the CEE trial were more ethnically diverse, had a less favorable cardiovascular disease risk profile, and frequently had undergone bilateral oophorectomy in addition to hysterectomy (Table 1).

Web Figures 1 and 2 show numbers of dual outcomes by randomization group, along with the estimated hazard ratios and 95% confidence intervals for pairs of clinical outcomes, and for pairs formed from these outcomes followed by death from any cause, during the intervention phases of the CEE and CEE + MPA trials, respectively. The dual-outcome intention-to-treat hazard ratios (and 95% confidence intervals) are also displayed graphically through "platter plots" above the main diagonal in these figures. Platters with an open interior indicate significant dual-outcome contrasts between randomization groups.

Web Figure 1 shows that few women experienced dual clinical outcomes during the intervention phase of the CEE trial. The dual-outcome hazard ratios for fractures and CHD, and for fractures and cancers other than breast cancer, were reduced in the intervention group. In addition, rates of gallbladder disease with CHD and gallbladder disease with hypertension were increased among intervention women, the latter having a hazard ratio of 1.96 (95% confidence interval (CI): 1.46, 2.64), with 127 and 69 participants having dual outcomes in the intervention and placebo groups, respectively.

Web Figure 2 shows, similarly, that few women experienced dual clinical outcomes during the intervention phase of the CEE + MPA trial. While significant intervention-related risk reductions were not evident, risk elevations were observed for hypertension in conjunction with multiple clinical outcomes, including venous thromboembolism, other cancer, gallbladder disease, and death from all causes—the latter two with hazard ratios of 2.30 (95% CI: 1.62, 3.25) and 1.77 (95% CI: 1.04, 3.00), respectively. There were also nominally significant elevations in stroke and CHD, and gallbladder disease and CHD, in the intervention group, though the numbers of participants with dual outcomes were small.

Additionally, updated univariate hazard ratios (and 95% confidence intervals) were simultaneously estimated over the follow-up periods defined here, and results are provided in Web Figure 3 for both trials. These findings differed little from those in our previous reports (7, 14).

Web Figures 4 and 5, with a median of 19.4 years of cumulative follow-up, show substantially larger numbers of women experiencing dual outcomes. For CEE (Web Figure 4), there are reductions among intervention participants in the dual outcomes of diabetes with either stroke or total mortality. Comparatively stronger risk elevations are observed for gallbladder disease in conjunction with each of CHD, coronary artery bypass graft/percutaneous coronary intervention, hypertension, breast cancer, diabetes, and total mortality. For example, the hazard ratio for gallbladder disease and breast cancer was 2.51 (95% CI: 1.20, 5.24), and that for gallbladder disease and hypertension was 1.74 (95% CI: 1.34, 2.25).

Somewhat similar patterns were observed for CEE + MPA over cumulative follow-up (Web Figure 5). Risk reductions among intervention group women were seen for hip fracture and other cancer, and for hip fracture and total mortality. In addition, risk elevations were observed in the intervention group for gallbladder disease in conjunction with each of breast cancer, stroke, coronary artery bypass graft/percutaneous coronary intervention, other cancer, hypertension, and death from any cause. There were further nominally significant elevations in hypertension and either breast cancer or total mortality, and in the dual outcomes of total mortality with either stroke or breast cancer.

Web Figures 6A and 6B break out the CEE hazard ratios of Web Figure 4 according to which of the 2 outcomes occurred first. Dual-outcome hazard ratio estimates for gallbladder disease and breast cancer, and for gallbladder disease and hip fracture, were relatively higher when the gallbladder disease preceded the other outcome. Additionally, the dual-outcome hazard ratio for diabetes and stroke was lower (P for interaction = 0.02) when diabetes preceded stroke (hazard ratio (HR) = 0.47, 95% CI: 0.28, 0.77) as compared with following stroke (HR = 1.42, 95% CI: 0.65, 3.11). Likewise, in the CEE + MPA trial (Web Figures 7A and 7B), hazard ratios were increased for gallbladder disease preceding breast cancer, stroke, coronary artery bypass graft/percutaneous coronary intervention, and other cancer, but only for venous thromboembolism when the other outcome preceded gallbladder disease. In addition, hazard ratios were increased significantly only when hypertension preceded breast cancer and other cancer. On the other hand, hazard ratios were similarly elevated for the dual outcomes of gallbladder disease and hypertension with either outcome pair ordering in both hormone therapy trials.

