nature medicine

Article

https://doi.org/10.1038/s41591-023-02235-5

Optimal dietary patterns for prevention of chronic disease

Received: 19 September 2022

Accepted: 26 January 2023

Published online: 13 March 2023

Check for updates

Peilu Wang[1,2] ✉, Mingyang Song[1,3,4,5], A. Heather Eliassen[1,3,6], Molin Wang[1,6,7], Teresa T. Fung[3,8], Steven K. Clinton[9], Eric B. Rimm[1,3,6], Frank B. Hu[1,3,6], Walter C. Willett[1,3], Fred K. Tabung[3,9,10] & Edward L. Giovannucci[1,3,10]

Multiple dietary patterns have been associated with different diseases; however, their comparability to improve overall health has yet to be determined. Here, in 205,852 healthcare professionals from three US cohorts followed for up to 32 years, we prospectively assessed two mechanism-based diets and six diets based on dietary recommendations in relation to major chronic disease, defined as a composite outcome of incident major cardiovascular disease (CVD), type 2 diabetes and cancer. We demonstrated that adherence to a healthy diet was generally associated with a lower risk of major chronic disease (hazard ratio (HR) comparing the 90th with the 10th percentile of dietary pattern scores = 0.58–0.80). Participants with low insulinemic (HR = 0.58, 95% confidence interval (CI) = 0.57, 0.60), low inflammatory (HR = 0.61, 95% CI = 0.60, 0.63) or diabetes risk-reducing (HR = 0.70, 95% CI = 0.69, 0.72) diet had the largest risk reduction for incident major CVD, type 2 diabetes and cancer as a composite and individually. Similar findings were observed across gender and diverse ethnic groups. Our results suggest that dietary patterns associated with markers of hyperinsulinemia and inflammation and diabetes development may inform on future dietary guidelines for chronic disease prevention.

Chronic diseases account for more than half of all premature deaths and >90% of yearly healthcare spending in the United States of America[1,2]. With 11 million deaths and 255 million disability-adjusted life-years globally attributable to poor nutrition, following a healthy diet can be a potentially cost-effective strategy for lowering the risk of chronic diseases[3,4]. Much existing diet-related research, however, focuses on specific foods and hence may not provide clear knowledge about the ideal diet for overall health. Dietary patterns, characterizing a variety of foods, nutrients and beverages, may serve as useful tools to represent the overall effects of diet on the risk of health outcomes.

Dietary patterns that emphasize high-quality foods, adherence to dietary recommendations and a focus on plant-based foods have been demonstrated to reduce the risk of cardiovascular disease (CVD), type 2 diabetes, cancer and all-cause mortality[5]. Based on this evidence, the 2015–2020 Dietary Guidelines for Americans propose healthy US-style, Mediterranean-style, vegetarian diets and Dietary Approaches to Stop

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [2]Department of Nutrition and Food Hygiene, School of Public Health, Institute of Nutrition, Fudan University, Shanghai, China. [3]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [4]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [5]Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [6]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [7]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [8]Department of Nutrition, Simmons University, Boston, MA, USA. [9]Division of Medical Oncology, Department of Internal Medicine, The Ohio State University College of Medicine and Comprehensive Cancer Center, Columbus, OH, USA. [10]These authors contributed equally: Fred K. Tabung, Edward L. Giovannucci. ✉e-mail: peiluwang@fudan.edu.cn

https://doi.org/10.1038/s41591-023-02235-5

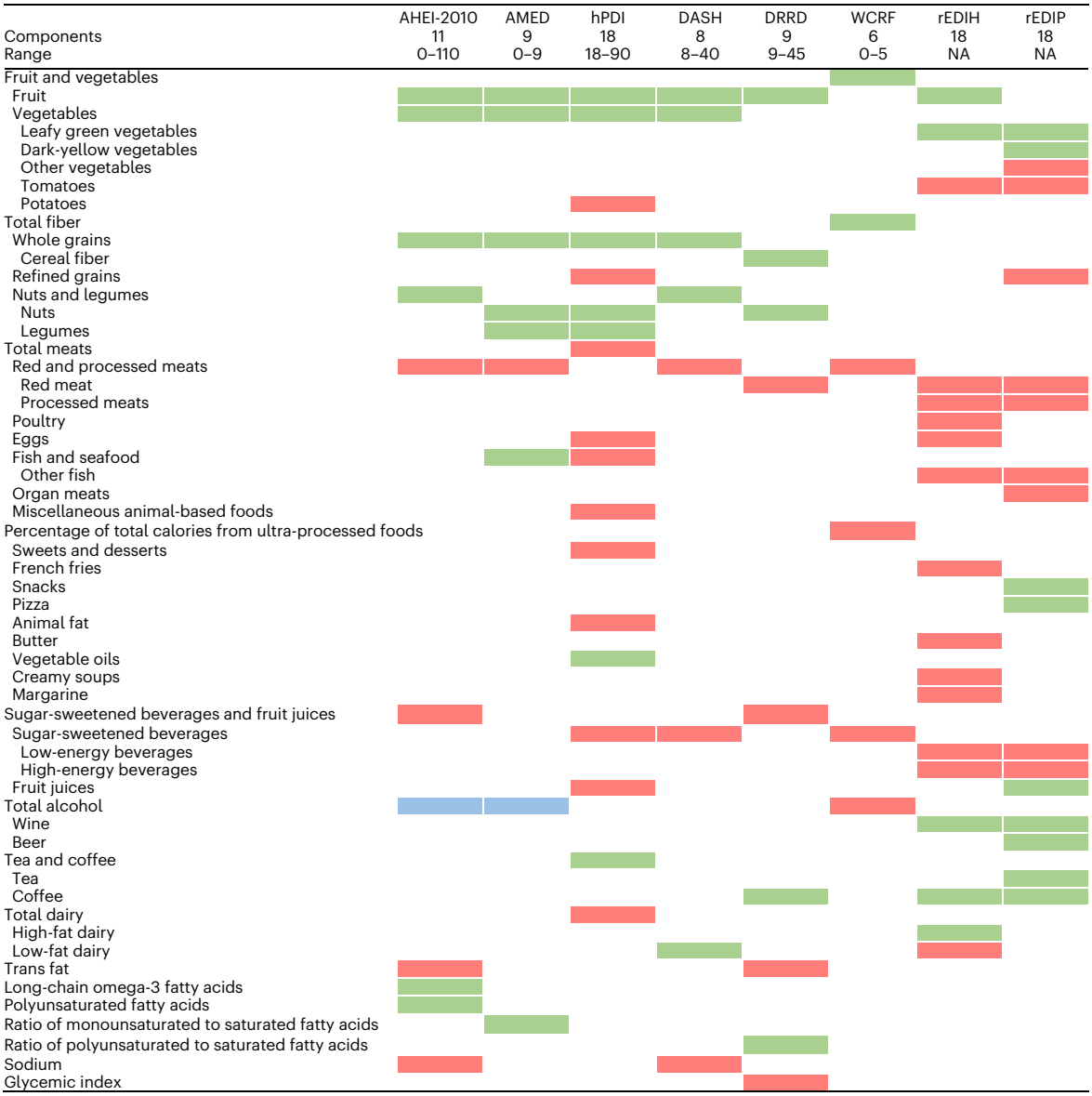


Fig. 1 | Detailed components of dietary patterns. Higher points or positive weights assigned to higher intakes of components are in green. Lower points or negative weights assigned to higher intakes of components are in red. Higher points assigned to moderate intakes of components are in blue. NA, not available.

Hypertension (DASH) as examples of healthy diets[6]. Despite this, only a few studies have explicitly examined these diets in the same context for their potential health impact.

The purpose of the present study is to compare the relative effectiveness of dietary patterns in improving general health by focusing on risk reduction of chronic diseases that contribute substantially to mortality in the UnitedStates of America, including CVD, cancer and diabetes. Using data from three US cohorts (Health Professionals Follow-up Study (HPFS), Nurses' Health Study (NHS) and NHS II) with up to 32 years of follow-up, we compared several dietary patterns that were promoted in the Dietary Guidelines for Americans and previously created for predicting major chronic diseases or important biological pathways for their associations with major chronic diseases (Fig. 1).

Results

Population characteristics

We evaluated the associations of eight dietary patterns with major chronic diseases as a composite and individually. Figure 1 shows the components of dietary patterns and Supplementary Table 1 shows the outcome definition. The present study included 162,667 women from the NHS and NHS II and 43,185 men from the HPFS (Extended Data Fig. 1). During a median follow-up of 26 years, we observed 44,975 major chronic disease events, 12,962 major CVDs, 18,615 cases of diabetes and 17,909 total cancers in the three cohorts.

Table 1 shows the age-standardized characteristics of the study population. Individuals in the highest quintile of dietary pattern scores were more likely to be older, exercise more, have lower body mass index (BMI), take multivitamins and use postmenopausal hormones (women), and were less likely to be current smokers. Individuals with the highest adherence to the reversed empirical dietary index for hyperinsulinemia (rEDIH) and reversed empirical dietary inflammatory pattern (rEDIP) drank more alcohol whereas those with the highest adherence to the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) dietary score drank less. Participants with higher scores on healthful Plant-based Diet Index (hPDI), Diabetes Risk Reduction Diet (DRRD), rEDIH and rEDIP consumed more coffee, whereas those with higher WCRF/AICR dietary scores consumed less.

https://doi.org/10.1038/s41591-023-02235-5

Table 1 | Age-standardized characteristics of the study population in the lowest and highest quintiles of energy-adjusted dietary patterns during the follow-up in the pooled data

Quintile	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5
Pattern	**AHEI-2010**		**AMED**		**hPDI**		**DASH**	
Median score[a]	39.0	65.0	2.3	6.4	47.2	62.5	17.9	29.1
Age[b] (years)	52.2 (11.6)	58.3 (11.1)	53.7 (11.4)	56.2 (11.7)	53.7 (11.6)	56.4 (11.7)	53.4 (11.1)	56.5 (12.0)
Family history of cancer (%)	35.9	43.1	39.3	39.2	39.7	39.1	38.8	39.3
Family history of diabetes (%)	24.0	28.7	26.6	25.8	26.3	26.8	26.2	25.8
Family history of CVD (%)	40.6	44.0	42.0	43.0	41.7	43.5	42.0	42.5
Physical activity (METS-h per week)	15.2 (17.1)	30.0 (27.0)	15.9 (17.6)	28.1 (25.6)	16.6 (17.8)	28.0 (26.8)	15.4 (17.1)	29.2 (27.0)
Height (cm)	167.3 (8.7)	166.9 (8.3)	167.2 (8.7)	167.0 (8.3)	167.3 (8.6)	166.9 (8.4)	167.0 (8.7)	167.2 (8.3)
BMI (kg m^{-2})	25.6 (4.8)	24.4 (4.0)	25.7 (4.9)	24.4 (3.9)	25.6 (4.8)	24.6 (4.1)	25.7 (4.9)	24.4 (4.0)
Alcohol consumption (g d^{-1})	5.9 (12.6)	6.8 (7.3)	5.8 (11.7)	6.7 (7.7)	5.2 (8.8)	6.7 (10.1)	6.7 (11.4)	5.3 (7.9)
Current smoking (%)	24.3	13.4	24.5	14.2	20.6	16.0	28.0	11.8
Regular aspirin use[c] (%)	26.5	23.2	25.2	24.9	25.2	24.7	25.6	23.8
Regular NSAID use[d] (%)	22.8	26.1	25.8	23.4	25.0	24.2	25.9	22.6
Multivitamin use (%)	42.5	57.2	43.6	56.0	45.0	54.9	40.8	58.0
Postmenopausal hormone use (%)	12.6	15.5	12.3	15.8	12.3	15.7	12.3	15.5
Coffee (cup per d)	1.8 (1.6)	1.9 (1.4)	1.9 (1.6)	1.8 (1.4)	1.5 (1.4)	2.2 (1.5)	1.9 (1.6)	1.8 (1.4)
Total energy intake (kcal d^{-1})	1836 (485)	1836 (497)	1824 (515)	1832 (445)	1841 (462)	1834 (508)	1839 (516)	1845 (451)
Pattern	**DRRD**		**WCRF/AICR**		**rEDIH**		**rEDIP**	
Median score[a]	20.6	33.4	1.5	3.2	−0.7	−0.2	−0.3	0.4
Age[b] (years)	53.6 (11.5)	56.4 (11.7)	52.9 (11.3)	56.9 (11.9)	52.4 (11.1)	57.6 (11.5)	53.9 (11.8)	55.6 (11.2)
Family history of cancer (%)	39.1	39.6	37.9	39.8	36.7	42.4	38.1	40.7
Family history of diabetes (%)	26.3	26.3	24.0	27.5	26.6	26.0	28.0	24.7
Family history of CVD (%)	41.7	43.6	41.7	42.7	41.6	42.9	42.5	42.6
Physical activity (METS-h per week)	15.9 (17.5)	28.5 (26.6)	16.6 (17.7)	27.7 (26.9)	17.4 (19.0)	28.0 (26.6)	18.8 (20.6)	24.8 (24.2)
Height (cm)	167.0 (8.7)	167.1 (8.3)	167.3 (8.6)	166.9 (8.4)	167.4 (8.8)	167.1 (8.2)	166.9 (8.7)	167.4 (8.4)
BMI (kg m^{-2})	25.7 (4.9)	24.5 (3.9)	25.0 (4.4)	25.0 (4.4)	26.7 (5.2)	23.7 (3.5)	26.5 (5.2)	24.2 (3.8)
Alcohol consumption (g d^{-1})	5.0 (9.3)	6.6 (9.4)	9.3 (11.7)	3.4 (6.6)	4.5 (8.7)	10.0 (12.0)	3.2 (7.0)	12.1 (13.5)
Current smoking (%)	21.2	15.0	26.0	12.4	21.3	17.2	17.6	22.5
Regular aspirin use[c] (%)	24.8	25.2	26.5	23.2	26.6	23.9	25.6	26.1
Regular NSAID use[d] (%)	24.9	24.3	24.9	23.1	25.2	25.5	25.2	26.2
Multivitamin use (%)	43.2	56.5	44.2	55.9	43.5	55.6	45.5	53.1
Postmenopausal hormone use (%)	12.5	15.9	13.0	15.1	13.2	14.5	13.4	14.6
Coffee (cup per d)	1.2 (1.3)	2.3 (1.5)	2.0 (1.5)	1.7 (1.5)	1.3 (1.4)	2.5 (1.6)	1.0 (1.1)	2.9 (1.6)
Total energy intake (kcal d^{-1})	1825 (499)	1827 (484)	1806 (466)	1801 (499)	1915 (534)	1923 (487)	1893 (531)	1889 (496)

METS, metabolic equivalent for task score; NSAID, nonsteroidal anti-inflammatory drug. Values are means (s.d.) for continuous variables and percentages for categorical variables if not specified otherwise. [a]Energy-adjusted scores are shown. [b]All variables are standardized to the age distribution of the study population, except for age. [c]Regular users are defined as participants who take at least two tablets of aspirin (325 mg per tablet) per week in the NHS and at least twice a week in the HPFS and NHS II. [d]Regular users are defined as participants who take at least twice a week.

Correlations between dietary patterns were comparable across cohorts (Extended Data Fig. 2). The Alternative Healthy Eating Index–2010 (AHEI-2010), Alternate Mediterranean Diet (AMED), DASH, hPDI, DRRD and WCRF/AICR dietary scores strongly correlated with each other, with Spearman's correlation coefficients ranging from 0.45 to 0.76 in the pooled data. The rEDIH and rEDIP scores had relatively lower correlations with other dietary patterns.

