Title

Metabolomics-Based Biomarker for Dietary Fat and Associations with Chronic Disease Risk in Postmenopausal Women.

Journal

Journal of Nutrition, 153(9)

Authors

Prentice, Ross
Vasan, Sowmya
Tinker, Lesley
et al.

Publication Date

2023-09-01

DOI

10.1016/j.tjnut.2023.05.021

Peer reviewed

The Journal of Nutrition 153 (2023) 2651–2662

American Society for Nutrition
Excellence in Nutrition Research and Practice

JN THE JOURNAL OF NUTRITION

journal homepage: https://jn.nutrition.org/

Nutritional Epidemiology

Metabolomics-Based Biomarker for Dietary Fat and Associations with Chronic Disease Risk in Postmenopausal Women

Ross L. Prentice [1,2,*], Sowmya Vasan [1], Lesley F. Tinker [1], Marian L. Neuhouser [1,2], Sandi L. Navarro [1], Daniel Raftery [3], GA Nagana Gowda [3], Mary Pettinger [1], Aaron K. Aragaki [1], Johanna W. Lampe [1,2], Ying Huang [1,2], Linda Van Horn [4], JoAnn E. Manson [5,6], Robert Wallace [7], Yasmin Mossavar-Rahmani [8], Jean Wactawski-Wende [9], Simin Liu [10], Linda Snetselaar [7], Barbara V. Howard [11], Rowan T. Chlebowski [12], Cheng Zheng [13]

[1] *Division of Public Health Sciences, Fred Hutchinson Cancer Center, Seattle, WA, United States;* [2] *School of Public Health, University of Washington, Seattle, WA, United States;* [3] *Department of Anesthesiology and Pain Medicine, University of Washington, Seattle, WA, United States;* [4] *Department of Preventive Medicine, Northwestern University, Chicago, IL, United States;* [5] *Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States;* [6] *Women's Hospital, Harvard Medical School, Boston, MA, United States;* [7] *College of Public Health, University of Iowa, Iowa City, IA, United States;* [8] *Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, United States;* [9] *Department of Epidemiology and Environmental Health, University at Buffalo, Buffalo, NY, United States;* [10] *Department of Epidemiology, School of Public Health, Brown University, Providence, RI, United States;* [11] *Department of Medicine, Georgetown University Medical Center, and MedStar Health Research Institute, Hyattsville, MD, United States;* [12] *Division of Medical Oncology and Hematology, The Lundquist Institute, Torrance, CA, United States;* [13] *Department of Biostatistics, University of Nebraska Medical Center, Omaha, NE, United States*

A B S T R A C T

Background: The Women's Health Initiative (WHI) randomized, controlled Dietary Modification (DM) trial of a low-fat dietary pattern suggested intervention benefits related to breast cancer, coronary heart disease (CHD), and diabetes. Here, we use WHI observational data for further insight into the chronic disease implications of adopting this type of low-fat dietary pattern.

Objectives: We aimed to use our earlier work on metabolomics-based biomarkers of carbohydrate and protein to develop a fat intake biomarker by subtraction, to use the resulting biomarker to develop calibration equations that adjusts self-reported fat intake for measurement error, and to study associations of biomarker-calibrated fat intake with chronic disease risk in WHI cohorts. Corresponding studies for specific fatty acids will follow separately.

Methods: Prospective disease association results are presented using WHI cohorts of postmenopausal women, aged 50–79 y when enrolled at 40 United States clinical centers. Biomarker equations were developed using an embedded human feeding study (n = 153). Calibration equations were developed using a WHI nutritional biomarker study (n = 436). Calibrated intakes were associated with cancer, cardiovascular diseases, and diabetes incidence in WHI cohorts (n = 81,954) over an approximate 20-y follow-up period.

Results: A biomarker for fat density was developed by subtracting protein, carbohydrate, and alcohol densities from one. A calibration equation was developed for fat density. Hazard ratios (95% confidence intervals) for 20% higher fat density were 1.16 (1.06, 1.27) for breast cancer, 1.13 (1.02, 1.26) for CHD, and 1.19 (1.13, 1.26) for diabetes, in substantial agreement with findings from the DM trial. With control for additional dietary variables, especially fiber, fat density was no longer associated with CHD, with hazard ratio (95% confidence interval) of 1.00 (0.88, 1.13), whereas that for breast cancer was 1.11 (1.00, 1.24).

Conclusions: WHI observational data support prior DM trial findings of low-fat dietary pattern benefits in this population of postmenopausal United States women.

Trial registration number: This study is registered with clinicaltrials.gov identifier: NCT00000611.

Keywords: biomarker, cancer, cardiovascular disease, diabetes, dietary fat, fat density, measurement error

DOI of original article: https://doi.org/10.1016/j.tjnut.2023.05.003.

Abbreviations used: DM, dietary modification; DM-C, dietary modification comparison group; NPAAS, Nutrition and Physical Activity Assessment Study; NPAAS-FS, NPAAS-Feeding Study; OS, Observational Study; T2D, type 2 diabetes; WHI, Women's Health Initiative; EI, energy intake; SFA, saturated fatty acids; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; FA, fatty acids; CHD, coronary heart disease; CVD, cardiovascular disease; FFQ, food frequency questionnaire; BMI, body mass index; MI, myocardial infarction; CABG/PCI, coronary artery bypass graft/percutaneous coronary intervention; CHF, congestive heart failure; HR, hazard ratio; CI, confidence interval; P-value, significance level.

* Corresponding author. *E-mail address:* rprentic@whi.org (R.L. Prentice).

https://doi.org/10.1016/j.tjnut.2023.05.021
Received 10 December 2022; Received in revised form 2 May 2023; Accepted 23 May 2023; Available online 26 May 2023

Introduction

The Women's Health Initiative (WHI) randomized controlled Dietary Modification (DM) trial studied the health effects of a low-fat dietary pattern intervention among postmenopausal United States women (n = 48,835). The intervention was designed to reduce the fraction of energy from total fat, hereafter fat density, without attempting to differentially alter the densities of SFA, MUFA, or PUFA, or to alter total energy intake (EI). Intervention goals included a reduction in fat density from customary values of ~35% to 20%, while increasing servings of vegetables and fruit to 5/d and grains servings to 6/d.

At 1 y after randomization based on FFQ assessments, EI in the intervention group diet was ~11% lower in fat, 10% higher in carbohydrate, and 1% higher in protein compared with the usual diet comparison group. Also, vegetable and fruit intake was higher by approximately one serving/d and grains by approximately one-half serving/d in the intervention group. Note that, according to these self-reported dietary data, the major dietary change in the intervention group is replacement of fat by carbohydrate [1].

After an 8.1-y (median) intervention period, intervention compared with usual diet comparison group contrasts were suggestive of benefit for primary outcome breast cancer (P = 0.09), and the composite outcome of breast cancer followed by death was nominally significantly (P = 0.02) reduced [2]. There was little evidence for an intervention effect on the coprimary colorectal cancer, or for the secondary CHD outcomes [3–5]. With longer-term nonintervention follow-up, randomization group differences were still NS for the primary or secondary outcomes, but reduction in the composite breast cancer followed by death continued, and death attributed to breast cancer was also significantly reduced in the intervention group [6], as was the incidence of diabetes requiring insulin [7]. Also, CHD incidence was reduced among the healthy, nonhypertensive subset of the trial cohort where, unlike complementary subsets, there was no evidence of postrandomization confounding by statin use [8]. See Prentice et al. [1,9] for recent reviews of DM trial results.

There is a massive literature reporting associations of specific FA intake categories with CVD and other chronic diseases. Our research group plans to separately report on the development and application of biomarkers for SFA, MUFA, and PUFA in WHI cohorts.

Here, we consider the construction and chronic disease application of a biomarker for total fat density using our recently proposed biomarkers for carbohydrate and protein densities [10], and we compare resulting biomarker-calibrated intake associations, in analyses that control for total energy, with findings from the DM randomized trial. The biomarker-calibrated association analyses presented here are best thought of applying to a high-fat dietary pattern. To examine the health implications of a change in dietary fat specifically, we augment these analyses by controlling for other dietary factors, including sodium, vegetable servings, fruit servings, and fiber.

As in our earlier studies [10,11], our approach to biomarker development and application involves 3 steps: the first involves biomarker equation development using a WHI feeding study (n = 153), the second uses biomarker values from this equation to develop calibration equations that aim to adjust self-reported dietary data for measurement error using a WHI biomarker study (n = 436), whereas the third uses calibrated intake values from this calibration equation to study calibrated intakes in relation to disease risk in larger WHI cohorts (n = 81,954). This approach will be elaborated in Methods, and related strengths and weaknesses will be briefly described in Discussion.