Figure 2 repeats the analyses of the final row of Web Figures 6B and 7B, but with censoring of clinical outcomes (including total fracture rather than just hip fracture) at the end of the program intervention period. Even though intervention associations with all-cause mortality were nonsignificant overall, reductions in total fracture risk during the intervention period followed by death during cumulative follow-up were evident in both trials (HR = 0.68 (95% CI: 0.57, 0.80) for CEE and HR = 0.78 (95% CI: 0.67, 0.91) for CEE + MPA); and elevations in risk of gallbladder disease in the intervention period followed by death during cumulative follow-up were evident in both trials (HR = 1.39 (95% CI: 1.09, 1.78) for CEE and HR = 1.74 (95% CI: 1.36, 2.22) for CEE + MPA). In contrast, reductions in hazard ratios for diabetes followed by death with CEE, and increases in hypertension followed by death with CEE + MPA, were not evident in the companion trial.

DISCUSSION

Dual-outcome hazard ratio analysis methods (9, 10) provide potential to augment previously presented single-outcome hazard ratio analyses for menopausal hormone therapy (7, 8, 14). Multivariate time-to-response outcome methods have typically modeled hazard rates for specific outcomes conditional on the preceding history of the correlated set of outcomes, using a "counting process intensity" approach (16, 17). By avoiding conditioning on the prior history of other outcomes, the present analyses estimate hazard ratios that have a useful population-averaged interpretation,

Table 1. Baseline Characteristics of Participants in the Women's Health Initiative Trials of Postmenopausal Hormone Therapy, 1993–2016[a]

Characteristic	Trial Arm and Distribution of Participants[b]							
	CEE				CEE + MPA			
	Active (*n* = 5,310)		Placebo (*n* = 5,429)		Active (*n* = 8,506)		Placebo (*n* = 8,102)	
	No.	%	No.	%	No.	%	No.	%
Age at screening, years[c]	63.6 (7.3)		63.6 (7.3)		63.2 (7.1)		63.3 (7.1)	
Age group at screening, years[d]								
50–59	1,639	30.9	1,674	30.8	2,837	33.4	2,683	33.1
60–69	2,386	44.9	2,465	45.4	3,854	45.3	3,655	45.1
70–79	1,285	24.2	1,290	23.8	1,815	21.3	1,764	21.8
Race/ethnicity								
White	4,009	75.5	4,075	75.1	7,141	84.0	6,805	84.0
Black	781	14.7	835	15.4	548	6.4	574	7.1
Hispanic	319	6.0	332	6.1	471	5.5	415	5.1
American Indian	41	0.8	34	0.6	25	0.3	30	0.4
Asian/Pacific Islander	86	1.6	78	1.4	194	2.3	169	2.1
Unknown	74	1.4	75	1.4	127	1.5	109	1.3
Menopausal hormone use								
Never use	2,769	52.2	2,769	51.0	6,277	73.8	6,022	74.4
Past use	1,871	35.2	1,947	35.9	1,671	19.7	1,587	19.6
Current use[e]	669	12.6	709	13.1	554	6.5	490	6.1
Vasomotor symptoms								
None	2,962	56.4	3,004	56.0	5,162	61.3	4,928	61.5
Mild	1,377	26.2	1,441	26.9	2,190	26.0	2,115	26.4
Moderate or severe	913	17.4	917	17.1	1,072	12.7	974	12.1
Body mass index[f,g]	29.2 (25.7–33.7)		29.2 (25.7–33.5)		27.5 (24.2–31.7)		27.5 (24.3–31.7)	
Systolic BP, mm Hg[c]	130.4 (17.5)		130.2 (17.6)		127.6 (17.6)		127.8 (17.5)	
Diastolic BP, mm Hg[c]	76.5 (9.2)		76.5 (9.4)		75.6 (9.1)		75.8(9.1)	
Smoking status								
Never smoker	2,723	51.9	2,705	50.4	4,178	49.6	3,999	50.0
Past smoker	1,986	37.8	2,090	38.9	3,362	39.9	3,157	39.5
Current smoker	542	10.3	571	10.6	880	10.5	838	10.5
Bilateral oophorectomy	1,938	39.5	2,111	42.0	29	0.3	24	0.3
Medical treatment								
Treated for diabetes	410	7.7	412	7.6	374	4.4	360	4.4
Treated for hypertension or BP ≥140/90 mm Hg	2,651	53.3	2,647	52.6	3,377	43.2	3,283	42.7
Elevated cholesterol levels requiring medication	763	14.4	829	15.3	1,018	12.0	1,027	12.7
Statin use at baseline	397	7.5	430	7.9	580	6.8	535	6.6
Aspirin use (≥80 mg/day at baseline)	1,050	19.8	1,081	19.9	1,652	19.4	1,654	20.4