As the point estimates in the pooled data and random effects meta-analysis based on three cohorts are almost identical (Supplementary Table 2), we presented the associations obtained in the pooled data for all analyses. Although the point estimates for the associations between patterns and outcomes varied by cohort, the general direction and magnitude ranking for the associations were similar.

Dietary patterns and major chronic diseases

Table 2 shows the hazard ratios (HRs) for major chronic disease (the composite of major CVD, type 2 diabetes and total cancer) comparing the 90th (high adherence) with the 10th (low adherence) percentile scores for each dietary pattern. The multivariable (MV)-adjusted associations were strongest for the rEDIH (HR = 0.58, 95% CI = 0.57, 0.60), rEDIP (HR = 0.61, 95% CI = 0.60, 0.63) and DRRD (HR = 0.70, 95% CI = 0.69, 0.72). Inverse associations with major chronic disease

https://doi.org/10.1038/s41591-023-02235-5

Table 2 | Associations of cumulative average dietary patterns (comparing the 90th with the 10th percentile) with major chronic disease and its major components in the pooled data

Outcome		Major chronic disease	Major CVD	Type 2 diabetes	Total cancer
Cases		44,975	12,962	18,615	17,909
Person-years		4,852,894	5,148,378	4,630,725	5,138,951
Pattern	Model	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
AHEI-2010	Age	0.64 (0.62, 0.66)	0.63 (0.60, 0.66)	0.49 (0.47, 0.51)	0.81 (0.78, 0.84)
	MV[a]	0.76 (0.74, 0.78)	0.77 (0.73, 0.81)	0.62 (0.59, 0.64)	0.94 (0.90, 0.98)
	MV+nSES[b]	0.78 (0.76, 0.80)	0.79 (0.75, 0.83)	0.66 (0.63, 0.69)	0.93 (0.89, 0.97)
	MV+BMI[c]	0.83 (0.80, 0.85)	0.79 (0.75, 0.83)	0.73 (0.70, 0.77)	0.95 (0.91, 0.99)
AMED	Age	0.65 (0.64, 0.67)	0.64 (0.61, 0.67)	0.56 (0.53, 0.58)	0.78 (0.75, 0.81)
	MV[a]	0.79 (0.77, 0.81)	0.80 (0.76, 0.84)	0.71 (0.68, 0.74)	0.92 (0.89, 0.96)
	MV+nSES[b]	0.81 (0.78, 0.83)	0.82 (0.78, 0.86)	0.74 (0.71, 0.77)	0.92 (0.88, 0.96)
	MV+BMI[c]	0.86 (0.84, 0.88)	0.83 (0.79, 0.87)	0.85 (0.82, 0.89)	0.94 (0.90, 0.98)
hPDI	Age	0.70 (0.69, 0.72)	0.72 (0.69, 0.76)	0.57 (0.55, 0.59)	0.86 (0.83, 0.90)
	MV[a]	0.80 (0.78, 0.82)	0.84 (0.80, 0.88)	0.70 (0.67, 0.72)	0.94 (0.90, 0.98)
	MV+nSES[b]	0.80 (0.78, 0.83)	0.84 (0.80, 0.88)	0.71 (0.68, 0.73)	0.94 (0.90, 0.98)
	MV+BMI[c]	0.84 (0.82, 0.87)	0.85 (0.81, 0.90)	0.78 (0.75, 0.81)	0.95 (0.91, 0.99)
DASH	Age	0.63 (0.61, 0.64)	0.62 (0.59, 0.65)	0.52 (0.50, 0.54)	0.76 (0.73, 0.79)
	MV[a]	0.78 (0.76, 0.80)	0.81 (0.77, 0.85)	0.66 (0.64, 0.69)	0.94 (0.91, 0.99)
	MV+nSES[b]	0.79 (0.77, 0.81)	0.83 (0.79, 0.87)	0.69 (0.66, 0.72)	0.94 (0.90, 0.98)
	MV+BMI[c]	0.83 (0.81, 0.86)	0.83 (0.79, 0.87)	0.77 (0.74, 0.81)	0.96 (0.92, 1.00)
DRRD	Age	0.61 (0.59, 0.62)	0.62 (0.59, 0.65)	0.44 (0.43, 0.46)	0.83 (0.80, 0.87)
	MV[a]	0.70 (0.69, 0.72)	0.73 (0.69, 0.76)	0.56 (0.54, 0.58)	0.92 (0.88, 0.96)
	MV+nSES[b]	0.72 (0.70, 0.73)	0.74 (0.71, 0.78)	0.58 (0.56, 0.60)	0.92 (0.88, 0.96)
	MV+BMI[c]	0.76 (0.74, 0.78)	0.75 (0.71, 0.79)	0.66 (0.63, 0.69)	0.94 (0.90, 0.98)
WCRF/AICR	Age	0.84 (0.82, 0.86)	0.80 (0.77, 0.84)	0.86 (0.83, 0.90)	0.85 (0.81, 0.88)
	MV[a]	1.00 (0.98, 1.03)	1.00 (0.96, 1.05)	1.05 (1.01, 1.10)	1.01 (0.97, 1.05)
	MV+nSES[b]	1.01 (0.98, 1.04)	1.01 (0.96, 1.06)	1.07 (1.02, 1.11)	1.01 (0.97, 1.05)
	MV+BMI[c]	0.96 (0.94, 0.99)	0.99 (0.94, 1.03)	0.95 (0.91, 0.99)	1.00 (0.96, 1.04)
rEDIH	Age	0.53 (0.52, 0.54)	0.61 (0.59, 0.64)	0.30 (0.29, 0.31)	0.85 (0.82, 0.88)
	MV[a]	0.58 (0.57, 0.60)	0.68 (0.65, 0.71)	0.35 (0.34, 0.36)	0.90 (0.87, 0.94)
	MV+nSES[b]	0.59 (0.58, 0.61)	0.70 (0.66, 0.73)	0.36 (0.35, 0.37)	0.90 (0.86, 0.93)
	MV+BMI[c]	0.75 (0.73, 0.77)	0.76 (0.72, 0.79)	0.57 (0.54, 0.59)	0.95 (0.91, 0.99)
rEDIP	Age	0.59 (0.57, 0.60)	0.67 (0.64, 0.70)	0.33 (0.32, 0.34)	0.93 (0.90, 0.97)
	MV[a]	0.61 (0.60, 0.63)	0.69 (0.66, 0.72)	0.38 (0.36, 0.39)	0.90 (0.87, 0.94)
	MV+nSES[b]	0.62 (0.60, 0.63)	0.70 (0.67, 0.73)	0.38 (0.37, 0.40)	0.90 (0.87, 0.94)
	MV+BMI[c]	0.75 (0.73, 0.77)	0.75 (0.72, 0.79)	0.57 (0.55, 0.59)	0.94 (0.91, 0.98)

[a]Analyses were stratified by age (in month), calendar year and cohort. MV model: adjusted for physical activity (<3.0, 3.0–8.9, 9.0–17.9, 18.0–26.9, 27.0–41.9 or ≥42 MET-h per week), cigarette smoking status (never, former quitting ≥10 years, former quitting <10 years, current), cigarette smoking pack-years (0, 1–4, 5–14, 15–24 or ≥25 pack-years), multivitamin use (yes or no), regular aspirin use (yes or no), regular NSAID use (yes or no), postmenopausal hormone use (premenopausal, never, former or current use) for women and total energy intake (quintiles). Alcohol consumption (<5.0, 5.0–14.9 or ≥15.0 g d^{-1}) was adjusted for DASH, hPDI and DRRD. For major chronic disease, the model was additionally adjusted for family history of diabetes (yes or no), family history of cancer (yes or no) and family history of CVD (yes or no). For type 2 diabetes, the model was additionally adjusted for family history of diabetes (yes or no). For major CVD, the model was additionally adjusted for family history of CVD (yes or no). For total cancer, the model was additionally adjusted for family history of cancer (yes or no) and height (continuous). [b]Adjusted for nSES (continuous). [c]Adjusted for BMI (continuous).

were also observed for AHEI-2010, AMED, hPDI and DASH, with HRs between 0.76 and 0.80. In contrast, the WCRF/AICR dietary score was not associated with major chronic disease. Although the spline analysis demonstrated statistically significant nonlinearity for the AMED, WCRF/AICR dietary score and rEDIP, the curves were largely monotonic for the rEDIP and AMED (Extended Data Fig. 3) and the relationship between the WCRF/AICR dietary score and major chronic disease also appeared monotonic after the alcohol component was removed from the score (data not shown). Similar findings were observed when using pattern scores in quintiles as the exposure (Supplementary Table 3).

We next investigated the associations between dietary pattern scores and each individual component of the major chronic disease as presented in Fig. 2, Table 2 and Extended Data Table 1. In general, the strongest relationships among the various outcomes (CVDs, cancer and type 2 diabetes) were those between dietary patterns and type 2 diabetes. Among the dietary patterns, rEDIH, rEDIP and DRRD scores

https://doi.org/10.1038/s41591-023-02235-5

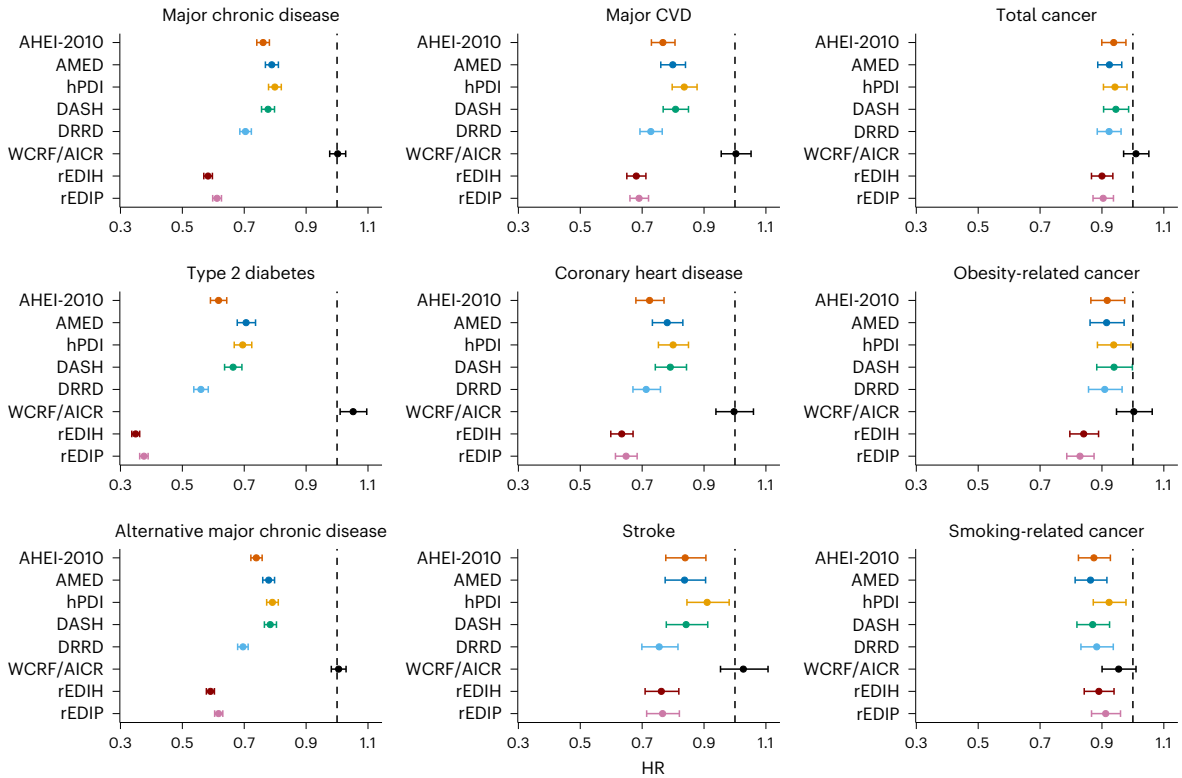


Fig. 2 | MV-adjusted associations of cumulative average dietary patterns (comparing the 90th with the 10th percentile) with major chronic disease and secondary outcomes in the pooled data of three cohorts (n = 205,852 participants). The analysis details and corresponding estimates are provided in Table 2 and Extended Data Table 1. The HRs (comparing the 90th with the 10th percentile) are indicated by the circles and the 95% CIs are reflected by the error bars.

had the strongest associations with type 2 diabetes (HR = 0.35–0.56). These three pattern scores were also among the top three patterns that were strongly associated with CVD-related outcomes, including major CVD, coronary heart disease and stroke (HR = 0.63–0.77). The estimates for AHEI-2010 were comparable with the DRRD score for coronary heart disease. The estimates for cancer-related outcomes were closer to null compared with type 2 diabetes or CVDs, for all dietary patterns. Participants with higher rEDIH, rEDIP or DRRD scores had a decreased risk of total cancer, obesity-related cancer and smoking-related cancer (HR = 0.83–0.92). For total cancer and smoking-related cancer, participants with higher AHEI-2010, AMED and DASH presented comparable or lower risk than those with higher rEDIH, rEDIP or DRRD scores. The WCRF/AICR dietary score was not associated with secondary outcomes except for a positive association with type 2 diabetes (HR =1.05, 95% CI = 1.01–1.10). As in the primary analysis, dietary patterns were similarly associated with an alternatively defined major chronic disease, which includes nontraumatic death in addition to major chronic disease incidence.

Figure 3 shows the associations between individual food groups with dietary patterns and major chronic diseases. The estimates for food groups were largely similar across major chronic diseases, although the associations were slightly stronger for type 2 diabetes. Coffee, whole grains, wine and desserts were inversely associated with major chronic disease, whereas processed meats, low-energy drinks, red meat, French fries, high-energy drinks and eggs were positively associated with major chronic disease. Compared with other patterns, the rEDIH, rEDIP and DRRD scores appeared to be more or at least similarly correlated with food groups that were associated with major chronic diseases (particularly type 2 diabetes), such as coffee, wine, processed meats, red meat, French fries and sugar-sweetened beverages. The rEDIH and rEDIP scores had relatively lower correlations with fruit and vegetables compared with other dietary patterns except for leafy green vegetables.

Subgroup and latency analyses

The associations between patterns and major chronic disease persisted in subgroups defined by age, BMI, gender, smoking status, alcohol intake, neighborhood socioeconomic status (nSES) and race/ethnicity (Fig. 4 and Extended Data Table 2). The inverse associations between dietary patterns and major chronic disease were generally stronger in participants who were younger, overweight or obese, women and Hispanic. Although WCRF/AICR exhibited an inverse association with major chronic disease in participants who drank more alcohol, the rEDIH and rEDIP showed stronger associations in participants who never smoked or drank less. Comparable associations with major chronic disease were observed in subgroups defined by nSES. With longer latency periods, the associations between dietary patterns and major chronic disease were slightly attenuated (Extended Data Fig. 4 and Extended Data Table 3). A similar trend was observed for major CVD and type 2 diabetes, but not for total cancer.