Methods

The context and resources for the dietary fat analyses reported here are the same as those for our recent report [10] on carbohydrate and protein biomarker development and disease associations.

Study cohorts

Specifically, during 1993–1998, 48,835 participants were randomly assigned in the WHI DM trial, with 29,294 assigned to the usual diet comparison group (DM-C), and 93,676 participants were enrolled in the companion prospective WHI Observational Study (OS) [12]. All participants were postmenopausal and in the age range 50–79 y when enrolled at 40 United States clinical centers. The WHI FFQ [13] targeted dietary intake over the preceding 3-mo period and was administered at baseline and year 1 in the DM trial, and approximately every 3 y thereafter during the trial intervention period (ended March 31, 2005), and the same FFQ was administered at baseline and at year 3 in the OS. Here, we used FFQs collected at 1 y after randomization in the DM-C, rather than at enrollment, to reduce assessment biases because of the trial eligibility criterion of FFQ fat density of ≥32%. The 1-y FFQ assessments will be referred to here as "baseline FFQs." FFQs at enrollment were used for baseline self-reported diet in the OS. All nutrient content estimates from self-report assessments were derived from the University of Minnesota's Nutrition Data System for Research (NDS-R version 2005). Participants completed core questionnaires at WHI enrollment, including medical history, reproductive history, family history, personal habits, medications and dietary supplements, physical measures, and also provided a fasting blood sample [12].

Nutrition and Physical Activity Assessment Study

After an initial Nutrition Biomarker Study in the DM trial cohort [14], we conducted a Nutrition and Physical Activity Assessment Study (NPAAS) [15] among 450 OS participants during 2007–2009. Its purposes were to examine the measurement properties of dietary self-report data for nutritional variables having an established biomarker, and to use biomarker data to correct corresponding dietary self-report data for measurement error in disease association analyses. We recruited WHI participants at 9 clinical centers to NPAAS, with an overrepresentation of Black and Hispanic women and of women having BMI >30.0 kg/m^2. Our study protocol required 2 clinic visits separated by 2 wk and included various at-home activities. A 20% reliability subsample repeated the protocol ~6 mo after their initial study participation. The first NPAAS visit included measured height and weight; DLW dosing for total energy expenditure assessment; completion of an FFQ, dietary supplement inventory, and other questionnaires; and collection of a fasting blood specimen. Participants received instructions and a

kit for 24-h urine collection for home completion. At the second clinic visit, participants brought 24-h urine specimens collected over the preceding day, provided a fasting blood specimen, and provided additional spot urine specimens to complete the DLW protocol. Baseline characteristics in the NPAAS cohort have been presented [15]. Participants were similar in age to other WHI participants, 60% were overweight or obese (i.e., BMI $\geq$25.0 kg/m^2), 95% were nonsmokers (never or past smokers), 51% had a college degree or higher education, 19%, 14%, and 64% self-classified as being of Black, Hispanic, or non-Hispanic White race/ethnicity, respectively.

In these biomarker studies, total EI was assessed using the DLW procedure [16], and total protein intake was assessed using urinary nitrogen [17] using specimens from the initial and replicate protocol applications.

NPAAS-Feeding Study

We conducted the NPAAS-Feeding Study (NPAAS-FS) among 153 WHI women in the Seattle area during 2010–2014. Of the 153 women, 14 had previously participated in NPAAS and were excluded from the calibration equation development described below. Participants were provided food and beverages over a 2-wk feeding period, with individualized diets that were intended to approximate their usual diets, so that blood and urine concentrations would stabilize quickly and intake variations in the study cohort would be substantially retained during the feeding period [18]. Biomarkers developed for the macronutrient intakes studied here rely on metabolomics profiles from the second clinic visit serum and 24-h urine specimens, along with the inclusion of readily available participant characteristic measures. Baseline demographic and lifestyle characteristics for participants in the NPAAS-FS have been reported [18]. Participants were well educated (83% college degree or higher), and nonsmokers (98%). Most were White (95%), overweight or obese (60%), and were of similar ages to other WHI enrollees.

Metabolite profiling

Serum and 24-h urine metabolomics profiles, obtained using specimens collected at the end of the NPAAS-FS feeding period, were derived as described by Zheng et al. [19].

Serum metabolite measurements

Briefly, serum samples from NPAAS-FS participants were analyzed by targeted LC-MS/MS using LC MS. A total of 303 metabolites were targeted, of which 155 were detected with $<$20% missing values. Separately, lipid metabolites were measured using a lipidyzer platform, including differential mobility spectrometry method that targeted 1070 lipids in 13 major lipid classes, resulting in 664 serum lipids that had $<$20% missing values.

Urine metabolite measurements

Metabolite profiles from 24-h urine samples were analyzed by NMR spectroscopy. Relative concentrations for 57 targeted metabolites were obtained. None of the metabolites had missing values. Urine metabolites were also analyzed by untargeted GC-MS resulting in the identification of 275 metabolites with $<$20% missing values.

Protein and carbohydrate biomarkers

Biomarker equations for protein density and carbohydrate density were developed as described previously [10]. These used serum and 24-h urine metabolite profiles, DLW energy and urinary nitrogen protein measures, and participant characteristics. Baseline FFQ data were added to the variables considered for biomarker specification [10] to prepare these biomarkers for use in disease association analyses, which condition on baseline FFQ data. Because of the relatively high dimensionality of the metabolomics data, biomarker development used a LASSO procedure [20] for variable selection, and used cross-validation for the study of biomarker properties.

Outcome ascertainment, follow-up, and disease categories

Clinical outcomes were reported biannually in the DM trial and annually in the OS, by self-administered questionnaire [21] throughout the time from enrollment in 1993–1998 to the end of the intervention period (March 31, 2005), and annually thereafter in both cohorts. An initial report of CVD during cohort follow-up was confirmed by review of medical records by physician-adjudicators. In addition, CHD (defined as nonfatal MI plus CHD death), stroke (ischemic plus hemorrhagic), heart failure, and all deaths were centrally reviewed by expert physician investigator committees. All invasive cancers, except nonmelanoma skin cancer, were centrally coded using the NCI's SEER procedures. Prevalent, treated type 2 diabetes (T2D) (by oral agents or insulin) at baseline was self-reported during eligibility screening. Incident treated T2D during follow-up was documented by self-report at each annual contact. These sources have been shown to be consistent with medication inventories of oral agents or insulin [22].

After the intervention period, WHI participants had the opportunity to enroll in follow-up through September 30, 2010, and subsequently for additional open-ended follow-up, with $>$80% of women doing so on each occasion. Cancer, diabetes, and all-cause mortality (including National Death Index matching) outcomes through December 31, 2020, are included here. Follow-up for CVD incidence is included only through September 30, 2010, because self-reports for most WHI participants were not adjudicated after that date. Also, heart failure adjudication in WHI cohorts stopped after March 31, 2005. The median follow-up duration is 11.3 y for CVD incidence, 7.8 y for heart failure, and $\sim$20 y for cancer, diabetes, and mortality outcomes. Disease outcome categories are those reported in our previous report on biomarker-calibrated protein and carbohydrate intake [10].

Statistical methods

Biomarker development for total fat in NPAAS-FS

The metabolite model building procedures described above for carbohydrate and protein density did not yield metabolite combinations that came close to meeting a 36% cross-validated regression R^2 criterion for percent of variation explained for fat density. Here, instead, we consider a fat density biomarker based on subtraction. Specifically, we subtract biomarker estimates of protein density and carbohydrate density, as well as the corresponding FFQ estimate of alcohol density, from unity. Note that alcohol comprised only a small fraction of energy consumption in the feeding study cohort, and that alcohol was not provided in

the feeding study. Instead, participants were allowed to continue their usual pattern of alcohol consumption, and alcohol intake was included in feeding study logs. We also consider participant characteristics and baseline FFQ assessments for possible inclusion in fat density biomarker equations, with P value of <0.10 for inclusion and retention in biomarker equation model building, and we use cross-validation to reduce overfitting. The participant characteristics considered included dietary supplement use, race/ethnicity, season, education, age, BMI, and self-reported leisure activity (metabolic equivalent unit hr/wk). As in our previous work [10], cross-validated linear regression fraction of provided dietary intake variation explained (CV-R^2) values were calculated as averages of R^2 values over 100 random splits of the NPAAS-FS data set into 2 approximately equal-sized subsets. Protein and carbohydrate biomarkers were kept intact in the cross-validation procedure. A 36% or larger CV-R^2 was a biomarker criterion, along with informal further biomarker sensitivity and specificity considerations.