Table continues

Table 1. Continued

Characteristic	Trial Arm and Distribution of Participants[b]							
	CEE				CEE + MPA			
	Active (*n* = 5,310)		Placebo (*n* = 5,429)		Active (*n* = 8,506)		Placebo (*n* = 8,102)	
	No.	%	No.	%	No.	%	No.	%
Medical history								
Myocardial infarction	165	3.1	173	3.2	139	1.6	157	1.9
Angina	402	7.6	388	7.2	318	3.8	331	4.1
CABG or PCI	120	2.3	114	2.1	95	1.1	120	1.5
Stroke	76	1.4	92	1.7	61	0.7	77	1.0
Deep vein thrombosis or pulmonary embolism	87	1.6	84	1.5	79	0.9	62	0.8
Family history of breast cancer[h]	892	17.9	870	17.1	1,286	16.0	1,175	15.3
Education more than high school diploma/GED	3,488	66.3	3,678	68.3	6,272	74.1	5,899	73.3
Family income ≥$50,000/year	1,148	22.9	1,167	22.9	2,447	30.4	2,401	31.4

Abbreviations: BP, blood pressure; CABG, coronary artery bypass graft; CEE, conjugated equine estrogens; GED, General Equivalency Diploma; MPA, medroxyprogesterone acetate; PCI, percutaneous coronary intervention; WHI, Women's Health Initiative.

[a] Participants were postmenopausal and aged 50–79 years when enrolled at 40 US clinical centers during 1993–1998. Data were missing for some participants.

[b] Values are expressed as number and percentage unless otherwise indicated.

[c] Values are expressed as mean (standard deviation).

[d] Percentages may not add to 100 because of rounding.

[e] Required a 3-month washout period prior to randomization.

[f] Weight (kg)/height (m)^2.

[g] Values are expressed as median (interquartile range).

[h] In a mother, sister, daughter, or grandmother.

while also permitting censoring rates to differ among the outcomes studied and while retaining an intention-to-treat interpretation. The dual-outcome intention-to-treat analyses require participants at risk for a dual outcome to be representative of the study population, given covariates. This requirement is somewhat stronger than the independent censoring assumption needed for the counting process intensity modeling approach. More generally, the 2 modeling approaches can be regarded as largely complementary, through their focus on the estimation of conditional and population-averaged associations, respectively. The dual-outcome analyses require specialized statistical software. Web Appendix 1 describes key aspects of the computation and provides a simulated data set (Web Table 1). R code (R Foundation for Statistical Computing, Vienna, Austria) with which to apply single- and dual-outcome hazard ratio estimation methods, with illustrative application to the simulated data set, is provided in Web Appendix 2.