Sensitivity analyses

After removing the alcohol component from the dietary pattern scores, the associations for AHEI-2010, AMED, rEDIH and rEDIP were attenuated only slightly but remained largely similar (Methods and Extended Data Table 4). In contrast to the main results, the WCRF/AICR dietary score without the alcohol component was inversely associated with major chronic disease, major CVDs and type 2 diabetes (HR = 0.77–0.88). After removing the coffee component, the associations of DRRD, rEDIH and rEDIP with major chronic disease attenuated but remained strong (Extended Data Table 5). The HRs for the association of dietary patterns with major chronic disease and its components barely changed after

https://doi.org/10.1038/s41591-023-02235-5

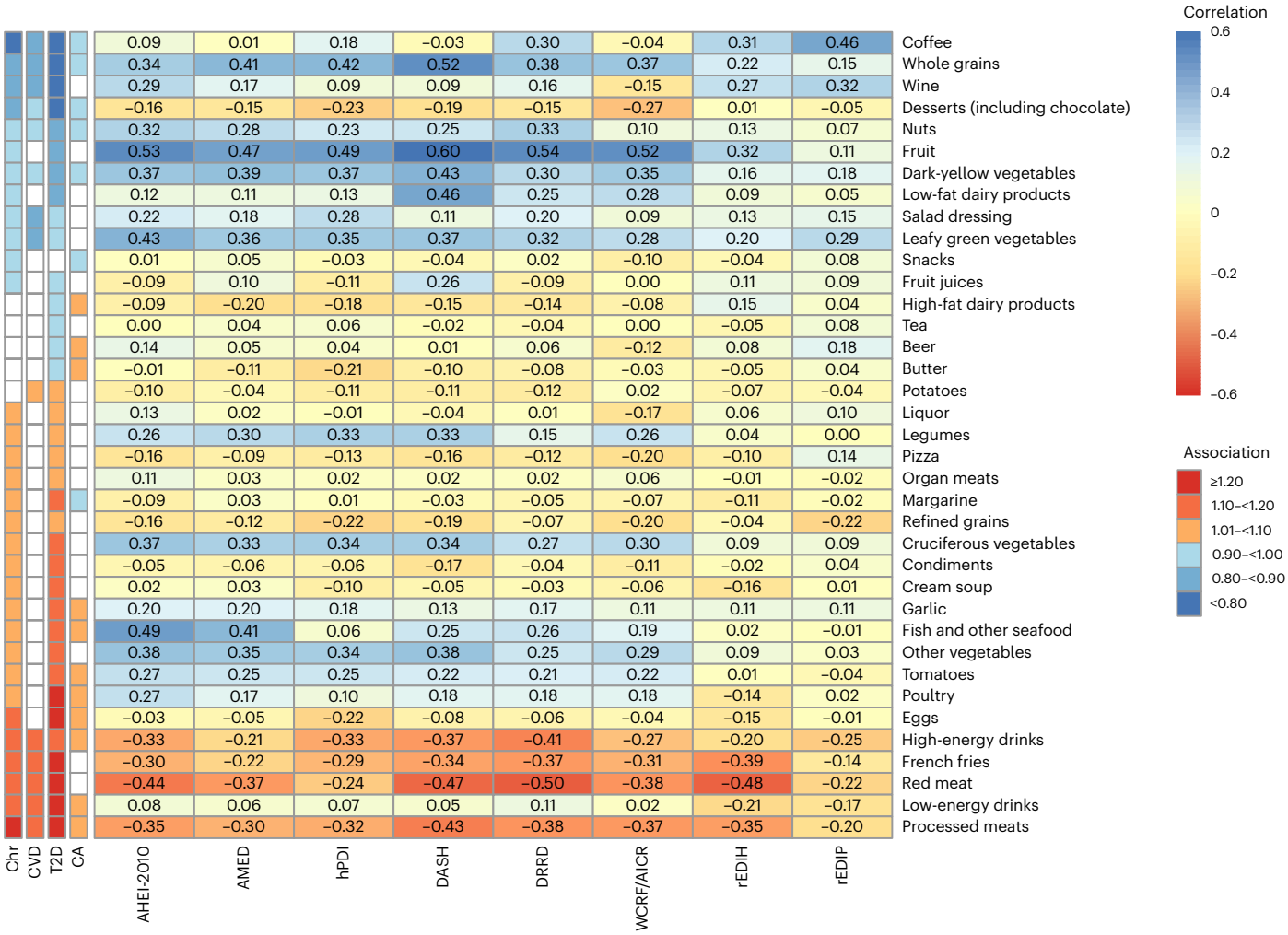


Fig. 3 | Baseline Spearman's correlations between energy-adjusted cumulative average dietary patterns and food groups in the pooled data of three cohorts (n = 205,852 participants). Spearman's correlation coefficients are shown and highlighted in color. Food groups are ordered based on the HRs of their associations with major chronic disease. The associations of food groups (comparing the 90th with the 10th percentile) with major chronic disease (Chr), major CVD, type 2 diabetes (T2D) or total cancer (CA) are indicated on the left of the figure. We reported unadjusted P values based on two-sided statistical tests. Significant associations ($P < 0.05$) are highlighted in color according to the magnitude of the HRs.

adjusting for the nSES score (HR change = 0.00–0.04) (Table 2 and Extended Data Table 1). After adjusting for the BMI, we observed an attenuated association with major chronic disease, especially for rEDIH and rEDIP, yet the rEDIH, rEDIP and DRRD scores remained showing the strongest associations (Table 2 and Extended Data Table 1). Additional sensitivity analyses showed that our findings remained unchanged (Supplementary Table 4). Compared with the same participants who were never in the highest quintiles of any pattern, the rEDIH, rEDIP and DRRD scrores remained the top three dietary patterns that were strongly associated with the risk of major chronic disease (Extended Data Table 6). The composite outcome in the pooled data did not meet the proportional hazards assumption, which can be related to effect modification by age ($P < 0.05$ for all patterns except WCRF/AICR). Thus, the results estimated in the pooled data should be interpreted as the average associations during the follow-up.

Discussion

The current understanding of dietary patterns on general health is based mostly on the synthesis of the literature for individual diseases rather than more composite outcomes of chronic diseases. Furthermore, the relationships between patterns and major chronic diseases have rarely been compared in the same study utilizing the same data collection and statistical analysis approach. In the present study, we compared two mechanism-based dietary patterns with six dietary patterns reflecting general or disease-specific dietary guidelines in 205,852 participants over the course of more than two decades. In general, adherence to a healthy diet was associated with a decreased risk of major chronic diseases. Participants who reported high adherence to low insulinemic, low inflammatory and diabetes risk-reducing dietary patterns displayed a decreased risk for major chronic diseases when examined as an individual or a composite outcome.

The rEDIH and rEDIP scores represent dietary patterns empirically constructed, based on foods that are either positively or inversely associated with biomarkers of two important and related biological pathways for chronic disease—hyperinsulinemia and chronic inflammation[7,8]. Assuming causal associations, the dietary patterns can be interpreted as individuals with high adherence to the rEDIH and rEDIP tending to have lower markers of insulin and inflammation, although these biomarkers could have other determinants. Therefore, strong associations were observed between these two patterns with chronic disease risk, which can be largely shaped by their connections to pathophysiological underpinnings of insulin resistance and type 2 diabetes. As shown in other studies, participants with low rEDIH or rEDIP scores had a higher risk of developing type 2 diabetes[9,10]. Despite this,

https://doi.org/10.1038/s41591-023-02235-5

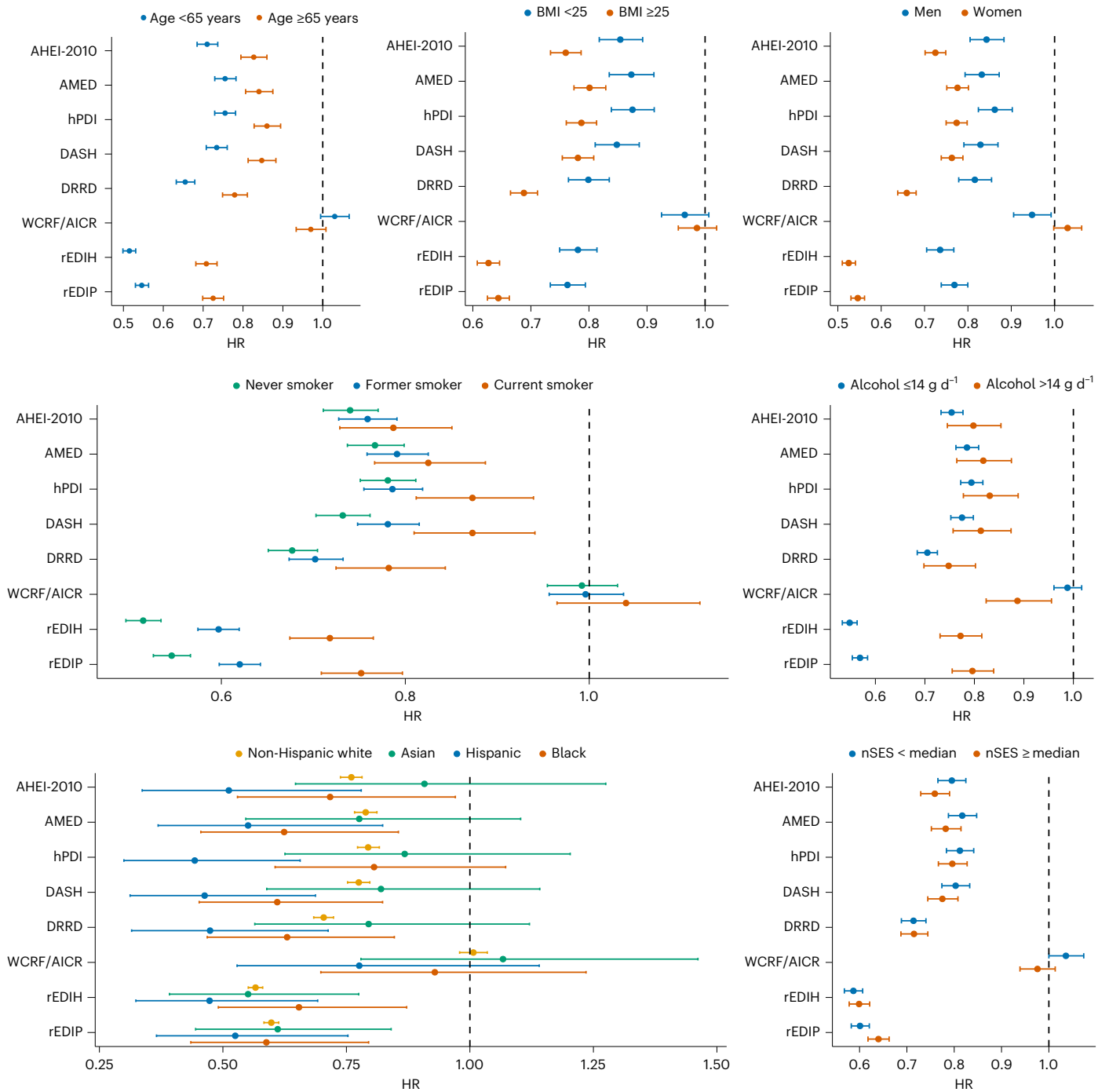


Fig. 4 | MV-adjusted associations between cumulative average dietary patterns (comparing the 90th with the 10th percentile) and major chronic disease in subgroups. Analysis details and corresponding estimates are provided in Extended Data Table 2. The HRs (comparing the 90th with the 10th percentile) are indicated by the circles and the 95% CIs are reflected by the error bars.

associations were observed between rEDIH or rEDIP and other elements of chronic diseases, which may be mediated by the same mechanistic pathways. In line with our findings, prior investigations in the NHS and HPFS showed an inverse association between rEDIP and CVD, as well as rEDIH and total mortality[11]. In addition, when compared with other patterns, the rEDIH and rEDIP showed greater associations with both composite and specific chronic diseases. Previous studies showed that these two diets are predictive of a diverse range of biomarkers for atherosclerosis, hyperlipidemia, hyperinsulinemia and systemic inflammation[12,13], supporting the hypothesis that rEDIH and rEDIP address fundamental biological pathways that are shared by multiple chronic diseases. The weighting of foods by a measured biological response may also account for differential measurement errors among specific foods.

Although not as strong as that for rEDIH or rEDIP, we observed an inverse association between DRRD and major chronic disease. Consisting of dietary components with sufficient evidence for type 2 diabetes, the DRRD has been shown to be associated with risk of type 2 diabetes across diverse racial and ethnic populations[14]. Participants with higher adherence to DRRD were associated with a lower risk of death from all

causes, CVD and cancer in a US-based study[15], which is compatible with our findings, indicating that the preventive role of DRRD may extend beyond type 2 diabetes to CVD, cancer and death. One possible explanation might be that a diet targeting type 2 diabetes captures dietary insulinemic potential that is etiologically crucial for developing a wide spectrum of chronic diseases.

Although their relationships with specific outcomes vary somewhat, the AHEI-2010, AMED, DASH and hPDI all showed an inverse association with major chronic disease. This aligns with numerous studies supporting the protective associations of AHEI-2010, AMED and DASH with CVD, type 2 diabetes and cancer[16–18]. As we accounted for smoking intensity and quit time, the inverse relationships of AHEI-2010, AMED and DASH with smoking-related cancer may not be entirely explained by residual confounding by smoking. Although dietary evidence for smoking-related cancer is sparse, a meta-analysis found that the inverse association between the Mediterranean diet and lung cancer was greater in former smokers[19], suggesting potential effect modification by smoking status. Previous research suggested that individuals who have increased their adherence to healthy plant-based diets had decreased risks of type 2 diabetes, overall CVD and overall mortality[20–22]. We found that, in agreement with findings from a meta-analysis on plant-based diets[20], the inverse relationship between hPDI and CVD was more pronounced for coronary heart disease than for stroke. One possibility is that the hPDI may have distinct relationships with ischemic and hemorrhagic stroke.

Despite moderate correlations with other dietary patterns, the WCRF/AICR dietary score includes the fewest items. That it was not specifically customized to prevent diseases other than cancer may account for the null association with major chronic disease. However, we found an unexpected positive association for type 2 diabetes and generally null associations for cancer-related outcomes. As the 2018 WCRF/AICR recommendations included behavioral components beyond nutritional guidance[23], this score based solely on dietary recommendations may not represent the optimal diet for cancer prevention. Prior studies showed that the inverse associations between the WCRF/AICR lifestyle score and cancer incidence and all-cause mortality, as well as the improvement of the biomarker profile, were predominantly driven by physical activity and body weight components, as opposed to the dietary components[24–26].

Due in part to the heterogeneity of cancer, the associations between dietary patterns and cancer-related outcomes were not as strong as those for other outcomes. Total cancer is a diverse constellation of unique diseases with a variety of causes, such as cigarette smoking, obesity, physical inactivity and infections[27]. Even within one type of cancer, there are multiple subtypes. Some cancers, such as colorectal cancer, are diet related whereas others are not[27]. A 10% reduction from dietary patterns alone (ignoring the potential effect of diet on weight control) would be considered relatively large based on current understanding. Although more directly related to cardiometabolic diseases, the rEDIH, rEDIP and DRRD displayed that a portion of cancers could be prevented by diet. In addition, following these three diets may be equally or even more beneficial for cancer prevention compared with other healthy diets, including the one that is specially designed for cancer.