Calibration equation development for total fat density in NPAAS

A biomarker equation essentially meeting CV-R^2 criteria was able to be developed for log-fat density, and this equation was used to calculate biomarker-based intake estimates for the 436 participants in NPAAS who were not a part of NPAAS-FS. These NPAAS biomarker values were regressed linearly on concurrent FFQ log-fat density assessments, and on a disease category-specific set of personal characteristics listed in Supplemental Table 1, for development of calibration equations for estimating fat density intakes in larger WHI cohorts. Briefly, CVD analyses included age (linear), BMI, family income, education, cigarette smoking history, alcohol consumption, leisure physical activity, any dietary supplement use, prior menopausal hormone use, hypertension, personal history of cancer, family history of MI, stroke, or diabetes, use of medications to lower blood pressure, blood lipids, or blood glucose, and season in which the NPAAS FFQ was completed. Invasive cancer analyses included these same variables, exclusive of prevalent CVD and of family history of CVD or diabetes, and inclusive of Gail model 5-y breast cancer risk score, family history of colorectal cancer, and personal history of colon polyp removal. T2D analyses included the same variables as the CVD analyses except for family history of MI or stroke. An assumption of independent measurement errors for the 2 assessments in the 14 participant NPAAS and NPAAS-FS overlap, which were based on specimen collections separated by ~4 y, leads to regression R^2 values that are adjusted for temporal variation in biomarker values. The adjustment involves dividing the linear equation R^2 values by the upper 90th percentile of the estimated paired correlation. This percentile rather than the estimated correlation itself was used to avoid an overly large adjustment that may otherwise arise with this very small replicate sample. An adjusted R^2 value of $\geq$36% was a criterion for a suitable calibration equation.

Disease association analyses in the DM-C and OS using Biomarker-Calibrated FFQ Data

Table 1, also given in Prentice et al. [10], presents baseline demographic and lifestyle characteristics for the 81,954 participants, 16,939 from the DM-C, and 65,015 from the OS. Participants averaged ~62 y of age at baseline. Approximately 60% were overweight or obese, 85% were White, >40% had a college degree or higher, and 94% were current nonsmokers. Participants having CVD, invasive cancer, or treated T2D before WHI enrollment in the OS, and before year 1 in the DM-C, were excluded from respective CVD, cancer, or diabetes analyses.

We entered calibrated intake values into Cox regression models [23], along with disease-specific potential confounding factors. We assumed a linear modeling of log-HR on log-fat density, and this implies a fixed HR for a fractional increase in fat density. For display purposes, we present HR estimates for a 20% increment in the fat density. A 20% increase is well within the intake variation estimated in WHI cohorts. Specifically, the FFQ geometric mean (95% confidence range) in the combined cohorts (n = 81,954) is 30% (17%, 53%) for fat density.

As in Prentice et al. [10], we stratified baseline hazard rates in the Cox model analyses on baseline age (i.e., year 1 in DM-C, enrollment in OS) in 5-y categories, race/ethnicity, on cohort (DM-C or OS), and, in the DM-C also on participation in the WHI hormone therapy trials (estrogen, estrogen placebo, estrogen plus progestin, estrogen plus progestin placebo, not randomized). Log-total energy was also included in the regression model. This implies that HRs for fat density estimate an HR factor beyond that for the fat contribution to total EI. The set of disease-specific potential confounding factors considered are those shown in Supplemental Table 1 and listed above for calibration equation model building. Missing data rates were generally low for specific covariates, but $\geq$20% participants had missing data on one or more modeled covariates in some analyses. Participants were excluded from outcome-specific analyses if any modeled covariate was missing. On the basis of sensitivity analyses that dropped covariates having relatively high missingness rates, thereby, including additional participants, this exclusion is not expected to materially affect disease association HR estimates. Corresponding HR estimates using FFQs without measurement error adjustment were also carried out and results are presented in main tables. Total energy was also biomarker calibrated in the calibrated total fat density analyses.

As in Prentice et al. [10], we defined disease occurrence time for a "case" developing a study outcome as days from "baseline" (year 1 in the DM-C and enrollment in the OS) to diagnosis. We defined censoring time for "noncases" as days from baseline to the earliest of date of death without the outcome under study, last contact, or March 31, 2005, for heart failure, September 30, 2010, for other CVD incidence outcomes, or December 31, 2020, for cancer, diabetes, and mortality outcomes. Because of uncertainty in the coefficients in the calibrated intake estimating equations, a "sandwich-type" estimator was used for the variance for the log-HR parameter estimates in calibrated intake analyses [24–26]. We present disease rates and numbers of included participants with events during follow-up in Supplemental Table 2.

Linearity of the associations between log-HR and log-fat density was studied by adding a quadratic term in log-fat density to the log-HR regression model, and examining evidence for nonzero quadratic coefficients.

Figure 1 shows cohorts and participant flow in the WHI DM-C and the OS, and in the NPAAS and NPAAS-FS subcohorts, over the intervention and postintervention phases of WHI. Note the 3 stages and 3 data sets used in these analyses: the NPAAS-FS (n =

TABLE 1
Baseline demographic and lifestyle characteristics of the analytic sample (n = 81,954) comprised of 16,939 women from the Women's Health Initiative Dietary Modification Trial Comparison Group and 65,015 from the Observational Study, enrolled during 1993–1998 at 40 United States clinical centers

Characteristic	OS (n = 65,015)		DM-C (n = 16,939)	
	n	%	n	%
Age (y)				
50–54	9126	14.0	1522	9.0
55–59	12,573	19.3	3634	21.5
60–64	14,381	22.1	4286	25.3
65–69	14,204	21.8	3902	23.0
70–74	10,259	15.8	2518	14.9
≥75	4472	6.9	1077	6.4
BMI (kg/m^2)				
<25	27,020	41.6	4579	27.0
25 to <30	22,140	34.1	6013	35.5
≥30	15,855	24.4	6347	37.5
Race/ethnicity				
White	56,032	86.2	14,250	84.1
Black	4122	6.3	1401	8.3
Hispanic	2022	3.1	536	3.2
American Indian	223	0.3	58	0.3
Asian/PI	1799	2.8	477	2.8
Unknown	817	1.3	217	1.3
Education				
<High school	2414	3.7	607	3.6
High school/GED	10,223	15.7	2876	17.0
School after high school	23,573	36.3	6648	39.2
College degree or higher	28,805	44.3	6808	40.2
Family income (USD/y)				
<$20k	9118	14.0	2258	13.3
$20k to <$35k	14,967	23.0	4084	24.1
$35k to <$50k	13,278	20.4	3664	21.6
$50k to <$75k	13,584	20.9	3671	21.7
≥$75k	14,068	21.6	3262	19.3
Season of FFQ completion				
Spring	16,755	25.8	4406	26.0
Summer	18,135	27.9	4172	24.6
Fall	15,148	23.3	4180	24.7
Winter	14,977	23.0	4181	24.7
Current smoker				
No	61,120	94.0	15,917	94.0
Yes	3895	6.0	1022	6.0
Alcohol[1]				
Nondrinker	18,410	28.3	5830	34.4
<1 drink/wk	20,583	31.7	4934	29.1
1 to <7 drinks/wk	17,424	26.8	4591	27.1
≥7 drinks/wk	8598	13.2	1584	9.4
Any dietary supplement use	36,358	55.9	8349	49.3
Medication use				
Antihyperlipidemic medication	5996	9.2	1562	9.2
Antidiabetic medication	1916	2.9	686	4.0
Antihypertensive medication	19,098	29.4	5611	33.1
Postmenopausal hormone use				
Never	25,334	39.0	6782	40.0
Past	9637	14.8	3357	19.8
Estrogens-alone	16,451	25.3	3932	23.2
Estrogens + Progestin	13,593	20.9	2868	16.9
Recreational physical activity, MET-h/wk				
None	8318	12.8	2952	17.4
>0 to ≤9.5	22,703	34.9	6910	40.8
>9.5 to ≤20.5	18,017	27.7	4110	24.3
>20.5	15,977	24.6	2967	17.5
History of CVD[2]				
No	61,934	95.3	16,263	96.0
Yes	3081	4.7	676	4.0
History of MI	1410	2.2	330	1.9
History of CABG/PCI	1139	1.8	215	1.3
History of CHF	643	1.0	136	0.8
History of stroke	833	1.3	184	1.1
History of cancer				
No	56,826	87.4	16,104	95.1
Yes	8189	12.6	835	4.9
Breast	3743	5.8	74	0.4
Colorectal	586	0.9	15	0.1
Ovary	427	0.7	72	0.4
Endometrium	1120	1.7	158	0.9
Thyroid	354	0.5	64	0.4
Cervix	794	1.2	211	1.2
Melanoma	877	1.3	113	0.7
Liver	24	0.0	1	0.0
Lung	145	0.2	15	0.1
Brain	32	0.0	6	0.0
Bone	42	0.1	9	0.1
Stomach	34	0.1	1	0.0
Leukemia	58	0.1	6	0.0
Bladder	120	0.2	12	0.1
Non-Hodgkin's lymphoma	148	0.2	6	0.0
Hodgkin's lymphoma	42	0.1	6	0.0
History of treated hypertension	15,954	24.5	5197	30.7
History of treated type 2 diabetes	2360	3.6	826	4.9
Family history of MI	33,803	52.0	8740	51.6
Family history of stroke	24,694	38.0	6404	37.8
Family history of breast cancer	9882	15.9	2333	14.4
Family history of colorectal cancer	10,831	16.7	2687	15.9
Family history of diabetes	20,889	32.1	5859	34.6
Gail model breast cancer risk score (tertiles)				
<1.26	18,972	29.2	5607	33.1
1.27–1.80	22,329	34.3	5900	34.8
>1.80	23,714	36.5	5432	32.1

Abbreviations: CABG/PCI, coronary artery bypass graft or percutaneous coronary intervention; CHF, congestive heart failure; CR, confidence range (2.5 percentile, 97.5 percentile); MET, metabolic equivalent unit.
[1] Drinks of alcohol defined as serving in mL (345 for beer, 177 for wine, 43 for liquor).
[2] Nonfatal MI, CABG/PCI, CHF, or stroke.