Dual-outcome hazard ratios reduce to the product of corresponding single-outcome hazard ratios for outcomes that arise from statistically independent processes, given randomization assignment and other modeled covariates. More generally, however, the joint distribution of dual time-to-response outcomes is not determined by their univariate hazard rates, and there is then no simple relationship between dual-outcome hazard ratios and corresponding single-outcome hazard ratios. Rather, the dual-outcome hazard function fully complements the single-outcome hazard functions in the sense that together they specify the joint survivor function for the pair of time-to-response outcomes given covariate histories (9, 10). The adequacy of proportionality and other single- and dual-outcome hazard ratio model specifications can be examined by including additional, typically time-dependent, terms in the hazard ratio models to reflect potential departures from assumed hazard ratio models, and by testing for zero values of corresponding regression coefficients.

For interpretation of hormone therapy associations, it is useful to simultaneously consider single-outcome hazard ratios and corresponding dual-outcome hazard ratios. CEE and CEE + MPA have associations with several clinical outcomes that are similar between the 2 preparations, but also some that are less favorable for CEE + MPA than for CEE, leading to a "global index" that is elevated for CEE + MPA but approximately neutral for CEE (7, 8, 14). A detailed study of cardiovascular disease biomarkers revealed

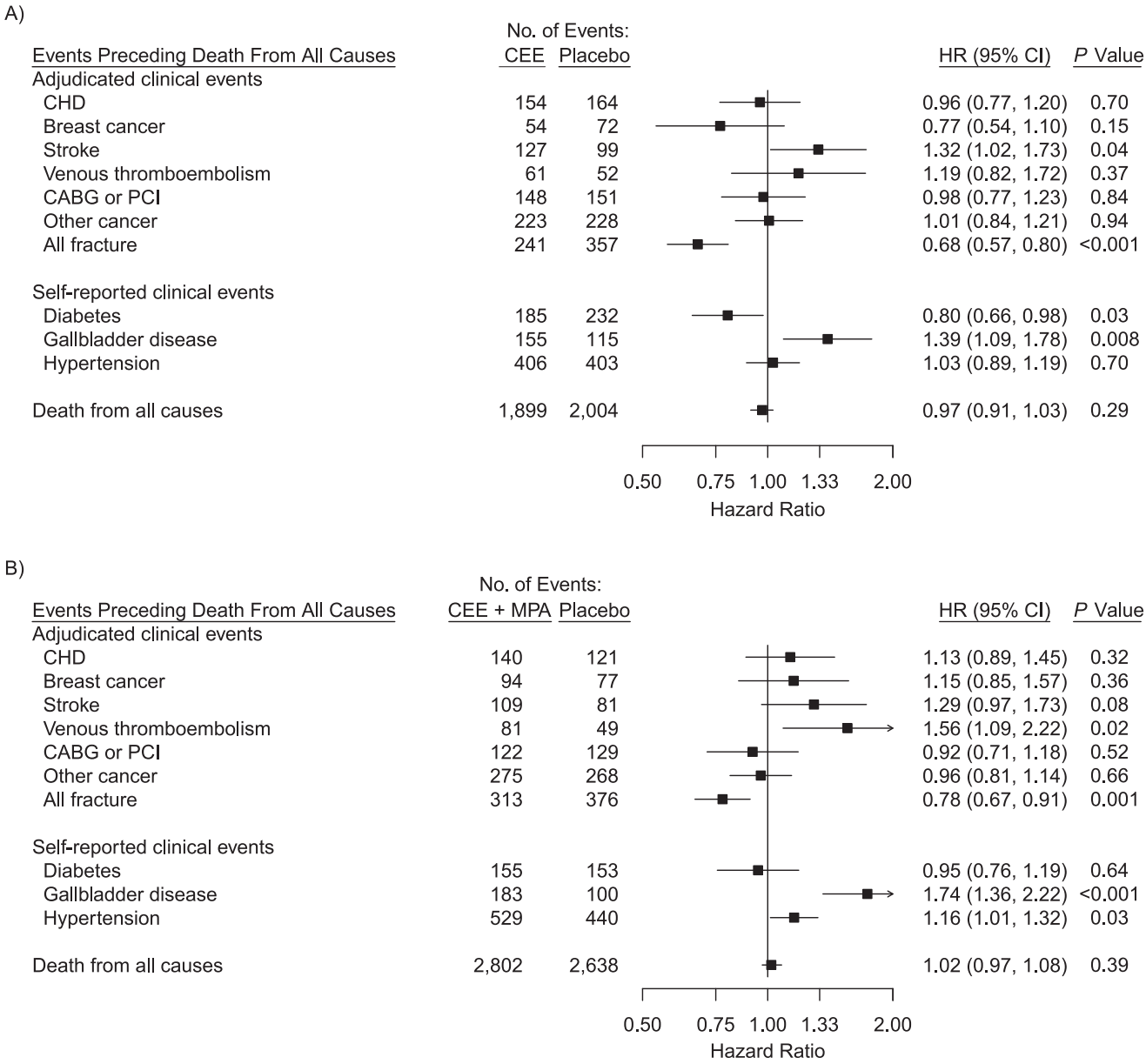

Figure 2. Dual-outcome hazard ratios (HRs) for clinical outcomes during the program intervention period of trials of conjugated equine estrogens (CEE) (A) and CEE + medroxyprogesterone acetate (MPA) (B) in conjunction with all-cause mortality over the cumulative duration of follow-up, Women's Health Initiative, 1993–2016. The program intervention period ended on March 31, 2005. *P* values (2-sided) are from tests of HR = 1 for the dual outcome. HR < 1 favors menopausal hormone therapy; HR > 1 favors placebo. All participants were postmenopausal and aged 50–79 years when enrolled at 40 US clinical centers during 1993–1998. Bars show the 95% confidence intervals (CIs). CABG, coronary artery bypass graft; CHD, coronary heart disease; PCI, percutaneous coronary intervention.