The individual food components and scoring methods of dietary patterns may partially explain the differential associations between dietary patterns and chronic disease. The WCRF/AICR score considers alcohol drinking as harmful due to strong evidence supporting its carcinogenic effect[27], which might obscure the potential benefits of moderate drinking for reducing coronary heart disease, ischemic stroke, type 2 diabetes and all-cause mortality[28–30]. The AHEI-2010 and AMED discourage low or heavy drinking while supporting moderate drinking. On the other hand, alcoholic beverages are given positive weights in the rEDIP and rEDIH because drinking alcohol is associated with lower levels of insulinemia and inflammation. Given the complex effects of alcohol on health, recommendations on alcohol intake should be personalized based on each person's risk profile. Nonetheless, the associations of these four scores with risk of subsequent chronic diseases changed only modestly when the alcohol component was eliminated, suggesting that health improvement may still be achieved without alcohol intake. Coffee drinking, which is considered in the DRRD, rEDIH and rEDIP, has been associated with a lower risk of cardiometabolic disease[31]. The lower level of coffee intake among participants who scored high for WCRF/AICR might also contribute to the slightly positive association between WCRF/AICR and type 2 diabetes. Dietary guidelines emphasizing foods contributing to caloric intake would not account for dietary factors such as coffee consumption.

Our results also highlight that dietary patterns could reflect the overall effects of diet beyond the sum of individual foods. From a reductionist perspective, specific dietary components, such as coffee or moderate alcohol intake, may contribute to the benefits of maintaining a healthy diet. The relatively weak correlations of rEDIH and rEDIP with some frequently recommended foods, such as cruciferous vegetables and legumes, suggest that the metabolic effects of diet may be cumulative and not dominated by a few components. The unexpected protective association of dessert intake with chronic diseases may be explained by its inverse correlations with foods that appeared unfavorable for metabolic diseases (such as poultry, tomatoes and eggs). The flavonoids in chocolate may also contribute to the inverse associations between dessert and chronic diseases[32,33]. Given the varying degrees of health benefits associated with dietary patterns, future research on the subtle difference between patterns may be important for maximizing the effectiveness of dietary interventions.

After adjusting for BMI, the relationships with chronic disease for rEDIH and rEDIP were attenuated most because these two scores had the strongest association with BMI, possibly because of the residual confounding (BMI) or mediation effect (long-term weight gain)[34]. Our results remained consistent across several sensitivity tests. The stronger associations found in the subgroup analysis imply that dietary modification may be more advantageous for women, younger people and people with a higher BMI. The associations with rEDIH, rEDIP and DRRD were particularly strong for those with a high BMI. This result would be expected because the influence of diet related to insulin resistance and inflammation is likely to be greater in individuals at risk for insulin resistance resulting from a high BMI. Yet, even in those with normal BMI, these scores still had the strongest inverse association with chronic disease. The strong associations between dietary patterns and major chronic disease in both ever- and never-smokers underscore the potential of dietary modification.

The goal of dietary guidelines for adults is to provide advice on food and beverage choices to meet nutritional needs and to help prevent diet-related chronic diseases. The pathophysiological processes underlying major chronic diseases include, among others, lipids, blood pressure, glycemia, insulinemia and inflammation. From an overall health perspective, the most important dietary components will probably be those that substantially affect the pathophysiological processes that affect sizable numbers of multiple chronic diseases. Our findings for rEDIH and rEDIP may reflect that hyperinsulinemia and inflammation are strongly influenced by diet, have overlapping factors and are quantitatively important for multiple diseases. The considerably stronger associations for rEDIH and rEDIP than other recommended dietary patterns may indicate that, when recommendations are formulated, greater emphasis may need to be given to specific dietary components that influence inflammation and insulinemia, especially in the context of rising trends in obesity.

The strengths of the present study include large sample size, long follow-up period, repeated assessments of dietary intake using validated instruments, detailed collection of lifestyle and medical data allowing for adjustment for potential confounders, and comprehensive comparisons of multiple dietary patterns with major chronic diseases within the same analytical framework.

Study limitations should be acknowledged. The composite outcome of major chronic disease did not include cognitive outcomes and other functional impairments. We included incident cases for diseases that are leading causes of death in the United States of America and well ascertained in the cohorts. To capture other chronic conditions severe enough to increase mortality, nontraumatic death was included in the alternative definition of major chronic disease. However, the possibility for reverse causation could increase after including mortality endpoints. We attempted to address this by conducting latency analyses and observed consistently strong associations for both individual and composite outcomes. With longer latency between diet and outcome assessment, the analysis may be conservative in not accounting for recent diet, but largely exclude the effect of reverse causation. Our study also implies that diet has more immediate effects on cardiometabolic diseases, but a longer latency period is required for cancer. The four dietary patterns—AMED, DASH, hPDI and DRRD—were determined based on the distribution of the research population using either the median or the quintiles as the cut-off, which may not represent populations with more diverse racial and ethnic compositions. Yet, we found comparable associations across different ethnic groups, supporting the generalizability of the diet and disease associations.

In conclusion, we found that maintaining a healthy diet was generally associated with a lower risk of developing major chronic diseases. Among the eight dietary patterns examined in the present study, those reflecting low insulinemic, low inflammatory and diabetes risk-reducing diets may confer the largest risk reduction for various chronic diseases. The rEDIH and rEDIP were developed empirically based on associations between specific foods and biomarkers of insulinemia and inflammation, which may contribute to their strong predictive capacity. Future research is needed to explore more specific biological mechanisms underlying the relationship between diet and overall health.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-023-02235-5.

References

1. Buttorff, C., Ruder, T. & Bauman, M. *Multiple Chronic Conditions in the United States* (RAND Corporation, 2017).
2. Murphy, S. L., Kochanek, K. D., Xu, J. & Arias, E. Mortality in the United States, 2020. NCHS Data Brief No. 427 (National Center for Health Statistics, 2021).
3. Rao, M., Afshin, A., Singh, G. & Mozaffarian, D. Do healthier foods and diet patterns cost more than less healthy options? A systematic review and meta-analysis. *BMJ Open* **3**, e004277 (2013).
4. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**, 1958–1972 (2019).
5. Dietary Guidelines Advisory Committee. *Scientific Report of the 2020 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture and the Secretary of Health and Human Services* (US Department of Agriculture, Agricultural Research Service, 2020).
6. *Dietary Guidelines for Americans, 2020–2025* (US Department of Agriculture and US Department of Health and Human Services, 2020).
7. Tabung, F. K. et al. Development and validation of an empirical dietary inflammatory index. *J. Nutr.* **146**, 1560–1570 (2016).
8. Tabung, F. K. et al. Development and validation of empirical indices to assess the insulinaemic potential of diet and lifestyle. *Br. J. Nutr.* **116**, 1787–1798 (2016).
9. Jin, Q. et al. Insulinemic and inflammatory dietary patterns show enhanced predictive potential for type 2 diabetes risk in postmenopausal women. *Diabetes Care* **44**, 707–714 (2021).
10. Lee, D. H. et al. Dietary inflammatory and insulinemic potential and risk of type 2 diabetes: results from three prospective U.S. cohort studies. *Diabetes Care* **43**, 2675–2683 (2020).
11. Wan, Y. et al. Dietary insulinemic potential and risk of total and cause-specific mortality in the Nurses' Health Study and the Health Professionals Follow-up Study. *Diabetes Care* **45**, 451–459 (2022).
12. Li, J. et al. Dietary inflammatory potential and risk of cardiovascular disease among men and women in the U.S. *J. Am. Coll. Cardiol.* **76**, 2181–2193 (2020).
13. Shi, N. et al. Proinflammatory and hyperinsulinemic dietary patterns are associated with specific profiles of biomarkers predictive of chronic inflammation, glucose-insulin dysregulation, and dyslipidemia in postmenopausal women. *Front Nutr.* **8**, 690428 (2021).
14. Rhee, J. J., Mattei, J., Hughes, M. D., Hu, F. B. & Willett, W. C. Dietary diabetes risk reduction score, race and ethnicity, and risk of type 2 diabetes in women. *Diabetes Care* **38**, 596–603 (2015).
15. Wang, C. R. et al. Type 2 diabetes-prevention diet and all-cause and cause-specific mortality: a prospective study. *Am. J. Epidemiol.* **191**, 472–486 (2022).
16. Morze, J., Danielewicz, A., Hoffmann, G. & Schwingshackl, L. Diet quality as assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and Health Outcomes: a second update of a systematic review and meta-analysis of cohort studies. *J. Acad. Nutr. Diet.* **120**, 1998–2031.e1915 (2020).
17. Sofi, F., Cesari, F., Abbate, R., Gensini, G. F. & Casini, A. Adherence to Mediterranean diet and health status: meta-analysis. *Br. Med. J.* **337**, a1344 (2008).
18. Bloomfield, H. E. et al. Effects on health outcomes of a Mediterranean diet qith no restriction on fat intake: a systematic review and meta-analysis. *Ann. Intern. Med.* **165**, 491–500 (2016).
19. Bahrami, A. et al. Adherence to the Mediterranean diet and the risk of lung cancer: a systematic review and dose-response meta-analysis of observational studies. *Nutr. Rev.* **80**, 1118–1128 (2022).
20. Gan, Z. H., Cheong, H. C., Tu, Y. K. & Kuo, P. H. Association between plant-based dietary patterns and risk of cardiovascular disease: a systematic review and meta-analysis of prospective cohort studies. *Nutrients* **13**, 3952 (2021).
21. Baden, M. Y. et al. Changes in plant-based diet quality and total and cause-specific mortality. *Circulation* **140**, 979–991 (2019).
22. Chen, Z. et al. Changes in plant-based diet indices and subsequent risk of type 2 diabetes in women and men: three U.S. prospective cohorts. *Diabetes Care* **44**, 663–671 (2021).
23. Shams-White, M. M. et al. Operationalizing the 2018 World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) Cancer Prevention Recommendations: a standardized scoring system. *Nutrients* **11**, 1572 (2019).
24. Kaluza, J., Harris, H. R., Hakansson, N. & Wolk, A. Adherence to the WCRF/AICR 2018 recommendations for cancer prevention and risk of cancer: prospective cohort studies of men and women. *Br. J. Cancer* **122**, 1562–1570 (2020).
25. Shams-White, M. M., Brockton, N. T., Mitrou, P., Kahle, L. L. & Reedy, J. The 2018 World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) score and all-cause, cancer, and cardiovascular disease mortality risk: a longitudinal analysis in the NIH-AARP diet and health study. *Curr. Dev. Nutr.* **6**, nzac096 (2022).

26. Tabung, F. K. et al. Associations between adherence to the World Cancer Research Fund/American Institute for Cancer Research cancer prevention recommendations and biomarkers of inflammation, hormonal, and insulin response. *Int J. Cancer* **140**, 764–776 (2017).
27. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. A review of human carcinogens. Part E: personal habits and indoor combustions. *IARC Monogr. Eval. Carcinog. Risks Hum.* **100**, 373–472 (2012).
28. Lichtenstein, A. H. et al. 2021 dietary guidance to improve cardiovascular health: a scientific statement from the American Heart Association. *Circulation* **144**, e472–e487 (2021).
29. Knott, C., Bell, S. & Britton, A. Alcohol consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of more than 1.9 million individuals from 38 observational studies. *Diabetes Care* **38**, 1804–1812 (2015).
30. Xi, B. et al. Relationship of alcohol consumption to all-cause, cardiovascular, and cancer-related mortality in U.S. adults. *J. Am. Coll. Cardiol.* **70**, 913–922 (2017).
31. Grosso, G., Godos, J., Galvano, F. & Giovannucci, E. L. Coffee, caffeine, and health outcomes: an umbrella review. *Annu Rev. Nutr.* **37**, 131–156 (2017).
32. Zhao, B. et al. Relationship between chocolate consumption and overall and cause-specific mortality, systematic review and updated meta-analysis. *Eur. J. Epidemiol.* **37**, 321–333 (2022).
33. Hooper, L. et al. Effects of chocolate, cocoa, and flavan-3-ols on cardiovascular health: a systematic review and meta-analysis of randomized trials. *Am. J. Clin. Nutr.* **95**, 740–751 (2012).
34. Tabung, F. K., Satija, A., Fung, T. T., Clinton, S. K. & Giovannucci, E. L. Long-term change in both dietary insulinemic and inflammatory potential is associated with weight gain in adult women and men. *J. Nutr.* **149**, 804–815 (2019).

Methods

Study population

The present study leveraged data collected in three prospective cohorts: the NHS, the NHS II and the HPFS. The NHS was established in 1976 and enrolled 121,700 female nurses aged 30–55 years at baseline. As a younger cohort, the NHS II enrolled 116,429 female nurses aged 25–42 in 1989. The HPFS enrolled 51,529 male health professionals aged 40–75 in 1986. Questionnaires were sent to participants in these cohorts to collect and update their lifestyle and medical history every two years. Participants provided data on their dietary intake during the preceding year using validated semiquantitative food frequency questionnaires (FFQs) every four years. The follow-up rates were around 90% in all three cohorts. In the present study, we used 1984 for the NHS, 1991 for the NHS II and 1986 for the HPFS as the baseline when detailed dietary data were first assessed using an expanded FFQ with >100 items.

We excluded participants who were missing data on dietary pattern scores, those with implausible energy intake, those with a baseline history of CVD, diabetes or cancer (except nonmelanoma skin cancer and nonfatal prostate cancer), and those with extreme BMI (<15 or >50 kg m^{-2}). To reduce potential reverse causation, participants were censored when they were aged 80 years. The institutional review boards (IRBs) of the Brigham and Women's Hospital, Harvard T.H. Chan School of Public Health and participating registries have approved the study protocol.

Dietary assessment

Participants were asked to specify their food consumption frequency of specified portion sizes in the FFQ. The nutrient intake was computed as the sum of the nutrient content of each contributing food multiplied by its consumption frequency. Several studies have evaluated the validity and reliability of self-reported food and nutrient measures. Comparing the estimates from FFQs with those from multiple 1-week diet records, the average correlation coefficient for food was 0.66 in the NHS and 0.63 in the HPFS, and for nutrients was 0.53 in the NHS and 0.66 in the HPFS[35–39]. The correlation coefficient for dietary patterns ranges from 0.50 to 0.80 in the NHS and the HPFS[40].

Details of each dietary pattern can be found in Fig. 1. As a measure of healthy US-style eating, the AHEI-2010 assigns 0–10 points to each of the 11 dietary components based on the portion size[41]. The AMED depicts a Mediterranean-style diet and assigns 0 or 1 point to each of the 9 components based on whether the intakes are higher than the population median[42]. Representing a vegetarian diet, the hPDI includes 18 food groups and each group receives 1–5 points based on its consumption quintile[43]. The DASH score contains 8 components, each of which receives 1–5 points according to its consumption quintile[44]. The DRRD score similarly assigns 1–5 points to each of the 9 components associated with type 2 diabetes[45]. Based on 5 dietary recommendations for cancer prevention, the WCRF/AICR dietary score assigns 0–1 point to each depending on adherence level[23]. To reflect the long-term dietary hyperinsulinemia potential, the EDIH was derived to predict fasting plasma C-peptide[8]. The EDIP was developed to simultaneously predict plasma interleukin-6, C-reactive protein and tumor necrosis factor α receptor 2 (ref. [7]).