153) for biomarker equation development, the NPAAS (n = 436) for calibration equation development, and the larger WHI cohorts (n = 81,954) for disease association analyses. Some strengths and weaknesses of this study design will be described in Discussion.

Ethics

The WHI is funded primarily by the NHLBI. Participants provided written informed consent for their overall WHI, NPAAS, and NPAAS-FS activities. Related protocols were approved by the institutional review boards at the Fred Hutchinson Cancer Research Center and at each participating clinical center (clinicaltrials.gov identifier: NCT00000611).

Results

Table 2 shows results for a potential biomarker development equation for fat density. The CV-R^2 for log-fat density is 35.6%,

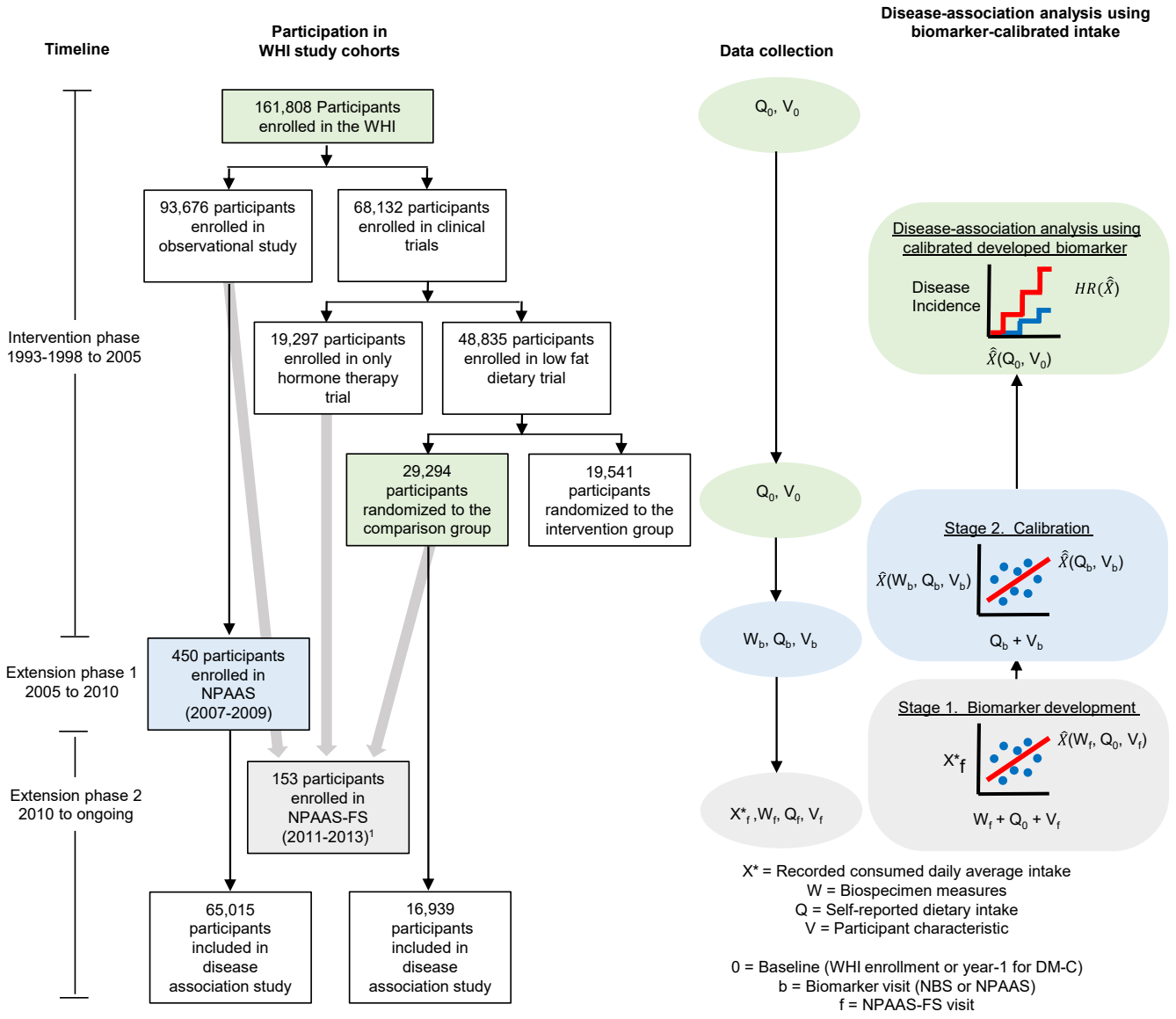

FIGURE 1. Study design for biomarker development, dietary intake calibration, and disease association analysis. Postmenopausal women were aged 50–79 y at enrollment during 1993–1998 at 40 United States clinical centers. Green, blue, and gray boxes indicate cohort (disease association, calibration, biomarker development), timing of data collection (WHI enrollment for OS or year 1 for DM-C, NPAAS, NPAAS-FS) and corresponding analysis stage, respectively. Gray arrows indicate parent-cohorts of participants who were recruited for NPAAS-FS. Participants in the disease association analyses were without prior personal history of the disease category under analysis, and had all data needed for intake calibration and confounding control. Notation for data collected (X*, W, Q, V) and analysis were based on Huang et al. [27] with pertinent regression variables shown along the X-axis (predictor variables), Y-axis (response variables), and line plots (developed biomarker; calibrated developed biomarker; estimated HR) of each regression-icon. [1]NPAAS-FS includes $n = 14$ who had previously participated in NPAAS. DM-C, dietary modification comparison group; FS, feeding study; NPAAS, Nutrition and Physical Activity Assessment Study; OS, Observational Study; WHI, Women's Health Initiative.

essentially meeting a 36% threshold, with most of this variation (32.9%) explained by the fat density construct reflecting EI not from protein, carbohydrate, or alcohol.

Table 3 shows a summary of R^2 and adjusted R^2 values for a potential log-fat density calibration equation. Adjusted R^2 values meet the 36% threshold for each of the 3 disease category-specific sets of potential confounding variables. The details of these calibration equations are presented in Supplemental Table 3.

Table 4 on the left side displays HRs (95% CIs) for a 20% increment in fat density relative to the incidence of various cancers. A 20% increment in calibrated fat density is associated with significant ($P < 0.05$) increases in breast, colon, obesity-related (defined as breast, colon, rectum, endometrium, or kidney cancer), and total invasive cancer, as well as leukemia and lymphoma. Significance levels were similar in the absence of calibration of FFQ fat density, but HRs were then strongly attenuated toward the null. For example, the breast cancer HR (95% CI) for a 20% fat density increment is 1.16 (1.06, 1.27) with biomarker calibration, compared with 1.03 (1.01, 1.05) without calibration. Several cancers, namely, rectum, endometrium, ovary, lung, bladder, kidney, and pancreas cancers, were not significantly related to fat density, with or without biomarker calibration.

TABLE 2
Cross-validated percent of variation in log-fat density explained by a log-fat density construct and other variables, in the NPAAS-Feeding Study ($n = 153$) conducted during 2010–2014

Log-fat density					
Variable	Estimate	SE	*P*	R^2 (%)	CV-R^2 (%)
(Intercept)	−1.791	0.297	0.000		
Diet supplement (Y/N)	0.049	0.029	0.094	0.3	0.2
Fat density construct[1]	0.686	0.079	0.000	37.2	32.9
Log-FFQ total energy[2]	−0.146	0.061	0.018	1.1	1.0
Log-FFQ fat (g/day)[2]	0.093	0.049	0.063	1.6	1.4
Total				40.2	35.6

Abbreviations: CV-R^2, cross-validated percent of variation explained[3]; NPAAS, Nutrition and Physical Activity Assessment Study; R^2, percent of variation explained.