differing concentrations of inflammatory and coagulation factors between intervention and placebo groups (18, 19), but it was unclear whether these changes could contribute to an explanation for early CHD risk elevations, or for more sustained stroke and venous disease elevations, with either treatment. Along the same lines, a proteomic discovery project comparing baseline protein concentrations with year 1 concentrations revealed plasma concentration changes for about half of the approximately 250 proteins quantified (20), with changes in many pathways relevant to clinical outcomes. These changes were quantitatively very similar for the 2 hormone therapy preparations.

Even though all-cause mortality, as well as cause-specific mortality, contrasts were approximately null for both preparations (8), the analyses presented here (Figure 2) show dual-outcome reduction in fracture followed by death and

dual-outcome increase in gallbladder disease followed by death. These findings add to the list of similar health risk alterations for the 2 preparations and may have clinical relevance for informing women, for example, as to how few years of hormone therapy might relate to risk of developing a fracture while under treatment and dying during the subsequent decade or two.

Dual-outcome hazard ratio treatment associations were also evident for pairs of clinical outcomes having nonfatal components, especially over cumulative follow-up (Web Figures 4 and 5). Dual outcomes involving gallbladder disease along with either hypertension or cardiovascular outcomes were frequently elevated. However, when these elevations were broken out by the ordering of the 2 outcomes (Web Figures 6 and 7), it was not evident that hazard ratios differed in any systematic way by outcome order, suggesting different disease processes and pathways for gallbladder disease and the other outcomes.