Covariate assessment

We extracted information from biennial questionnaires for race, family history of diabetes, family history of cancer, family history of CVD, physical activity, BMI, height, cigarette smoking (status, pack-years and time since quitting), multivitamin use, regular aspirin use, regular NSAID use and postmenopausal hormone use for women. Data on census tract-level variables were obtained by linking the US Census to participants' geocoded addresses. A summary score for nSES was calculated based on 9 census tract variables, including median family income, median home value, >25% with college or higher degree, percentage of families receiving interest dividends or rent income, percentage who occupied housing units, percentage white, percentage black, percentage foreign born and percentage aged >16 who were unemployed[46].

Outcome definition

The primary outcome was major chronic disease, defined as the first occurrence of incident major CVD, type 2 diabetes or total cancer (excluding nonmelanoma skin cancer and nonfatal prostate cancer). Secondary outcomes were the components of major chronic disease: major CVD (coronary heart disease and stroke), type 2 diabetes and total cancer (obesity-related cancer[47] and smoking-related cancer[27]). To test the influence of conditions not included in the primary outcome, we examined alternative major chronic disease, defined as the first occurrence of incident major CVD, type 2 diabetes, total cancer or nontraumatic death from all other causes. The diseases and corresponding *International Classification of Diseases*, 8th edn (ICD-8)[48] codes are provided in Supplementary Table 1.

Participants who reported a new diagnosis of CVD or cancer were asked for permission to obtain their medical records and pathological reports[49]. Deaths were identified through the next of kin or post office when questionnaires were mailed and through searches of the National Death Index. Death ascertainment using the National Death Index was reported to have a high sensitivity (98%) and specificity (100%)[50,51]. Permission was obtained from the next of kin or other contact person to review the medical records. Physicians who were blinded to the exposure information reviewed the medical records to confirm the diagnosis or determine the cause of death. Fatal coronary heart disease and fatal stroke were confirmed by death certificate and additional pathological evidence from either autopsy reports or medical records. Nonfatal myocardial infarction was confirmed according to the World Health Organization criteria and nonfatal stroke was confirmed according to the National Survey of Stroke criteria[52,53]. Type 2 diabetes was confirmed according to the National Diabetes Data Group criteria (before 1988) or the American Diabetes Association criteria (after 1988) using a supplementary questionnaire[54,55]. The confirmation rate based on medical records was estimated to range from 97% to 98% (refs. [56,57]).

Statistical analysis

Person-time of follow-up was accumulated from baseline until the occurrence of the outcome, death, age 80 or the end of follow-up (January 2016 for the HPFS, June 2016 for the NHS and June 2017 for NHS II), whichever came first. The length of follow-up differed because separate analyses were conducted for each outcome. The main analysis used cumulative averages of dietary pattern scores as the exposure to capture long-term intake. To limit the potential influence of outliers, pattern scores were winsorized at the 0.5 and 99.5 percentiles[58]. We used the residual method to compute energy-adjusted scores by fitting each pattern score against the total energy intake[59]. Nonmissing values from the preceding data cycle were used to fill in missing dietary variables and covariates. To facilitate comparison across pattern scores, we reversed the EDIH and EDIP scores so that the highest levels for both scores were regarded as the healthiest, similar to the other scores.

We assessed the relationship of energy-adjusted pattern scores to each other using Spearman's correlation coefficients. Time-dependent Cox's proportional hazards regression models with age as the time scale were fitted to estimate the associations of patterns with the risk of major chronic diseases and secondary outcomes. Each pattern score was modeled as a continuous variable standardized by its increment from the 10th to the 90th percentile. The potential nonlinear relationship between dietary pattern scores and outcome was examined by restricted cubic splines[60]. We also examined the associations for patterns in quintiles. The proportional hazards assumption was tested by adding an interaction term between each major dietary pattern and the time-scale age.

Analyses were performed in each cohort as well as the pooled data of three cohorts. All the analyses were stratified by age in months and calendar year of the questionnaire. In the pooled data, the model was additionally stratified by cohort. The estimates from pooled data were compared with the random effects meta-analyses based on three cohorts. MV models were adjusted for family history of diabetes, family history of cancer, family history of CVD, physical activity, cigarette smoking (status, pack-years and time since quitting), multivitamin use, regular aspirin use, regular NSAID use, postmenopausal hormone use for women and total energy intake. The model was also adjusted for alcohol intake for dietary patterns that did not include alcohol, such as DASH, hPDI and DRRD. We did not adjust for BMI in the main analysis, but did so in a sensitivity analysis because adiposity is a potential mediator for the diet and chronic disease relationship.

Furthermore, we explored which foods might explain the associations. The definition of food groups has been described previously[38]. The relationships between pattern scores and food groups were evaluated using Spearman's correlation coefficients. The associations of food groups with major chronic diseases were examined using Cox's proportional hazards regression models. Each food group was modeled as a continuous variable standardized by its increment from the 10th to the 90th percentile. The models were adjusted for the same set of covariates used in the main analysis.

We conducted subgroup analyses by age, BMI, gender, smoking status, alcohol, nSES and race/ethnicity. Potential interaction was assessed using Wald's test (binary variable) or the likelihood ratio test (categorical variable). To better understand possible latency, we investigated dietary pattern scores with different latency periods (0–4, 4–8, 8–12 or 12–16 years)[61]. For example, in a 4- to 8-year latency analysis, the pattern score constructed based on the 1990 FFQ was used as the exposure for the follow-up period between 1994 and 1998.

Finally, we conducted several sensitivity analyses. We adjusted for the nSES score to evaluate potential residual confounding. For the rEDIH, rEDIP, AHEI-2010, AMED and WCRF/AICR dietary score, we evaluated whether removal of alcohol from the pattern scores influenced the associations. Given that coffee consumption was inversely associated with cardiometabolic disease[31], we also examined the associations of DRRD, rEDIH and rEDIP with major chronic disease after removing coffee from the pattern scores. As diabetes was identified based on questionnaires, we examined the associations for diabetes with additional censoring at the last questionnaire response. Participants were similarly censored in a sensitivity analysis for major chronic disease. Participants may change their dietary habits after a diagnosis with an intermediate endpoint, such as hypertension, hypercholesterolemia, angina, transient ischemic attack or coronary artery bypass graft surgery. In a sensitivity analysis, we stopped updating the dietary information at the diagnosis of these intermediate endpoints. Using the same reference group, we compared individuals who were in the highest quintile (the healthiest level) of each dietary pattern with those who were never in the highest quintiles of any pattern.

Statistical analyses were performed using SAS 9.4. We reported unadjusted *P* values based on two-sided statistical tests. We did not adjust for multiple testing because we aimed to compare the patterns, some of which have previously been examined individually.

Ethics

This is an observational study. The study protocol was approved by the IRBs of the Brigham and Women's Hospital, Harvard T.H. Chan School of Public Health and participating registries (IRB protocol nos. 2001P001945/BWH and 10372). The IRBs allowed participants' completion of questionnaires to be considered as implied consent for participation in these studies of health professionals. Written informed consent was required for biomarker collection and medical record acquisition. The study was performed in accordance with the Declaration of Helsinki.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

As a result of participant confidentiality and privacy concerns, data are available upon written request. According to standard controlled access procedure, applications to use the NHS, NHS II and HPFS resources will be reviewed by our External Collaborators Committee for scientific aims, evaluation of the fit of the data for the proposed methodology, and verification that the proposed use meets the guidelines of the Ethics and Governance Framework and of the consent that was provided by the participants. Investigators wishing to use the NHS, NHS II and HPFS data are asked to submit a brief description of the proposed project. Further information including the procedures to obtain and access data from the NHS, NHS II and HPFS is described at https://www.nurseshealthstudy.org/researchers (contact email: nhsaccess@channing.harvard.edu) and https://sites.sph.harvard.edu/hpfs/for-collaborators.

Code availability

The analysis programs are publicly available through https://github.com/pwangepi/DPs-and-chronic-disease.

References

35. Salvini, S. et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int. J. Epidemiol.* **18**, 858–867 (1989).
36. Yuan, C. et al. Validity of a dietary questionnaire assessed by comparison with multiple weighed dietary records or 24-hour recalls. *Am. J. Epidemiol.* **185**, 570–584 (2017).
37. Feskanich, D. et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J. Am. Diet. Assoc.* **93**, 790–796 (1993).
38. Hu, F. B. et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am. J. Clin. Nutr.* **69**, 243–249 (1999).
39. Al-Shaar, L. et al. Reproducibility and validity of a semiquantitative food frequency questionnaire in men assessed by multiple methods. *Am. J. Epidemiol.* **190**, 1122–1132 (2021).
40. Yue, Y. et al. Reproducibility and validity of diet quality scores derived from food-frequency questionnaires. *Am. J. Clin. Nutr.* **115**, 843–853 (2022).
41. Chiuve, S. E. et al. Alternative dietary indices both strongly predict risk of chronic disease. *J. Nutr.* **142**, 1009–1018 (2012).
42. Fung, T. T. et al. Diet quality is associated with the risk of estrogen receptor-negative breast cancer in postmenopausal women. *J. Nutr.* **136**, 466–472 (2006).
43. Satija, A. et al. Plant-based dietary patterns and incidence of type 2 diabetes in US men and women: results from three prospective cohort studies. *PLoS Med.* **13**, e1002039 (2016).
44. Fung, T. T. et al. Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. *Arch. Intern. Med.* **168**, 713–720 (2008).
45. Wang, T. et al. Diabetes risk reduction diet and survival after breast cancer diagnosis. *Cancer Res.* **81**, 4155–4162 (2021).
46. Iyer, H. S. et al. Impact of neighborhood socioeconomic status, income segregation, and greenness on blood biomarkers of inflammation. *Environ. Int.* **162**, 107164 (2022).
47. Lauby-Secretan, B. et al. Body fatness and cancer—viewpoint of the IARC Working Group. *N. Engl. J. Med.* **375**, 794–798 (2016).
48. World Health Assembly, 19. *Report of the International Conference for the Eighth Revision of the International Classification of Diseases* (World Health Organization, 1966); https://apps.who.int/iris/handle/10665/137300

49. Colditz, G. A. et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am. J. Epidemiol.* **123**, 894–900 (1986).
50. Stampfer, M. J. et al. Test of the National Death Index. *Am. J. Epidemiol.* **119**, 837–839 (1984).
51. Rich-Edwards, J. W., Corsano, K. A. & Stampfer, M. J. Test of the National Death Index and Equifax Nationwide Death Search. *Am. J. Epidemiol.* **140**, 1016–1019 (1994).
52. Walker, A. E., Robins, M. & Weinfeld, F. D. The National Survey of Stroke. Clinical findings. *Stroke* **12**, I13–I44 (1981).
53. Mendis, S. et al. World Health Organization definition of myocardial infarction: 2008–09 revision. *Int J. Epidemiol.* **40**, 139–146 (2011).
54. American Diabetes Association. Standards of medical care in diabetes—2010. *Diabetes Care* **33**, S11–S61 (2010).
55. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* **28**, 1039–1057 (1979).
56. Manson, J. E. et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *Lancet* **338**, 774–778 (1991).
57. Hu, F. B. et al. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch. Intern. Med.* **161**, 1542–1548 (2001).
58. Mozaffarian, D., Hao, T., Rimm, E. B., Willett, W. C. & Hu, F. B. Changes in diet and lifestyle and long-term weight gain in women and men. *N. Engl. J. Med.* **364**, 2392–2404 (2011).
59. Willett, W. & Stampfer, M. J. Total energy intake: implications for epidemiologic analyses. *Am. J. Epidemiol.* **124**, 17–27 (1986).
60. Durrleman, S. & Simon, R. Flexible regression models with cubic splines. *Stat. Med.* **8**, 551–561 (1989).
61. Lee, J. E. et al. Folate intake and risk of colorectal cancer and adenoma: modification by time. *Am. J. Clin. Nutr.* **93**, 817–825 (2011).

Acknowledgements

We acknowledge the contribution to the present study from central cancer registries supported through the Centers for Disease Control and Prevention's National Program of Cancer Registries and/or the National Cancer Institute's Surveillance, Epidemiology, and End Results Program. Central registries may also be supported by state agencies, universities and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Delaware, Colorado, Connecticut, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia and Wyoming. This work was supported by grants from the National Institutes of Health (nos. UM1 CA186107 to A.H.E., P01 CA087969 to A.H.E. and E.L.G., R01 HL034594, R01 HL088521 and U01 CA176726 to W.C.W. and A.H.E., U01 HL145386 and U01 CA167552 to W.C.W., R01 HL035464 to E.B.R., U01 CA261961 to M.S. and K99 CA207736 to F.K.T.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

P.W., F.K.T. and E.L.G. conceived and designed the study. M.S., A.H.E., E.B.R., W.C.W., F.K.T. and E.L.G. acquired the data and obtained funding. P.W. conducted statistical analysis and wrote the first draft of the paper. M.S. provided technical review. All authors interpreted the results and revised the paper. F.K.T. and E.L.G. supervised the study. All authors approved the final paper as submitted.

Competing interests

The authors declare no competing interests.

Additional information

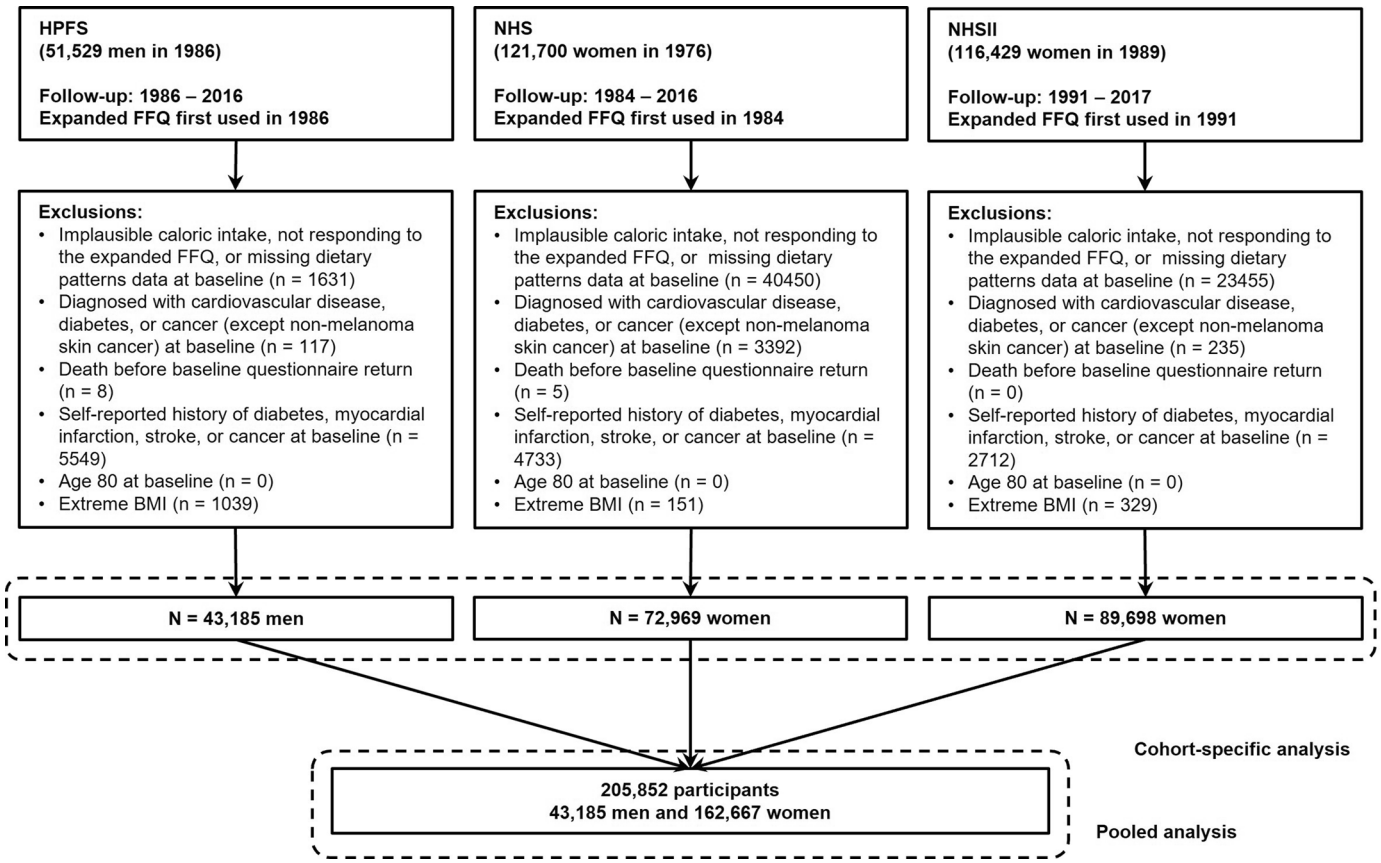
Extended data is available for this paper at https://doi.org/10.1038/s41591-023-02235-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-023-02235-5.