[1] Fat density construct defined as 1 minus carbohydrate density biomarker, minus protein density biomarker, minus FFQ alcohol density. The variables meeting criteria [10] for inclusion in the carbohydrate density biomarker were triacylglycerol (TAG52.4.FA20.2) (serum), alanine (serum), maltose (urine), phosphatidylcholine (PC.18.0.22.5) (serum), triacylglycerol (TAG54.1.FA20.0) (serum), ethylalcohol (urine), glycochenodeoxycholate (serum), creatine (urine), lysophosphatidylcholine (LPC.22.5) (serum), glutamine (serum), 4-hydroxybenzoic acid (serum), triacylglycerol (TAG50.4.FA14.1) (serum), phosphatidylcholine (PC.18.1.20.2) (serum), triacylglycerol (TAG50.4.FA18.2) (serum), and phosphatidylcholine (PC.18.1.22.5) (serum). Those meeting criteria for inclusion in the protein density biomarker were lysophosphatidylethanolamine (LPE.16.0) (serum), urea (urine), propanediol (urine), creatine (serum), 2-oxoisovalerate (serum), gentiobiose (urine), 2-hydroxyglutarate (serum), methylglycocholate (urine), urine nitrogen, phosphatidylcholine (PC.15.0.20.4) (serum), cholesteryl ester (CE.18.3) (serum), glutamine (serum), 2-hydroxybutyrate (serum), cholesteryl ester (CE.22.6) (serum), and creatine (urine).

[2] Baseline FFQ if in OS; year 1 FFQ if in DM trial.

[3] Cross-validated R^2 based on average of validation set R^2 values from 100 equal-sized random splits of the data set into test and validation components.

The estimated log-HR associations with log-fat density in these analyses were substantially linear. Specifically, only for bladder cancer with biomarker calibration, among 26 tests conducted, was there evidence ($P = 0.03$) for a nonzero quadratic coefficient in log-fat density.

The right side of Table 4 gives corresponding HRs after including FFQ estimates of log-sodium, log (1 + vegetable servings/d), log (1 + fruit servings/d), and log-fiber in the outcome models. There is some modest HR attenuation with these additions, but fat density associations with cancer outcomes are mostly retained. The log-HR associations with log-fat density were again close to linear with only bladder cancer with calibrated intakes as a possible exception.

TABLE 3
Percent of biomarker-assessed log-fat density variation explained intake (R^2) by linear regression on log-FFQ fat density and disease-specific covariates in a Nutrition and Physical Activity Assessment Study biomarker study ($n = 436$) conducted during 2007–2009 among Women's Health Initiative participants, all of whom were postmenopausal and in the age range 50–79 y when enrolled during 1994–1998 at 9 United States clinical centers

Covariate set	Regression R^2 values (%)	Adjusted R^2 values (%)[1]
Cancer		
Log(FFQ)	8.5	16.0
Total[2]	25.8	48.8
CVD		
Log(FFQ)	7.4	14.0
Total	21.9	41.3
Diabetes		
Log(FFQ)	7.5	14.2
Total	22.5	42.4

Abbreviation: R^2, percent of variation explained.

[1] Adjustment by dividing R^2 values by the upper 90% percentile of the estimated paired correlation between biomarker values for the 14 participants in both NPAAS and NPAAS-FS.

[2] Total R^2 from linear regression of biomarker log-fat density on log-FFQ fat density and participant characteristics selected in model building.

Table 5 (left side) gives CVD HRs (95% CIs) for a 20% increment in fat density. Significant elevations after biomarker calibration (left side) are observed for nonfatal MI, total CHD, ischemic stroke, total stroke, combined CHD and stroke, and total CVD. The significance levels are again almost identical in the absence of calibration of fat density, but the HRs are substantially attenuated toward the null. For example, the HR (95% CI) for total CHD is 1.13 (1.02, 1.26) with fat density calibration, compared with 1.03 (1.01, 1.06) for FFQ fat density. The association of log-HR with log-fat density was again close to linear, with or without calibration, with only hemorrhagic stroke as a possible exception.

In sharp contrast to the cancer analyses, the CVD HRs in Table 5 all cease to be significantly related to fat density when the other 4 dietary variables are added to the disease risk model (Table 5, right side). Further analyses (not shown) indicate that HR changes from the left to the right side of the table are almost completely explained by the addition of fiber to the disease model.

As shown on the left side of Table 6, a 20% increment in calibrated fat density corresponded to an estimated 19% higher T2D incidence, as compared with an estimated 4% higher incidence for a 20% increment in FFQ fat density, with these elevations being highly significant. The inclusion of the other 4 dietary variables reduced this to an estimated 13% risk elevation with calibration, and 3% without calibration, which remain highly significantly elevated. There was little evidence of departure from linear HR models in relation to log-fat density in these T2D analyses.

Discussion

Previously described [10] biomarkers for protein and carbohydrate densities, based primarily on metabolomics profiles in

TABLE 4
Cancer incidence HRs and 95% CIs for a 20% increment in fat density, with and without biomarker calibration of FFQ assessments, and with and without the inclusion of FFQ assessment for additional dietary variables (sodium, vegetable servings, fruit servings, and fiber) in the disease risk models in analyses that include total EI, in Women's Health Initiative cohorts ($n = 81,894$) of postmenopausal United States women enrolled during 1993–1998 at 40 United States clinical centers and followed through December 2020

Cancer site (participants with events)	Fat density and total energy as dietary variables				With additional dietary variables			
	With biomarker calibration		Without biomarker calibration		With biomarker calibration		Without biomarker calibration	
	HR (95% CI)[1]	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Breast (5311)	1.16 (1.06, 1.27)	0.001	1.03 (1.01, 1.05)	0.001	1.11 (1.00, 1.24)	0.06	1.02 (1.00, 1.05)	0.06
Colon (1101)	1.26 (1.04, 1.53)	0.02	1.05 (1.01, 1.10)	0.02	1.31 (1.04, 1.66)	0.02	1.06 (1.01, 1.12)	0.03
Rectum (162)	0.82 (0.50, 1.34)	0.43	0.96 (0.86, 1.07)	0.44	1.06 (0.59, 1.88)	0.85	1.01 (0.88, 1.16)	0.86
Endometrium (916)	1.06 (0.84, 1.33)	0.65	1.01 (0.97, 1.06)	0.63	1.13 (0.84, 1.50)	0.42	1.03 (0.97, 1.09)	0.38
Ovary (479)	0.88 (0.66, 1.17)	0.37	0.97 (0.91, 1.04)	0.38	1.25 (0.88, 1.78)	0.22	1.05 (0.97, 1.13)	0.22
Leukemia (456)	1.78 (1.30, 2.42)	<0.001	1.13 (1.06, 1.21)	<0.001	1.75 (1.19, 2.59)	0.005	1.13 (1.04, 1.23)	0.005
Lung (1500)	1.00 (0.84, 1.18)	0.98	1.00 (0.97, 1.04)	0.98	0.81 (0.67, 0.99)	0.04	0.96 (0.92, 1.00)	0.04
Lymphoma (852)	1.32 (1.06, 1.64)	0.01	1.06 (1.01, 1.11)	0.02	1.32 (1.00, 1.75)	0.05	1.06 (1.01, 1.13)	0.05
Bladder (179)	1.01 (0.60, 1.70)	0.97	1.00 (0.96, 1.11)	0.97	1.34 (0.69, 2.61)	0.39	1.06 (0.93, 1.21)	0.35
Kidney (326)	1.09 (0.76, 1.58)	0.63	1.02 (0.94, 1.11)	0.64	0.85 (0.55, 1.31)	0.46	0.97 (0.88, 1.06)	0.48
Pancreas (433)	0.95 (0.71, 1.27)	0.71	0.99 (0.92, 1.06)	0.72	0.82 (0.58, 1.16)	0.26	0.96 (0.88, 1.04)	0.28
Obesity-related[2] (7563)	1.14 (1.06, 1.23)	<0.001	1.03 (1.01, 1.05)	<0.001	1.11 (1.01, 1.22)	0.02	1.02 (1.00, 1.04)	0.02
Total Invasive (13,290)	1.11 (1.05, 1.18)	<0.001	1.02 (1.01, 1.04)	<0.001	1.07 (1.10, 1.14)	0.06	1.01 (1.00, 1.03)	0.06

[1] HR estimates and 95% CIs are based on Cox models with baseline hazard rates stratified on study component (DM-C or OS), hormone therapy trial status (estrogen plus progestin, estrogen plus progestin placebo, estrogen-alone, estrogen-alone placebo, not randomized), age at enrollment (50–54, 55–59, 60–64, 65–69, 70–74, and ≥75 y), and race/ethnicity, and with adjustment for a disease-specific set of potential confounding factors.