On the other hand, some dual-outcome hazard ratio associations appeared to differ between the 2 preparations, providing leads to further understanding of health-related differences between the 2 treatment regimens. For example, the hazard ratio for hypertension and death was 0.76 (95% CI: 0.48, 1.21) for CEE (Web Figure 1) and 1.77 (95% CI: 1.04, 3.00) for CEE + MPA (Web Figure 2) during the trial intervention periods and 1.03 (95% CI: 0.89, 1.19) for CEE and 1.16 (95% CI: 1.01, 1.32) for CEE + MPA for hypertension during the intervention period followed by death over cumulative follow-up (Figure 2). This suggests that the small observed blood pressure increases with hormone therapy, with similar hypertension hazard ratios for the 2 preparations (14), may have greater relative mortality risk implications for the generally healthier CEE + MPA trial population than for the posthysterectomy CEE trial population. This, along with some blunting of favorable blood cholesterol changes with CEE + MPA as compared with CEE (18), may help to explain less favorable cardiovascular hazard ratios with CEE + MPA compared with CEE.

Similarly, the dual-outcome hazard ratio for diabetes during the intervention period followed by death over cumulative follow-up was 0.80 (95% CI: 0.66, 0.98) for CEE and 0.95 (95% CI: 0.76, 1.19) for CEE + MPA. This suggests a possible detrimental role for MPA in relation to these dual outcomes, in spite of similar reductions in fasting glucose, insulin, and insulin resistance for the 2 regimens (21, 22).

The most striking difference in hazard ratios between the 2 hormone therapy formulations was seen for invasive breast cancer, with overall risk elevation with CEE + MPA and risk reduction with CEE (8, 23, 24) (Web Figure 3). While reasons for these discrepant patterns are still under investigation and may involve timing issues such as age (8) or gap time from menopause to hormone therapy initiation (25), it is interesting that the dual outcome of gallbladder disease and breast cancer had a hazard ratio of 2.51 (95% CI: 1.20, 5.24) for CEE (Web Figure 4) and a hazard ratio of 2.14 (95% CI: 1.35, 3.38) for CEE + MPA (Web Figure 5) over cumulative follow-up. For CEE, the risk elevation derives primarily from dual outcomes where the gallbladder disease precedes the breast cancer occurrence, while there is no suggested hazard ratio dependence on outcome ordering for CEE + MPA (Web Figures 6 and 7). This finding is consistent with our report that associations of baseline sex hormone concentrations with breast cancer risk tend to disappear in the presence of MPA (26), and it suggests that breast cancer benefits that would otherwise be associated with CEE may be reduced or lost when MPA is added to the treatment regimen. Additional analyses showed that participants assigned to active CEE or CEE + MPA who experienced gallbladder disease did not have subsequent reduced intervention adherence. Hence, one could speculate that participants who develop gallbladder disease early may have reduced or altered metabolism of CEE, rendering them less able to realize breast cancer benefits from CEE, whereas no such benefits are suggested when MPA is included in the regimen regardless of whether or not gallbladder disease precedes breast cancer occurrence.

In summary, we have presented an application of dual-outcome hazard ratio estimation methods in the context of the WHI randomized, controlled trials of menopausal hormones among US women. In conjunction with corresponding single-outcome hazard ratio analyses, these methods lead to insights and hypotheses to help explain a complex profile of health benefits and risks with CEE and CEE + MPA, as well as differences between the 2 formulations.

Strengths of this contribution include the randomized, placebo-controlled trial design, with high quality and long-term outcome ascertainment. Limitations include multiple testing issues related to the large number of outcome pairs considered. Multiple testing considerations, along with the fact that the present analyses were not prespecified in the trial protocols, implies that these analyses should mostly be regarded as exploratory, with the principal goal of developing leads for the further elucidation of hormone therapy influences on clinical outcomes, in this study population of postmenopausal US women. Some of the estimated dual-outcome hazard ratios differed sufficiently from unity to remain significant at the 0.05 level following Bonferroni adjustment for the set of paired outcomes considered, in which case a stronger interpretation may be justified.