Correspondence and requests for materials should be addressed to Peilu Wang.

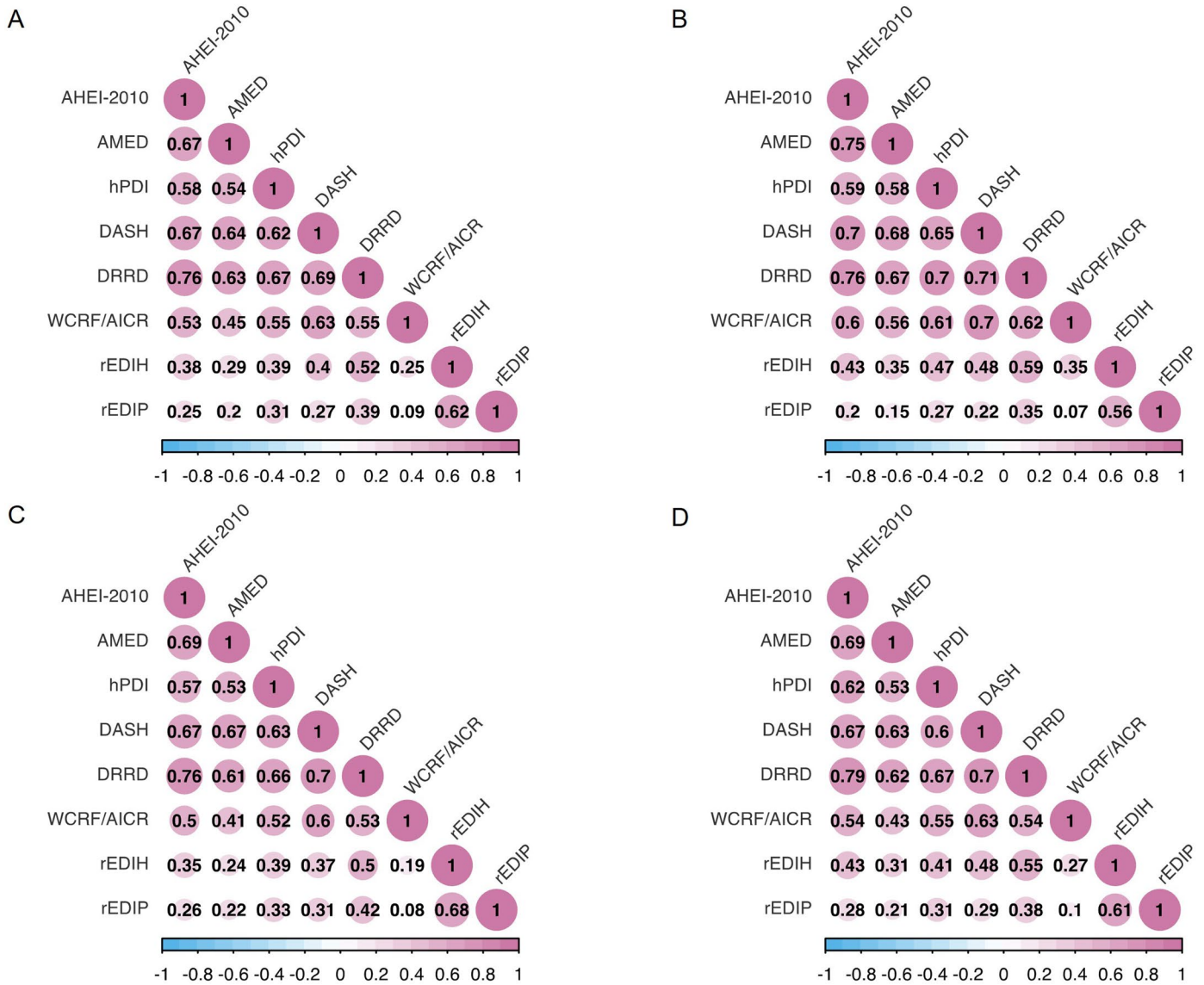
Peer review information *Nature Medicine* thanks Yan-Bo Zhang, Keren Papier and Bernard Srour for their contributions to the peer review of this paper. Primary Handling Editor: Ming Yang, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-023-02235-5



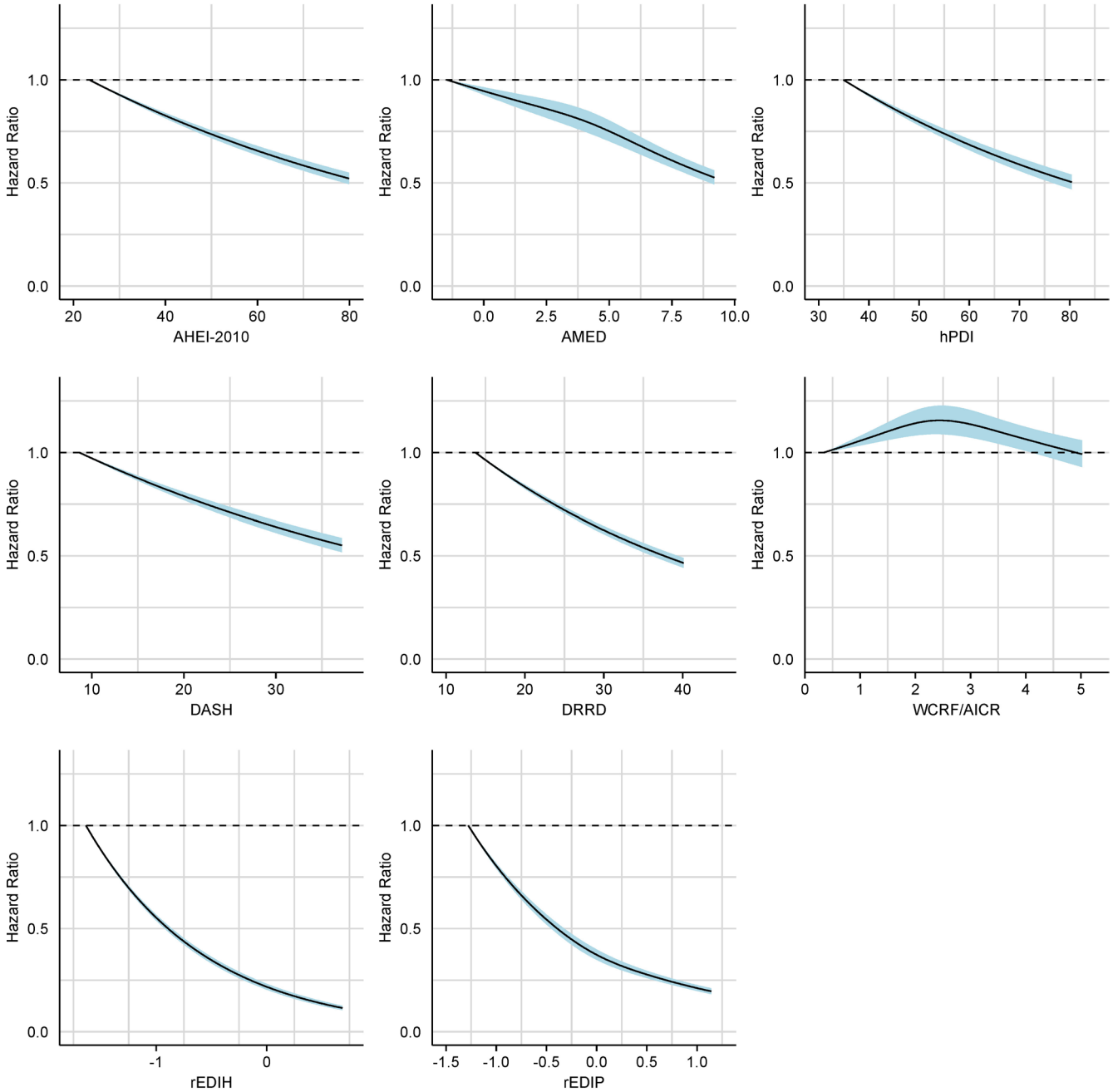
Extended Data Fig. 1 | Flowchart of participants included in the main analysis. BMI, body mass index; FFQ, food frequency questionnaires; Nurses' Health Study (NHS); Health Professionals Follow-up Study (HPFS).

https://doi.org/10.1038/s41591-023-02235-5



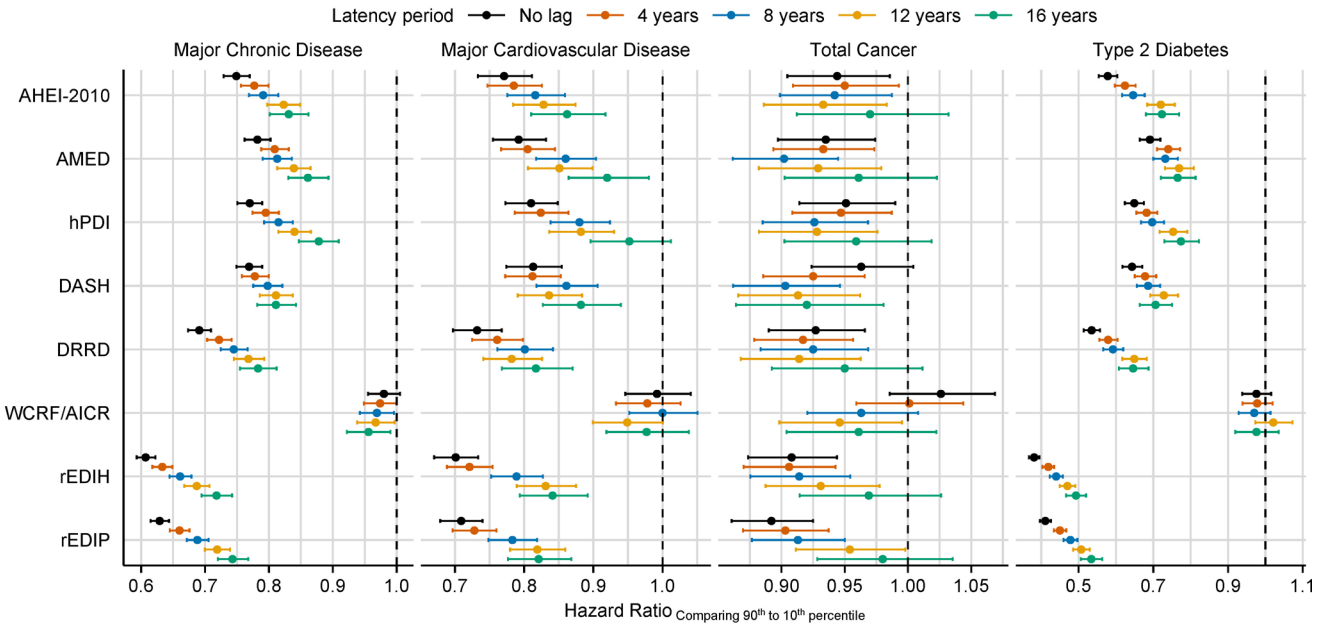
Extended Data Fig. 2 | Baseline Spearman correlations between energy-adjusted cumulative average dietary patterns in (a) all cohorts, (b) the Health Professionals Follow-up Study, (c) the Nurses' Health Study, and (d) Nurses' Health Study II. *P* values based on the two-sided tests were <0.0001 for all correlations (not adjusted for multiple comparisons). AHEI-2010, Alternative Healthy Eating Index-2010; AMED, Alternate Mediterranean Diet score; DASH, Dietary Approaches to Stop Hypertension score; DRRD, Diabetes Risk Reduction Diet; hPDI, Healthful plant-based diet index; rEDIH, reversed Empirical dietary index for hyperinsulinemia; rEDIP, reversed Empirical dietary inflammation pattern; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) dietary score.

https://doi.org/10.1038/s41591-023-02235-5



Extended Data Fig. 3 | Multivariable-adjusted spline analysis of dietary pattern scores with risk of major chronic disease. *P* values for nonlinearity based on the two-sided tests were statistically significant for AMED ($P = 0.03$), WCRF/AICR ($P = 0.007$), and rEDIP ($P = 0.008$) (not adjusted for multiple comparisons). The hazard ratios (black line) and the 95% confidence intervals (grey bands) are shown. The models were adjusted for the same list of covariates as in Table 2. AHEI-2010, Alternative Healthy Eating Index-2010; AMED, Alternate Mediterranean Diet score; DASH, Dietary Approaches to Stop Hypertension score; DRRD, Diabetes Risk Reduction Diet; hPDI, Healthful plant-based diet index; HR, Hazard ratio; rEDIH, reversed Empirical dietary index for hyperinsulinemia; rEDIP, reversed Empirical dietary inflammation pattern; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) dietary score.

https://doi.org/10.1038/s41591-023-02235-5



Extended Data Fig. 4 | Multivariable-adjusted associations between cumulative average dietary patterns (comparing the 90th to 10th percentile) and major chronic disease and major components in the pooled data of three cohorts (n = 205,852 participants) with different lags. Analyses details and corresponding estimates are provided in Extended Data Table 3. The hazard ratios are indicated by the circles and the 95% confidence intervals are reflected by the error bars. AHEI-2010, Alternative Healthy Eating Index-2010; AMED, Alternate Mediterranean Diet score; DASH, Dietary Approaches to Stop Hypertension score; DRRD, Diabetes Risk Reduction Diet; hPDI, Healthful plant-based diet index; rEDIH, reversed Empirical dietary index for hyperinsulinemia; rEDIP, reversed Empirical dietary inflammation pattern; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) dietary score.

https://doi.org/10.1038/s41591-023-02235-5

Extended Data Table 1 | Associations between cumulative average dietary patterns (comparing the 90th with the 10th percentile) and secondary outcomes in the pooled data