[2] Obesity-related cancer defined here as breast, colon, rectum, endometrium or kidney cancer.

serum and urine, and self-reported alcohol intake, led to a novel biomarker for the assessment of fat density having a correlation (cross-validated R values) of 0.60 with feeding study intakes for these variables. Informal sensitivity and specificity justifications for this biomarker derive from those for the protein, carbohydrate, and total energy biomarkers employed [10]. Biomarker equations developed in the NPAAS-FS ($n = 153$) were used to calculate biomarker values in the NPAAS cohort ($n = 436$) and these were used to develop calibration equations for fat density that adjust FFQ intakes for measurement error while also using other participant characteristics to strengthen the assessment. The resulting calibration equations had adjusted R^2 values of 36% or more in each of the 3 clinical outcome settings, and these were used to calculate calibrated fat density intakes in WHI cohorts ($n = 81,954$) for use in disease association analyses.

The resulting HRs for a 20% increment in fat density, in analyses that also include DLW-calibrated total energy, can be considered as a conceptual substitution of fat for other energy sources, primarily carbohydrate. For example, the correlations (P values for test of zero correlation) in NPAAS between log (fat density) biomarker values is −0.08 (0.10) with log (1+ FFQ vegetable servings/d), −0.07 (0.17) with log (1+ FFQ fruit servings/d), and −0.12 (0.02) with log (fiber density). It follows that a lower fat density corresponds to a somewhat higher vegetable and fruit intake, and especially a higher fiber intake, similar to that for the low-fat dietary intervention group in the DM trial [1]. The left side of Tables 4–6 shows this increment to associate positively with breast and colorectal cancer incidence, CHD, stroke, and T2D among other outcomes. These HRs reflect a relatively high-fat dietary pattern at a specified total EI, with its attendant dietary correlates. For further insight into the role of fat density per se, the right side of Tables 4–6 simultaneously models the intakes of vegetable servings, fruit servings, fiber, and sodium. The CVD associations were no longer evident after controlling these additional dietary variables, especially after controlling for the fiber component of carbohydrate. The HRs for cancer outcomes and T2D were altered only modestly by these additions, with many significant associations, although the P value for breast cancer is increased to $P = 0.06$, after these additions. These results align with our previous publication [11] in which favorable disease risks at higher carbohydrate intake were largely explained by the fiber content of the carbohydrate for CVD outcomes, but not for cancer outcomes or T2D.

The biomarker calibration procedure used in these analyses involves measurement model assumptions for both the biomarker equation and the calibration equation. For the former, log-fat density in the NPAAS-FS is written as log-fat density biomarker plus random error that is assumed to be independent of the biomarker, a so-called Berkson measurement model, as befits the related linear regression model building. We allow the possibility of baseline characteristics, including baseline FFQ assessments, in the biomarker equation to avoid bias in subsequent disease association analyses that may otherwise occur [27] if the measurement model error component is substantial (that is, CV-R^2 is not large). Similarly, for the calibration equation, the log-fat density biomarker is written as calibrated log-FFQ fat density plus random error that is independent of this calibrated intake. It is a limitation of the calibration procedure that 2 measurement modeling specifications are needed. This compares, for example, to a potential case-control approach using biomarker intakes in larger WHI cohorts, which could completely avoid the calibration modeling component. Note, however, that the random error components in the 2 measurement models used here do not accumulate. Specifically, under the measurement model assumptions just mentioned, the utility of the calibrated log-fat density values is little affected by the magnitude of the error component in the biomarker equation,

TABLE 5
CVD incidence HRs and 95% CIs for a 20% increment in fat density, with and without biomarker calibration of FFQ assessments, and with and without the inclusion of FFQ assessment for additional dietary variables (sodium, vegetable servings, fruit servings, and fiber) in the disease risk models in analyses that include total EI, in Women's Health Initiative cohorts ($n = 81,894$) of postmenopausal United States women enrolled during 1993–1998 at 40 United States clinical centers and followed through December 2020

Outcome (participants with events)	Fat density and total energy as dietary variables				With additional dietary variables			
	With biomarker calibration		Without biomarker calibration		With biomarker calibration		Without biomarker calibration	
	HR (95% CI)[1]	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Nonfatal MI (2102)	1.17 (1.03, 1.32)	0.02	1.04 (1.01, 1.07)	0.02	1.01 (0.87, 1.17)	0.89	1.00 (0.97, 1.04)	0.89
Coronary death (897)	1.14 (0.94, 1.38)	0.20	1.03 (0.99, 1.08)	0.19	1.02 (0.81, 1.29)	0.87	1.01 (0.95, 1.06)	0.86
Total CHD (2869)	1.13 (1.02, 1.26)	0.02	1.03 (1.01, 1.06)	0.02	1.00 (0.88, 1.13)	0.96	1.00 (0.97, 1.03)	0.96
Ischemic stroke (1776)	1.19 (1.04, 1.36)	0.01	1.04 (1.01, 1.08)	0.01	1.16 (0.98, 1.37)	0.08	1.04 (1.00, 1.08)	0.08
Hemorrhagic stroke (395)	1.01 (0.79, 1.29)	0.94	1.00 (0.94, 1.07)	0.95	0.96 (0.70, 1.30)	0.78	0.99 (0.91, 1.08)	0.80
Total stroke (2425)	1.14 (1.02, 1.28)	0.03	1.03 (1.00, 1.06)	0.03	1.09 (0.95, 1.25)	0.22	1.02 (0.99, 1.06)	0.23
CHD + stroke (5023)	1.14 (1.05, 1.24)	0.001	1.03 (1.01, 1.05)	0.001	1.04 (0.95, 1.15)	0.37	1.01 (0.99, 1.04)	0.38
CABG + PCI (3119)	1.09 (0.98, 1.20)	0.08	1.02 (0.99, 1.05)	0.10	1.03 (0.91, 1.17)	0.61	1.01 (0.98, 1.04)	0.61
Total CVD[2] (6964)	1.12 (1.04, 1.19)	0.001	1.03 (1.01, 1.05)	0.002	1.03 (0.95, 1.12)	0.48	1.01 (0.99, 1.03)	0.48
Heart failure (1381)	1.16 (0.98, 1.36)	0.08	1.04 (1.00, 1.08)	0.08	1.08 (0.89, 1.31)	0.43	1.02 (0.97, 1.07)	0.42

Abbreviations: CABG/PCI, coronary artery bypass graft or percutaneous coronary intervention.
[1] HR estimates and 95% CIs are based on Cox models with baseline hazard rates stratified on study component (DM-C or OS), hormone therapy trial status (estrogen plus progestin, estrogen plus progestin placebo, estrogen-alone, estrogen-alone placebo, not randomized), age at enrollment (50–54, 55–59, 60–64, 65–69, 70–74, and ≥75 y), and race/ethnicity, and with adjustment for a disease-specific set of potential confounding factors.
[2] Total CVD comprised of CHD + CABG + PCI + stroke.

provided the feeding study used to generate this equation has adequate sample size.

An interesting feature of the Tables 4–6 results is the near identity of null hypothesis tests, whether or not fat density and total energy are biomarker calibrated. This presumably happens because the measurement error for the log-FFQ fat density assessment is close to a classical additive error model with measurement error primarily causing attenuation of HR estimates. In fact, the fat density HRs without calibration are substantially attenuated toward the null compared with the biomarker-calibrated HRs, as is to be expected with a correlation of only 0.27–0.29, depending which covariate set is used (Table 3), between biomarker-calibrated and FFQ log-fat density assessments (all $P < 0.0001$) suggesting a quite large random error component for the log-FFQ assessments. Also, the agreement between null hypothesis *P* values with or without biomarker calibration makes it unlikely that our biomarker assessment is attended by any serious lack of sensitivity or specificity, because such a lack would be expected to result in a weakened null hypothesis tests using the biomarker-calibrated intake procedure.

Taken together, the left side of Tables 4–6 suggests practically important chronic disease risk elevations for a conceptual isocaloric substitution of a relatively high-fat dietary pattern with reduced fat and fiber-containing higher carbohydrate diet in the WHI population. The CVD risk elevations are evidently largely attributable to a lower intake of fiber. When expressed instead as a substitution of fiber-containing carbohydrate for fat, the resulting low-fat dietary pattern resembles that implemented in the DM trial, giving the possibility of adding precision to, or otherwise extending, findings from the WHI randomized, controlled trial.