ACKNOWLEDGMENTS

Author affiliations: Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington (Ross L. Prentice, Aaron K. Aragaki, Garnet L. Anderson); Division of Medical Oncology/Hematology, Los Angeles Biomedical Institute at Harbor-UCLA Medical Center, Torrance, California (Rowan T. Chlebowski); Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina (Shanshan Zhao); Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, Bethesda, Maryland (Jacques E. Rossouw); Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, Iowa (Robert Wallace); Department of Epidemiology and Environmental Health, School of Public

Health and Health Professions, State University of New York at Buffalo, Buffalo, New York (Hailey Banack); Department of Family Medicine and Public Health, School of Medicine, University of California, San Diego, San Diego, California (Aladdin H. Shadyab); Department of Public Health Sciences, School of Medicine, University of California, Davis, Davis, California (Lihong Qi); Department of Biostatistics and Data Science, Wake Forest School of Medicine, Wake Forest University, Winston-Salem, North Carolina (Beverly M. Snively); North American Menopause Society, Mayfield Heights, Ohio (Margery Gass); and Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts (JoAnn E. Manson).

This work was supported by the National Cancer Institute (grants R01 CA119171 and R01 CA210921) and by the National Heart, Lung and Blood Institute, which supports the infrastructure of the Women's Health Initiative (WHI) (contracts HHSN268201100046C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, HHSN268201600004C, and HHSN271201600004C).

We acknowledge the following investigators in the WHI Program: *Program Office* (National Heart, Lung, and Blood Institute, Bethesda, Maryland)—Jacques Rossouw, Shari Ludlam, Joan McGowan, Leslie Ford, and Nancy Geller; *Clinical Coordinating Center* (Fred Hutchinson Cancer Research Center, Seattle, Washington)—Garnet Anderson, Ross Prentice, Andrea LaCroix, and Charles Kooperberg; *investigators and academic centers*: Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts—JoAnn E. Manson; MedStar Health Research Institute/Howard University, Washington, DC—Barbara V. Howard; Stanford Prevention Research Center, Stanford, California—Marcia L. Stefanick; The Ohio State University, Columbus, Ohio—Rebecca Jackson; University of Arizona, Tucson/Phoenix, Arizona—Cynthia A. Thomson; State University of New York at Buffalo, Buffalo, New York—Jean Wactawski-Wende; University of Florida, Gainesville/Jacksonville, Florida—Marian Limacher; University of Iowa, Iowa City/Davenport, Iowa—Jennifer Robinson; University of Pittsburgh, Pittsburgh, Pennsylvania—Lewis Kuller; Wake Forest University School of Medicine, Winston-Salem, North Carolina—Sally Shumaker; University of Nevada, Reno, Nevada—Robert Brunner; Women's Health Initiative Memory Study, Wake Forest University School of Medicine, Winston-Salem, North Carolina—Mark Espeland.

For a list of all WHI investigators, please visit https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf.

The views expressed in this article are our own and do not necessarily represent the views of the National Heart, Lung, and Blood Institute, the National Institutes of Health, or the US Department of Health and Human Services.

R.T.C. reports consulting arrangements with Novartis International AG (Basel, Switzerland), AstraZeneca AB (Cambridge, United Kingdom), Amgen, Inc. (Thousand Oaks, California), Immunomedics (Morris Plains, New Jersey), Puma Biotechnology, Inc. (Los Angeles, California), and Genentech, Inc. (South San Francisco, California). The other authors have no pertinent disclosures.