Outcome		Coronary Heart Disease	Stroke	Obesity-related Cancer	Smoking-related Cancer	Alternative Major Chronic Disease
Cases		8,083	5,398	8,660	9,118	53,524
Person-year		5,185,291	5,203,823	5,189,782	5,204,833	4,852,894
Pattern	Model	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
AHEI-2010	Age	0.57 (0.54, 0.61)	0.72 (0.67, 0.77)	0.83 (0.78, 0.88)	0.68 (0.64, 0.72)	0.60 (0.59, 0.62)
	MV[a]	0.72 (0.68, 0.77)	0.84 (0.78, 0.91)	0.92 (0.86, 0.97)	0.87 (0.82, 0.93)	0.74 (0.72, 0.76)
	MV + nSES[b]	0.75 (0.70, 0.80)	0.86 (0.79, 0.93)	0.92 (0.87, 0.98)	0.87 (0.82, 0.93)	0.76 (0.74, 0.78)
	MV + BMI[c]	0.75 (0.71, 0.80)	0.86 (0.79, 0.93)	0.94 (0.89, 1.00)	0.88 (0.83, 0.94)	0.79 (0.77, 0.81)
AMED	Age	0.60 (0.57, 0.64)	0.71 (0.66, 0.76)	0.81 (0.77, 0.86)	0.64 (0.61, 0.68)	0.62 (0.61, 0.63)
	MV[a]	0.78 (0.73, 0.83)	0.84 (0.77, 0.91)	0.91 (0.86, 0.97)	0.86 (0.81, 0.92)	0.78 (0.76, 0.80)
	MV + nSES[b]	0.80 (0.75, 0.85)	0.85 (0.79, 0.92)	0.92 (0.86, 0.97)	0.86 (0.81, 0.92)	0.79 (0.77, 0.81)
	MV + BMI[c]	0.81 (0.76, 0.87)	0.86 (0.79, 0.93)	0.94 (0.89, 1.00)	0.87 (0.82, 0.92)	0.84 (0.82, 0.86)
hPDI	Age	0.67 (0.63, 0.71)	0.81 (0.76, 0.87)	0.87 (0.83, 0.92)	0.79 (0.75, 0.84)	0.68 (0.67, 0.70)
	MV[a]	0.80 (0.75, 0.85)	0.91 (0.84, 0.98)	0.94 (0.89, 0.99)	0.92 (0.87, 0.98)	0.79 (0.77, 0.81)
	MV + nSES[b]	0.80 (0.76, 0.86)	0.92 (0.85, 0.99)	0.94 (0.89, 0.99)	0.92 (0.87, 0.98)	0.80 (0.78, 0.82)
	MV + BMI[c]	0.82 (0.77, 0.87)	0.92 (0.86, 1.00)	0.96 (0.90, 1.01)	0.93 (0.88, 0.98)	0.83 (0.81, 0.85)
DASH	Age	0.59 (0.55, 0.62)	0.69 (0.64, 0.74)	0.81 (0.76, 0.86)	0.60 (0.57, 0.64)	0.60 (0.59, 0.62)
	MV[a]	0.79 (0.74, 0.84)	0.84 (0.78, 0.91)	0.94 (0.88, 1.00)	0.87 (0.82, 0.92)	0.78 (0.77, 0.80)
	MV + nSES[b]	0.81 (0.76, 0.87)	0.86 (0.79, 0.93)	0.94 (0.89, 1.00)	0.87 (0.82, 0.92)	0.80 (0.78, 0.82)
	MV + BMI[c]	0.81 (0.76, 0.87)	0.86 (0.79, 0.93)	0.96 (0.90, 1.02)	0.87 (0.82, 0.93)	0.83 (0.81, 0.85)
DRRD	Age	0.58 (0.55, 0.62)	0.67 (0.62, 0.72)	0.84 (0.79, 0.89)	0.73 (0.69, 0.77)	0.58 (0.57, 0.60)
	MV[a]	0.71 (0.67, 0.76)	0.76 (0.70, 0.82)	0.91 (0.86, 0.97)	0.88 (0.83, 0.94)	0.70 (0.68, 0.71)
	MV + nSES[b]	0.73 (0.69, 0.78)	0.77 (0.71, 0.83)	0.91 (0.86, 0.97)	0.88 (0.83, 0.94)	0.71 (0.69, 0.72)
	MV + BMI[c]	0.74 (0.69, 0.79)	0.77 (0.71, 0.83)	0.94 (0.88, 0.99)	0.89 (0.84, 0.94)	0.74 (0.72, 0.76)
WCRF/AICR	Age	0.77 (0.73, 0.82)	0.87 (0.81, 0.93)	0.89 (0.85, 0.94)	0.71 (0.67, 0.75)	0.81 (0.79, 0.83)
	MV[a]	1.00 (0.94, 1.06)	1.03 (0.95, 1.11)	1.00 (0.95, 1.06)	0.95 (0.90, 1.01)	1.00 (0.98, 1.03)
	MV + nSES[b]	1.01 (0.95, 1.07)	1.03 (0.96, 1.11)	1.00 (0.95, 1.06)	0.95 (0.90, 1.01)	1.01 (0.99, 1.04)
	MV + BMI[c]	0.98 (0.92, 1.04)	1.01 (0.94, 1.09)	0.98 (0.93, 1.04)	0.95 (0.89, 1.00)	0.97 (0.95, 1.00)
rEDIH	Age	0.56 (0.53, 0.59)	0.71 (0.66, 0.76)	0.81 (0.76, 0.85)	0.80 (0.76, 0.85)	0.53 (0.51, 0.54)
	MV[a]	0.63 (0.60, 0.67)	0.76 (0.71, 0.82)	0.84 (0.80, 0.89)	0.89 (0.84, 0.94)	0.59 (0.58, 0.60)
	MV + nSES[b]	0.65 (0.61, 0.69)	0.77 (0.72, 0.83)	0.84 (0.80, 0.89)	0.89 (0.84, 0.94)	0.60 (0.59, 0.61)
	MV + BMI[c]	0.71 (0.67, 0.76)	0.82 (0.77, 0.89)	0.93 (0.88, 0.99)	0.92 (0.87, 0.97)	0.73 (0.71, 0.75)
rEDIP	Age	0.63 (0.59, 0.66)	0.75 (0.70, 0.81)	0.83 (0.79, 0.88)	0.95 (0.90, 1.00)	0.59 (0.58, 0.60)
	MV[a]	0.65 (0.61, 0.68)	0.77 (0.71, 0.82)	0.83 (0.79, 0.87)	0.91 (0.87, 0.96)	0.62 (0.60, 0.63)
	MV + nSES[b]	0.66 (0.62, 0.69)	0.77 (0.72, 0.83)	0.83 (0.79, 0.88)	0.91 (0.87, 0.96)	0.62 (0.61, 0.64)
	MV + BMI[c]	0.71 (0.67, 0.75)	0.82 (0.76, 0.88)	0.90 (0.86, 0.96)	0.94 (0.89, 0.99)	0.73 (0.72, 0.75)

MV, multivariable; nSES, neighborhood socioeconomic status. [a]Analyses were stratified by age (in month), calendar year and cohort. MV model: adjusted for physical activity (<3.0, 3.0–8.9, 9.0–17.9, 18.0–26.9, 27.0–41.9 or ≥42 MET-h per week), cigarette smoking status (never, former quitting ≥10 years, former quitting <10 years, current), cigarette smoking pack-years (0, 1–4, 5–14, 15–24 or ≥25 pack-years), multivitamin use (yes or no), regular aspirin use (yes or no), regular NSAID use (yes or no), postmenopausal hormone use (premenopausal, never, former or current use) for women and total energy intake (quintiles). Alcohol consumption (<5.0, 5.0–14.9 or ≥15.0 g d^{-1}) was adjusted for DASH, hPDI and DRRD. For alternative major chronic disease, the model was additionally adjusted for family history of diabetes (yes or no), family history of cancer (yes or no) and family history of CVD (yes or no). For coronary heart disease and stroke, the model was additionally adjusted for family history of CVD (yes or no). For obesity-related or smoking-related cancer, the model was additionally adjusted for family history of cancer (yes or no) and height (continuous). [b]Additionally adjusted for nSES (continuous). [c]Additionally adjusted for BMI (continuous).

https://doi.org/10.1038/s41591-023-02235-5

Extended Data Table 2 | Multivariable-adjusted associations of the cumulative average dietary patterns (comparing the 90th with the 10th percentile) with major chronic disease in subgroups

	Age[a], year		BMI[b], kg/m^2		Sex[c]	
	<65	≥65	<25	≥25	Men	Women
Case	24,246	20,729	16,250	28,725	14427	30,548
Person-year	3,719,096	1,133,799	2,785,883	2,067,011	879,987	3,972,907
AHEI-2010	0.71 (0.68, 0.74)	0.83 (0.79, 0.86)	0.85 (0.82, 0.89)	0.76 (0.73, 0.79)	0.84 (0.81, 0.88)	0.72 (0.70, 0.75)
AMED	0.76 (0.73, 0.78)	0.84 (0.81, 0.87)	0.87 (0.83, 0.91)	0.80 (0.77, 0.83)	0.83 (0.79, 0.87)	0.78 (0.75, 0.80)
hPDI	0.75 (0.73, 0.78)	0.86 (0.83, 0.89)	0.87 (0.84, 0.91)	0.79 (0.76, 0.81)	0.86 (0.82, 0.90)	0.77 (0.75, 0.80)
DASH	0.73 (0.71, 0.76)	0.85 (0.81, 0.88)	0.85 (0.81, 0.89)	0.78 (0.75, 0.81)	0.83 (0.79, 0.87)	0.76 (0.74, 0.79)
DRRD	0.66 (0.63, 0.68)	0.78 (0.75, 0.81)	0.80 (0.76, 0.84)	0.69 (0.66, 0.71)	0.82 (0.78, 0.85)	0.66 (0.64, 0.68)
WCRF/AICR	1.03 (0.99, 1.07)	0.97 (0.93, 1.01)	0.96 (0.93, 1.01)	0.99 (0.95, 1.02)	0.95 (0.91, 0.99)	1.03 (1.00, 1.06)
rEDIH	0.51 (0.50, 0.53)	0.71 (0.68, 0.73)	0.78 (0.75, 0.81)	0.63 (0.61, 0.65)	0.74 (0.70, 0.77)	0.53 (0.51, 0.54)
rEDIP	0.55 (0.53, 0.56)	0.72 (0.70, 0.75)	0.76 (0.73, 0.79)	0.64 (0.63, 0.66)	0.77 (0.74, 0.80)	0.55 (0.53, 0.56)
	Smoking Status[d]		Alcohol[e], g/day		nSES[f]	
	Never	Former	Current	≤14	>14	< Median
Case	19,941	18,719	6,315	37,456	7,519	24,693
Person-year	2,689,238	1,671,370	492,287	4,228,443	624,451	2,424,086
AHEI-2010	0.74 (0.71, 0.77)	0.76 (0.73, 0.79)	0.79 (0.73, 0.85)	0.75 (0.73, 0.78)	0.80 (0.75, 0.85)	0.79 (0.77, 0.82)
AMED	0.77 (0.74, 0.80)	0.79 (0.76, 0.82)	0.82 (0.77, 0.89)	0.79 (0.76, 0.81)	0.82 (0.76, 0.87)	0.82 (0.79, 0.85)
hPDI	0.78 (0.75, 0.81)	0.79 (0.75, 0.82)	0.87 (0.81, 0.94)	0.79 (0.77, 0.82)	0.83 (0.78, 0.89)	0.81 (0.78, 0.84)
DASH	0.73 (0.70, 0.76)	0.78 (0.75, 0.82)	0.87 (0.81, 0.94)	0.77 (0.75, 0.80)	0.81 (0.76, 0.87)	0.80 (0.77, 0.83)
DRRD	0.68 (0.65, 0.70)	0.70 (0.67, 0.73)	0.78 (0.72, 0.84)	0.70 (0.68, 0.73)	0.75 (0.70, 0.80)	0.71 (0.69, 0.74)
WCRF/AICR	0.99 (0.95, 1.03)	1.00 (0.96, 1.04)	1.04 (0.96, 1.12)	0.99 (0.96, 1.02)	0.89 (0.82, 0.96)	1.04 (1.00, 1.07)
rEDIH	0.51 (0.50, 0.53)	0.60 (0.57, 0.62)	0.72 (0.67, 0.77)	0.55 (0.53, 0.56)	0.77 (0.73, 0.81)	0.59 (0.57, 0.61)
rEDIP	0.55 (0.53, 0.57)	0.62 (0.60, 0.64)	0.75 (0.71, 0.80)	0.57 (0.55, 0.58)	0.80 (0.76, 0.84)	0.60 (0.58, 0.62)
	nSES[g]				Race/Ethnicity[g]	
	≥ Median	Non-Hispanic White	Asian	Hispanic	Black	Unknown
Case	20,282	39,433	554	354	686	3,948
Person-year	2,428,808	4,501,212	62,284	47,875	58,149	183,374
AHEI-2010	0.76 (0.73, 0.79)	0.76 (0.74, 0.78)	0.91 (0.65, 1.28)	0.51 (0.34, 0.78)	0.72 (0.53, 0.97)	0.85 (0.78, 0.94)
AMED	0.78 (0.75, 0.81)	0.79 (0.77, 0.81)	0.78 (0.55, 1.10)	0.55 (0.37, 0.82)	0.62 (0.46, 0.86)	0.88 (0.80, 0.96)
hPDI	0.80 (0.77, 0.83)	0.79 (0.77, 0.82)	0.87 (0.63, 1.20)	0.44 (0.30, 0.66)	0.81 (0.61, 1.07)	0.96 (0.88, 1.05)
DASH	0.78 (0.74, 0.81)	0.77 (0.75, 0.80)	0.82 (0.59, 1.14)	0.46 (0.31, 0.69)	0.61 (0.45, 0.82)	0.92 (0.84, 1.00)
DRRD	0.72 (0.69, 0.74)	0.70 (0.68, 0.72)	0.80 (0.56, 1.12)	0.47 (0.32, 0.71)	0.63 (0.47, 0.85)	0.84 (0.76, 0.92)
WCRF/AICR	0.98 (0.94, 1.01)	1.01 (0.98, 1.04)	1.07 (0.78, 1.46)	0.78 (0.53, 1.14)	0.93 (0.70, 1.24)	0.97 (0.89, 1.06)
rEDIH	0.60 (0.58, 0.62)	0.57 (0.55, 0.58)	0.55 (0.39, 0.77)	0.47 (0.32, 0.69)	0.65 (0.49, 0.87)	0.87 (0.80, 0.94)
rEDIP	0.64 (0.62, 0.66)	0.60 (0.58, 0.61)	0.61 (0.44, 0.84)	0.52 (0.37, 0.75)	0.59 (0.44, 0.79)	0.86 (0.79, 0.92)

Analyses were stratified by age (in month), calendar year and cohort. Models were adjusted for the same list of covariates as in Table 2. We reported unadjusted *P* values based on two-sided statistical tests. [a]*P* values for Wald's test of interaction term (pattern×age) were <0.0001 for all pattern scores except WCRF/AICR (*P*=0.21). [b]*P* values for Wald's test of interaction term (pattern×BMI) were <0.0001 for all pattern scores except AMED (*P*=0.80), hPDI (*P*=0.006), DASH (*P*=0.46) and WCRF/AICR (*P*=0.54). [c]*P* values for Wald's test of interaction term (pattern×gender) were <0.0001 for all pattern scores except WCRF/AICR (*P*=0.30). [d]*P* values for the likelihood ratio test of interaction term (pattern×smoking status) were <0.0001 for all pattern scores. [e]*P* values for Wald's test of interaction term (pattern×alcohol) were <0.05 for DRRD (*P*=0.007), WCRF/AICR (*P*=0.04), rEDIH (*P*<0.0001) and rEDIP (*P*<0.0001). [f]*P* values for Wald's test of interaction term (pattern×nSES) were <0.05 for AMED (*P*=0.0009), DASH (*P*=0.02), DRRD (*P*=0.04), WCRF/AICR (*P*=0.004) and rEDIP (*P*=0.02). [g]*P* values for the likelihood ratio test of interaction terms (pattern×race/ethnicity) were <0.0001 for all pattern scores except WCRF/AICR (*P*=0.03).