As noted above, the DM trial provides evidence concerning the chronic disease effects of an actual replacement of total fat (11% reduction) by carbohydrate (10% increase), and protein (1% increase), in conjunction with an increase in vegetables, fruit, and grains in this same study population. The left side of Figure 2 shows projected HRs (95% CIs) under the biomarker-calibrated analyses on the left side of Tables 4–6 for an 11% reduction in fat density, whereas the right side shows corresponding HRs (95% CIs) from intention-to-treat comparisons between randomized groups during the intervention phase of the

TABLE 6
Type 2 diabetes disease incidence HRs and 95% CIs for a 20% increment in fat density, with and without biomarker calibration of FFQ assessments, and with and without the inclusion of FFQ assessment for additional dietary variables (sodium, vegetable servings, fruit servings, and fiber) in the disease risk models in analyses that include total EI, in Women's Health Initiative cohorts ($n = 81,894$) of postmenopausal United States women enrolled during 1993–1998 at 40 United States clinical centers and followed through December 2020

Outcome (participants with outcome)	Fat density and total energy as dietary variables				With additional dietary variables			
	With biomarker calibration		Without biomarker calibration		With biomarker calibration		Without biomarker calibration	
	HR (95% CI)[1]	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
T2D (12,605)	1.19 (1.13, 1.26)	<0.001	1.04 (1.03, 1.06)	<0.001	1.13 (1.06, 1.21)	<0.001	1.03 (1.02, 1.05)	<0.001

[1] HR estimates and 95% CIs are based on Cox models with baseline hazard rates stratified on study component (DM-C or OS), hormone therapy trial status (estrogen plus progestin, estrogen plus progestin placebo, estrogen-alone, estrogen-alone placebo, not randomized), age at enrollment (50–54, 55–59, 60–64, 65–69, 70–74, and ≥75 y), and race/ethnicity, and with adjustment for a disease-specific set of potential confounding factors.

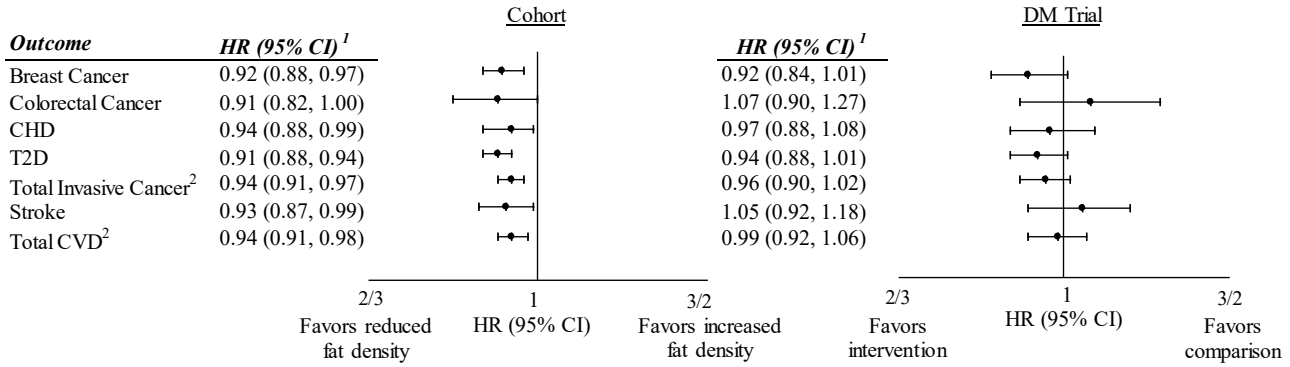

FIGURE 2. Comparison of HRs for 11% lower total fat density from calibrated intake cohort analyses with randomized group contrasts during the intervention phase of the WHI low-fat Dietary Modification (DM) trial. [1]For each outcome, cohort component HR estimates and 95% CIs are based on Cox models with baseline hazard rates stratified on study component (DM-C or OS), hormone therapy trial status (estrogen plus progestin, estrogen plus progestin placebo, estrogen-alone, estrogen-alone placebo, not randomized), age at enrollment (50–54, 55–59, 60–64, 65–69, 70–74, and ≥75 y), and race/ethnicity, and with adjustment for a disease-specific set of potential confounding factors that does not include BMI. [2]Total invasive (cancer) includes all incident cancers except nonmelanoma skin cancer; total CVD comprised of CHD + CABG + PCI + stroke. CABG/PCI, coronary artery bypass graft or percutaneous coronary intervention; DM-C, dietary modification comparison group; OS, Observational Study; T2D, type 2 diabetes; WHI, Women's Health Initiative.

DM trial [1,8]. Results are generally consistent from the 2 sources, with identical HR estimates for breast cancer incidence. Complete agreement between these sources should not be expected for multiple reasons: the DM trial, although large with a lengthy intervention period, still has considerable imprecision in HR estimates, for example for colorectal cancer. Also, CVD HRs in the (unblinded) DM trial may incorporate bias toward the null, because statin use increased rapidly during the trial intervention phase, and there was evidence of related postrandomization confounding [8]. Also, the time interval from dietary change to full DM trial intervention effect is uncertain for these outcomes, and estimated randomized group dietary differences declined somewhat over the intervention period and these were based on FFQs without measurement error correction. Overall, Figure 2 can be viewed as suggesting substantial agreement between randomized trial and observational cohort results for a low-fat dietary pattern. The cohort component, with its large numbers of cases and biomarker-calibrated intakes, serves to add precision to and extend DM trial findings. These sources, in conjunction with randomized trial evidence of some benefits, especially concerning long-term breast cancer mortality [4], and lack of any clear health risks, combine to indicate favorable chronic disease health benefits compared with risks for the low-fat dietary pattern intervention studied in the WHI population.

The ability to attribute DM trial low-fat dietary pattern findings to fat intake per se on the right sides of Tables 4–6 is limited by the reliance on FFQ assessments, which could be differentially biased between randomization groups. Note, however, there was a significant reduction [8] in HDL-C in the DM intervention group at year 1, as is typical for low-fat diet interventions. LDL-C was also reduced among intervention group women, except among the small subsets with prior CVD where postrandomization confounding by statin use may have intervened [8].

It is interesting to compare DM trial results with those of the low-fat diet intervention implemented in the Canadian breast cancer prevention trial (n = 4,690) by Martin et al. [28] which also showed reduction in HDL-C. Breast cancer results were nonsignificantly in the direction of higher risk. Note, however, that this trial intervention focused on fat reduction only with a more stringent fat reduction goal (to 15% of energy). Furthermore, trial participants had extensive mammographic density (>50%), and most participants (73%) were premenopausal.

Strengths of the present study include the development of a novel fat density biomarker in an embedded feeding study context, and the application of this biomarker for intake assessment in large well-characterized WHI cohorts having extensive quality follow-up for outcome ascertainment. Importantly, our study also serves to integrate these OS findings with those from the randomized, controlled WHI DM trial of a low-fat dietary pattern intervention.

Limitations include the observational nature of the low-fat dietary pattern association results presented, which will benefit from further evaluation in other settings. Also, the measurement error modeling is rather complex, involving both biomarker and calibration equation development. Application of these or other measurement error correction approaches in other large cohorts will be valuable. Also, the work is conducted in postmenopausal United States women enrolled during 1993–1998 and results in other populations could differ. Similarly results could differ in more recently enrolled cohorts of postmenopausal United States women in view of changes over the past 30 y in the United States food supply, including reductions in trans-FAs and increases in olive oil intake. Additional analyses taking advantage of longitudinal dietary data over WHI cohort follow-up may provide further insight, but are not considered here for reasons of space. Additional study of dietary adherence in the DM trial using the fat density biomarker proposed here will also be helpful, although requisite metabolomics data are not yet available for this purpose.

In summary, analyses in WHI cohorts using biomarker-based dietary assessment provide evidence additional to that from the WHI DM trial that a change to a low-fat dietary pattern, including increases in fiber-rich vegetables, fruits, and grains, led to risk reductions of public health importance for breast cancer, total invasive cancer, CHD, and diabetes, among other outcomes in the WHI population of postmenopausal United States women

enrolled during 1993–1998. Analyses that control for additional self-reported dietary variables, including fiber intake, lead to an association of somewhat reduced magnitude for breast cancer, whereas that for CHD is no longer evident.

Funding

The WHI program is funded by the NHLBI, NIH, United States Department of Health and Human Services through contracts 75N92021D00001, 75N92021D00002, 75N92021D00003, 75N92021D00004, 75N92021D00005; National Institute for Diabetes, Digestive and Kidney Diseases grant P30DK035816; National Cancer Institute grants R01 CA119171 and P30 CA15704, and NIH instrumentation grant S10 OD021562. Decisions concerning study design, data collection and analysis, interpretation of the results, the preparation of the manuscript, and the decision to submit the manuscript for publication resided with committees comprised of WHI investigators that included NHLBI representatives. The contents of the paper are solely the responsibility of the authors.