REFERENCES

1. Ettinger B, Wang SM, Leslie RS, et al. Evolution of postmenopausal hormone therapy between 2002 and 2009. *Menopause*. 2012;19(6):610–615.
2. Steinkellner AR, Denison SE, Eldridge SL, et al. A decade of postmenopausal hormone therapy prescribing in the United States: long-term effects of the Women's Health Initiative. *Menopause*. 2012;19(6):616–621.
3. Weissfeld JL, Liu W, Woods C, et al. Trends in oral and vaginally administered estrogen use among US women 50 years of age and older with commercial health insurance. *Menopause*. 2018;25(6):611–614.
4. Crawford SL, Crandall CJ, Derby CA, et al. Menopausal hormone therapy trends before versus after 2002: impact of the Women's Health Initiative study results. *Menopause*. 2018;26(6):588–597.
5. Ravdin PM, Cronin KA, Howlader N, et al. The decrease in breast-cancer incidence in 2003 in the United States. *N Engl J Med*. 2007;356(16):1670–1674.
6. Roth JA, Etzioni R, Waters JM, et al. Economic return of the Women's Health Initiative Estrogen plus Progestin Trial: a modeling study. *Ann Intern Med*. 2014;160(9):594–602.
7. Manson JE, Chlebowski RT, Stefanick ML, et al. Menopausal hormone therapy and health outcomes during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials. *JAMA*. 2013;310(13):1353–1368.
8. Manson JE, Aragaki AK, Rossouw JE, et al. Menopausal hormone therapy and long-term all-cause and cause-specific mortality: the Women's Health Initiative randomized trials. *JAMA*. 2017;318(10):927–938.
9. Prentice RL, Zhao S. Regression models and multivariate life tables [published online ahead of print February 18, 2020]. *J Am Stat Assoc*. (doi:10.1080/01621459.2020.1713792).
10. Prentice RL, Zhao S. *The Statistical Analysis of Multivariate Failure Time Data: A Marginal Modeling Approach*. Boca Raton, FL: Chapman & Hall/CRC Press; 2019.
11. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials*. 1998;19(1):61–109.
12. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321–333.
13. Anderson GL, Limacher M, Assaf AR, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA*. 2004;291(14):1701–1712.
14. Swica Y, Warren MP, Manson JE, et al. Effects of oral conjugated equine estrogens with or without medroxyprogesterone acetate on incident hypertension in the Women's Health Initiative hormone therapy trials. *Menopause*. 2018;25(7):753–757.
15. Anderson GL, Kooperberg C, Geller N, et al. Monitoring and reporting of the Women's Health Initiative randomized hormone therapy trials. *Clin Trials*. 2007;4(3):207–217.
16. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat*. 1982;10(4):1100–1120.

17. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc.; 2002.
18. Rossouw JE, Cushman M, Greenland P, et al. Inflammatory, lipid, thrombotic, and genetic markers of coronary heart disease risk in the Women's Health Initiative trials of hormone therapy. *Arch Intern Med*. 2008;168(20): 2245–2253.
19. Kooperberg C, Cushman M, Hsia J, et al. Can biomarkers identify women at increased stroke risk? The Women's Health Initiative hormone trials. *PLoS Clin Trials*. 2007;2(6):e28.
20. Pitteri SJ, Hanash SM, Aragaki A, et al. Postmenopausal estrogen and progestin effects on the serum proteome. *Genome Med*. 2009;1(12):121.1–121.14.
21. Margolis KL, Bonds DE, Rodabough RJ, et al. Effect of oestrogen plus progestin on the incidence of diabetes in postmenopausal women: results from the Women's Health Initiative Hormone Trial. *Diabetologia*. 2004;47(7): 1175–1187.
22. Bonds DE, Lasser N, Qi L, et al. The effect of conjugated equine oestrogen on diabetes incidence: the Women's Health Initiative randomised trial. *Diabetologia*. 2006;49(3): 459–468.
23. Chlebowski RC, Anderson GL, Gass M, et al. Estrogen plus progestin and breast cancer incidence and mortality in postmenopausal women. *JAMA*. 2010;304(15):1684–1692.
24. Anderson GL, Chlebowski RT, Aragaki AK, et al. Conjugated equine oestrogen and breast cancer incidence and mortality in postmenopausal women with hysterectomy: extended follow-up of the Women's Health Initiative randomised placebo-controlled trial. *Lancet Oncol*. 2012;13(5):476–486.
25. Prentice RL, Manson JE, Langer RD, et al. Benefits and risks of postmenopausal hormone therapy when it is initiated soon after menopause. *Am J Epidemiol*. 2009;170(1):12–23.
26. Zhao S, Chlebowski RT, Anderson GL, et al. Sex hormone associations with breast cancer risk and the mediation of randomized trial postmenopausal hormone therapy effects. *Breast Cancer Res*. 2014;16(2):R30.