Author disclosures

RLP, SV, LFT, SLN, DR, GANG, MP, AKA, JWL, YH, LVH, JEM, RW, YM-R, JW-W, SL, LS, BVH, and CZ, no conflicts of interest. RTC reports receiving personal fees from Novartis, AstraZeneca, Pfizer Amgen, Genentech, Puma, and Up-To-Date. MLN is an Associate Editor on the *Journal of Nutrition* and played no role in the Journal's evaluation of the manuscript.

Author contributions

The authors' responsibilities were as follows – RLP, MLN, DR, CZ, GANG, YH, LFT, JWL: designed the research; RLP, MP, MLN, DR, GANG, CZ, LFT, YH, JLW: conducted the research, and drafted the manuscript; all authors: participated actively in critical evaluation of the manuscript, and read and approved the final manuscript; and RLP: had primary responsibility for the final content.

Data availability

Data, codebook, and analytic code used in this report may be accessed in a collaborative mode as described on the Women's Health Initiative website (www.whi.org).

Acknowledgments

We acknowledge the following investigators in the WHI program:

Program Office: (NHLBI, Bethesda, MD) Jacques Rossouw, Shari Ludlam, Joan McGowan, Leslie Ford, and Nancy Geller.

Clinical Coordinating Center: (Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, Ross L. Prentice, Andrea LaCroix, and Charles Kooperberg.

Investigators and Academic Centers: (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn E. Manson; (MedStar Health Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Jennifer Robinson; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; (University of Nevada, Reno, NV) Robert Brunner.

Women's Health Initiative Memory Study: (Wake Forest University School of Medicine, Winston-Salem, NC) Mark Espeland.

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tjnut.2023.05.021.

References

[1] R.L. Prentice, A.K. Aragaki, B.V. Howard, R.T. Chlebowski, C.A. Thomson, L. Van Horn, et al., Low-fat dietary pattern among postmenopausal women influences long-term cancer, cardiovascular disease, and diabetes outcomes, J. Nutr. 149 (2019) 1565–1574.

[2] R.T. Chlebowski, A.K. Aragaki, G.L. Anderson, C.A. Thomson, J.E. Manson, M.S. Simon, et al., Low-fat dietary pattern and breast cancer mortality in the Women's Health Initiative randomized controlled trial, J. Clin. Oncol. 35 (2017) 2919–2926.

[3] R.L. Prentice, B. Caan, R.T. Chlebowski, R. Patterson, L.H. Kuller, J.K. Ockene, et al., Low-fat dietary pattern and risk of invasive breast cancer: the Women's Health Initiative randomized controlled Dietary Modification trial, JAMA 295 (2006) 629–642.

[4] S.A.A. Beresford, K.C. Johnson, C. Ritenbaugh, N.L. Lasser, L.G. Snetselaar, H.R. Black, et al., Low-fat dietary pattern and risk of colorectal cancer: the Women's Health Initiative randomized controlled Dietary Modification trial, JAMA 295 (2006) 643–654.

[5] B.V. Howard, L. Van Horn, J. Hsia, J.E. Manson, M.L. Stefanick, S. Wassertheil-Smoller, et al., Low-fat dietary pattern and risk of cardiovascular disease: the Women's Health Initiative randomized controlled Dietary Modification trial, JAMA 295 (2006) 655–666.

[6] R.T. Chlebowski, A.K. Aragaki, G.L. Anderson, K. Pan, M.L. Neuhouser, J.E. Manson, et al., Dietary modification and breast cancer mortality: long-term follow-up of the Women's Health Initiative randomized trial, J. Clin. Oncol. 38 (2020) 1419–1428.

[7] B.V. Howard, A.K. Aragaki, L.F. Tinker, M. Allison, M.D. Hingle, K.C. Johnson, et al., A low-fat dietary pattern and diabetes: a secondary analysis from the Women's Health Initiative Dietary Modification trial, Diabetes Care 41 (2018) 680–687.

[8] R.L. Prentice, A.K. Aragaki, L. Van Horn, C.A. Thomson, S.A.A. Beresford, J. Robinson, et al., Low-fat dietary pattern and cardiovascular disease: results from the Women's Health Initiative randomized controlled trial, Am. J. Clin. Nutr. 106 (2017) 35–43.

[9] R.L. Prentice, B.V. Howard, L. Van Horn, M.L. Neuhouser, G.L. Anderson, L.F. Tinker, et al., Nutritional epidemiology and the Women's Health Initiative: a review, Am. J. Clin. Nutr. 113 (2021) 1083–1092.

[10] R.L. Prentice, M. Pettinger, M.L. Neuhouser, D. Raftery, C. Zheng, G.A.N. Gowda, et al., Biomarker-calibrated macronutrient intake and chronic disease risk among postmenopausal women, J. Nutr. 151 (2021) 2330–2341.

[11] R.L. Prentice, M. Pettinger, C. Zheng, M.L. Neuhouser, D. Raftery, G.A.N. Gowda, et al., Biomarkers for components of dietary protein and carbohydrate with application to chronic disease risk in postmenopausal women, J. Nutr. 152 (2022) 1107–1117.

[12] Women's Health Initiative Study Group, Design of the Women's Health Initiative clinical trial and observational study, Control, Clin. Trials. 19 (1998) 61–109.

[13] R.E. Patterson, A.R. Kristal, L.F. Tinker, R.A. Carter, M.P. Bolton, T. Agurs-Collins, Measurement characteristics of the Women's Health Initiative food frequency questionnaire, Ann. Epidemiol. 9 (1999) 178–187.
[14] M.L. Neuhouser, L. Tinker, P.A. Shaw, D. Schoeller, S.A. Bingham, L. Van Horn, et al., Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative, Am. J. Epidemiol. 167 (2008) 1247–1259.
[15] R.L. Prentice, Y. Mossavar-Rahmani, Y. Huang, L. Van Horn, S.A. Beresford, B. Caan, et al., Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment using recovery biomarkers, Am. J. Epidemiol. 174 (2011) 591–603.
[16] D.A. Schoeller, J.M. Hnilicka, Reliability of the doubly labeled water method for the measurement of total daily energy expenditure in free-living subjects, J. Nutr. 126 (1996) 348S–354S.
[17] S.A. Bingham, Urine nitrogen as a biomarker for the validation of dietary protein intake, J. Nutr. 133 (2003) 921S–924S.
[18] J.W. Lampe, Y. Huang, M.L. Neuhouser, L.F. Tinker, X. Song, D.A. Schoeller, et al., Dietary biomarker evaluation in a controlled feeding study in women from the Women's Health Initiative cohort, Am. J. Clin. Nutr. 105 (2017) 466–475.
[19] C. Zheng, G.A.N. Gowda, D. Raftery, M.L. Neuhouser, L.F. Tinker, R.L. Prentice, et al., Evaluation of potential metabolomics-based biomarkers of protein, carbohydrate and fat intakes using a controlled feeding study, Eur. J. Nutr. 60 (2021) 4207–4218.
[20] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. B. 58 (1996) 267–288.
[21] J.D. Curb, A. McTiernan, S.R. Heckbert, C. Kooperberg, J. Stanford, M. Nevitt, et al., Outcomes ascertainment and adjudication methods in the Women's Health Initiative, Ann. Epidemiol. 13 (2003) S122–S128.
[22] K.L. Margolis, L. Qi, R. Brzyski, D.E. Bonds, B.V. Howard, S. Kempainen, et al., Validity of diabetes self-reports in the Women's Health Initiative: comparison with medication inventories and fasting glucose measurements, Clin. Trials. 5 (2008) 240–247.
[23] D.R. Cox, Regression analysis and life tables (with discussion), J. R. Stat. Soc. B. 34 (1972) 187–220.
[24] R.L. Prentice, Covariate measurement errors and parameter estimation in a failure time regression model, Biometrika 69 (1982) 331–342.
[25] C.Y. Wang, L. Hsu, Z.D. Feng, R.L. Prentice, Regression calibration in failure time regression, Biometrics 53 (1997) 131–145.
[26] R.J. Carroll, D. Ruppert, L.A. Stefanski, C.M. Crainiceanu, Measurement Error in Nonlinear Models, A Modern Perspective, 2nd ed., Chapman and Hall/CRC, Boca Raton, FL, 2006.
[27] Y. Huang, C. Zheng, L.F. Tinker, M.L. Neuhouser, R.L. Prentice, Biomarker-based methods and study designs to calibrate dietary intake for assessing diet–disease associations, J. Nutr. 152 (2022) 899–906.
[28] L.J. Martin, Q. Li, O. Melnichouk, C. Greenberg, S. Minkin, G. Hislop, et al., A randomized trial of dietary intervention for breast cancer prevention, Cancer Res 71 (2011) 123–133.