https://doi.org/10.1038/s41591-023-02235-5

Extended Data Table 3 | Association between cumulative average dietary patterns (comparing the 90th with the 10th percentile) and major chronic disease and major components in the pooled data with different latency periods

Latency period		No lag	4 years	8 years	12 years	16 years
Outcome	Pattern	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Major Chronic Disease	Cases	44,975	42,698	38,080	31,232	22,585
	Person-year	4,852,894	4,135,546	3,399,860	2,652,663	1,925,289
	AHEI-2010	0.75 (0.73, 0.77)	0.78 (0.76, 0.80)	0.79 (0.77, 0.81)	0.82 (0.80, 0.85)	0.83 (0.80, 0.86)
	AMED	0.78 (0.76, 0.80)	0.81 (0.79, 0.83)	0.81 (0.79, 0.84)	0.84 (0.81, 0.87)	0.86 (0.83, 0.89)
	hPDI	0.77 (0.75, 0.79)	0.79 (0.77, 0.82)	0.81 (0.79, 0.84)	0.84 (0.81, 0.87)	0.88 (0.85, 0.91)
	DASH	0.77 (0.75, 0.79)	0.78 (0.76, 0.80)	0.80 (0.78, 0.82)	0.81 (0.79, 0.84)	0.81 (0.78, 0.84)
	DRRD	0.69 (0.67, 0.71)	0.72 (0.70, 0.74)	0.75 (0.72, 0.77)	0.77 (0.74, 0.79)	0.78 (0.75, 0.81)
	WCRF/AICR	0.98 (0.96, 1.01)	0.97 (0.95, 1.00)	0.97 (0.94, 1.00)	0.97 (0.94, 1.00)	0.96 (0.92, 0.99)
	rEDIH	0.61 (0.59, 0.62)	0.63 (0.62, 0.65)	0.66 (0.64, 0.68)	0.69 (0.67, 0.71)	0.72 (0.69, 0.74)
	rEDIP	0.63 (0.61, 0.64)	0.66 (0.64, 0.68)	0.69 (0.67, 0.71)	0.72 (0.70, 0.74)	0.74 (0.72, 0.77)
Major Cardiovascular Disease	Cases	12,962	12,844	11,960	10,114	7,694
	Person-year	5,148,378	4,402,319	3,616,470	2,805,738	2,012,374
	AHEI-2010	0.77 (0.73, 0.81)	0.79 (0.75, 0.83)	0.82 (0.78, 0.86)	0.83 (0.78, 0.87)	0.86 (0.81, 0.92)
	AMED	0.79 (0.75, 0.83)	0.80 (0.77, 0.84)	0.86 (0.82, 0.90)	0.85 (0.81, 0.90)	0.92 (0.86, 0.98)
	hPDI	0.81 (0.77, 0.85)	0.82 (0.79, 0.86)	0.88 (0.84, 0.92)	0.88 (0.84, 0.93)	0.95 (0.90, 1.01)
	DASH	0.81 (0.77, 0.85)	0.81 (0.77, 0.85)	0.86 (0.82, 0.91)	0.84 (0.79, 0.88)	0.88 (0.83, 0.94)
	DRRD	0.73 (0.70, 0.77)	0.76 (0.72, 0.80)	0.80 (0.76, 0.84)	0.78 (0.74, 0.83)	0.82 (0.77, 0.87)
	WCRF/AICR	0.99 (0.95, 1.04)	0.98 (0.93, 1.03)	1.00 (0.95, 1.05)	0.95 (0.90, 1.00)	0.98 (0.92, 1.04)
	rEDIH	0.70 (0.67, 0.73)	0.72 (0.69, 0.75)	0.79 (0.75, 0.83)	0.83 (0.79, 0.87)	0.84 (0.79, 0.89)
	rEDIP	0.71 (0.68, 0.74)	0.73 (0.70, 0.76)	0.78 (0.75, 0.82)	0.82 (0.78, 0.86)	0.82 (0.78, 0.87)
Type 2 Diabetes	Cases	18,615	16,588	15,056	11,951	8,077
	Person-year	4,630,725	3,726,657	3,124,205	2,365,983	1,646,645
	AHEI-2010	0.58 (0.55, 0.60)	0.62 (0.60, 0.65)	0.65 (0.62, 0.68)	0.72 (0.68, 0.76)	0.72 (0.68, 0.77)
	AMED	0.69 (0.66, 0.72)	0.74 (0.71, 0.77)	0.73 (0.70, 0.77)	0.77 (0.73, 0.81)	0.77 (0.72, 0.81)
	hPDI	0.65 (0.62, 0.67)	0.68 (0.65, 0.71)	0.70 (0.67, 0.73)	0.75 (0.72, 0.79)	0.77 (0.73, 0.82)
	DASH	0.64 (0.62, 0.67)	0.68 (0.65, 0.71)	0.69 (0.66, 0.72)	0.73 (0.69, 0.77)	0.71 (0.66, 0.75)
	DRRD	0.54 (0.51, 0.56)	0.58 (0.56, 0.60)	0.59 (0.57, 0.62)	0.65 (0.62, 0.68)	0.65 (0.61, 0.69)
	WCRF/AICR	0.98 (0.94, 1.02)	0.98 (0.94, 1.02)	0.97 (0.93, 1.01)	1.02 (0.97, 1.07)	0.98 (0.92, 1.04)
	rEDIH	0.38 (0.37, 0.39)	0.42 (0.40, 0.43)	0.44 (0.42, 0.46)	0.47 (0.45, 0.49)	0.49 (0.47, 0.52)
	rEDIP	0.41 (0.40, 0.43)	0.45 (0.43, 0.47)	0.48 (0.46, 0.50)	0.51 (0.49, 0.53)	0.53 (0.51, 0.56)
Total Cancer	Cases	17,909	16,595	14,176	11,127	7,682
	Person-year	5,138,951	4,397,783	3,616,916	2,808,237	2,015,999
	AHEI-2010	0.94 (0.90, 0.99)	0.95 (0.91, 0.99)	0.94 (0.90, 0.99)	0.93 (0.89, 0.98)	0.97 (0.91, 1.03)
	AMED	0.93 (0.90, 0.97)	0.93 (0.89, 0.97)	0.90 (0.86, 0.94)	0.93 (0.88, 0.98)	0.96 (0.90, 1.02)
	hPDI	0.95 (0.91, 0.99)	0.95 (0.91, 0.99)	0.93 (0.89, 0.97)	0.93 (0.88, 0.98)	0.96 (0.90, 1.02)
	DASH	0.96 (0.92, 1.00)	0.92 (0.89, 0.97)	0.90 (0.86, 0.95)	0.91 (0.87, 0.96)	0.92 (0.86, 0.98)
	DRRD	0.93 (0.89, 0.97)	0.92 (0.88, 0.96)	0.93 (0.88, 0.97)	0.91 (0.87, 0.96)	0.95 (0.89, 1.01)
	WCRF/AICR	1.03 (0.99, 1.07)	1.00 (0.96, 1.04)	0.96 (0.92, 1.01)	0.95 (0.90, 1.00)	0.96 (0.90, 1.02)
	rEDIH	0.91 (0.87, 0.94)	0.91 (0.87, 0.94)	0.91 (0.88, 0.95)	0.93 (0.89, 0.98)	0.97 (0.91, 1.03)
	rEDIP	0.89 (0.86, 0.92)	0.90 (0.87, 0.94)	0.91 (0.88, 0.95)	0.95 (0.91, 1.00)	0.98 (0.93, 1.04)

Analyses were stratified by age (in month), calendar year and cohort. MV model: adjusted for physical activity (<3.0, 3.0–8.9, 9.0–17.9, 18.0–26.9, 27.0–41.9 or ≥42 MET-h per week), cigarette smoking status (never, former quitting ≥10 years, former quitting <10 years, current), cigarette smoking pack-years (0, 1–4, 5–14, 15–24 or ≥25 pack-years), multivitamin use (yes or no), regular aspirin use (yes or no), regular NSAID use (yes or no), postmenopausal hormone use (premenopausal, never, former or current use) for women and total energy intake (quintiles). For DASH, hPDI and DRRD, alcohol consumption (<5.0, 5.0–14.9 or ≥15.0 g d^{-1}) was additionally adjusted for. For major chronic disease, the model was additionally adjusted for family history of diabetes (yes or no), family history of cancer (yes or no) and family history of CVD (yes or no). For type 2 diabetes, the model was additionally adjusted for family history of diabetes (yes or no). For major CVD, the model was additionally adjusted for family history of CVD (yes or no). For total cancer, the model was additionally adjusted for family history of cancer (yes or no) and height (continuous).

https://doi.org/10.1038/s41591-023-02235-5

Extended Data Table 4 | Association between cumulative average dietary patterns without alcohol component (comparing the 90th with the 10th percentile) and outcomes in the pooled data

Outcome		Major Chronic Disease	Major Cardiovascular Disease	Type 2 Diabetes	Total Cancer
Cases		44,975	12,962	18,615	17,909
Person-year		4,852,894	5,148,378	4,630,725	5,138,951
AHEI-2010	Age	0.69 (0.67, 0.71)	0.66 (0.63, 0.69)	0.57 (0.55, 0.59)	0.82 (0.79, 0.86)
	MV	0.83 (0.81, 0.85)	0.82 (0.78, 0.86)	0.74 (0.71, 0.77)	0.97 (0.93, 1.01)
AMED	Age	0.68 (0.67, 0.70)	0.66 (0.63, 0.69)	0.62 (0.60, 0.65)	0.78 (0.75, 0.81)
	MV	0.84 (0.82, 0.86)	0.83 (0.79, 0.87)	0.81 (0.77, 0.84)	0.94 (0.90, 0.98)
WCRF/AICR	Age	0.72 (0.70, 0.73)	0.72 (0.68, 0.75)	0.57 (0.55, 0.59)	0.89 (0.85, 0.92)
	MV	0.86 (0.84, 0.89)	0.88 (0.84, 0.92)	0.77 (0.73, 0.80)	1.00 (0.96, 1.05)
rEDIH	Age	0.55 (0.54, 0.56)	0.64 (0.61, 0.67)	0.33 (0.32, 0.35)	0.83 (0.80, 0.86)
	MV	0.62 (0.60, 0.63)	0.72 (0.68, 0.75)	0.41 (0.39, 0.42)	0.89 (0.85, 0.92)
rEDIP	Age	0.61 (0.60, 0.63)	0.69 (0.66, 0.72)	0.39 (0.38, 0.41)	0.87 (0.84, 0.90)
	MV	0.66 (0.64, 0.67)	0.73 (0.70, 0.76)	0.47 (0.45, 0.48)	0.87 (0.84, 0.90)

Analyses were stratified by age (in month), calendar year and cohort. MV model: adjusted for physical activity (<3.0, 3.0–8.9, 9.0–17.9, 18.0–26.9, 27.0–41.9 or ≥42 MET-h per week), cigarette smoking status (never, former quitting ≥10 years, former quitting <10 years, current), cigarette smoking pack-years (0, 1–4, 5–14, 15–24 or ≥25 pack-years), multivitamin use (yes or no), regular aspirin use (yes or no), regular NSAID use (yes or no), postmenopausal hormone use (premenopausal, never, former or current use) for women and total energy intake (quintiles). Alcohol consumption (<5.0, 5.0–14.9 or ≥15.0 g d^{-1}) was adjusted for DASH, hPDI and DRRD. For major chronic disease, the model was additionally adjusted for family history of diabetes (yes or no), family history of cancer (yes or no) and family history of CVD (yes or no). For type 2 diabetes, the model was additionally adjusted for family history of diabetes (yes or no). For major CVD, the model was additionally adjusted for family history of CVD (yes or no). For total cancer, the model was additionally adjusted for family history of cancer (yes or no) and height (continuous).

https://doi.org/10.1038/s41591-023-02235-5

Extended Data Table 5 | Association between cumulative average dietary patterns without coffee component (comparing the 90th with the 10th percentile) and outcomes in the pooled data

Outcome		Major Chronic Disease	Major Cardiovascular Disease	Type 2 Diabetes	Total Cancer
Cases		44,975	12,962	18,615	17,909
Person-year		4,852,894	5,148,378	4,630,725	5,138,951
DRRD	Age	0.62 (0.60, 0.63)	0.61 (0.58, 0.64)	0.49 (0.47, 0.51)	0.79 (0.76, 0.82)
	MV	0.74 (0.72, 0.76)	0.76 (0.72, 0.80)	0.62 (0.60, 0.65)	0.93 (0.89, 0.97)
rEDIH	Age	0.53 (0.52, 0.54)	0.60 (0.58, 0.63)	0.31 (0.30, 0.32)	0.81 (0.78, 0.84)
	MV	0.61 (0.59, 0.62)	0.71 (0.67, 0.74)	0.38 (0.36, 0.39)	0.91 (0.87, 0.94)
rEDIP	Age	0.58 (0.57, 0.59)	0.64 (0.62, 0.67)	0.36 (0.34, 0.37)	0.85 (0.82, 0.88)
	MV	0.66 (0.64, 0.67)	0.73 (0.70, 0.76)	0.42 (0.41, 0.44)	0.92 (0.88, 0.95)

Analyses were stratified by age (in month), calendar year and cohort. MV model: adjusted for physical activity (<3.0, 3.0–8.9, 9.0–17.9, 18.0–26.9, 27.0–41.9 or ≥42 MET-h per week), cigarette smoking status (never, former quitting ≥10 years, former quitting <10 years, current), cigarette smoking pack-years (0, 1–4, 5–14, 15–24 or ≥25 pack-years), multivitamin use (yes or no), regular aspirin use (yes or no), regular NSAID use (yes or no), postmenopausal hormone use (premenopausal, never, former or current use) for women and total energy intake (quintiles). Alcohol consumption (<5.0, 5.0–14.9 or ≥15.0 g d^{-1}) was adjusted for DASH, hPDI and DRRD. For major chronic disease, the model was additionally adjusted for family history of diabetes (yes or no), family history of cancer (yes or no) and family history of CVD (yes or no). For type 2 diabetes, the model was additionally adjusted for family history of diabetes (yes or no). For major CVD, the model was additionally adjusted for family history of CVD (yes or no). For total cancer, the model was additionally adjusted for family history of cancer (yes or no) and height (continuous).

https://doi.org/10.1038/s41591-023-02235-5

Extended Data Table 6 | Associations between cumulative average dietary patterns and risk of major chronic disease using the same reference group

Pattern	Model	Reference group	Highest Quintile
AHEI-2010	Cases	23,708	7,660
	Person-year	2,292,086	967,242
	Age	1	0.66 (0.64, 0.68)
	MV	1	0.75 (0.73, 0.77)
AMED	Cases	23,708	7,625
	Person-year	2,292,086	972,409
	Age	1	0.67 (0.65, 0.68)
	MV	1	0.77 (0.75, 0.79)
hPDI	Cases	23,708	8,063
	Person-year	2,292,086	972,078
	Age	1	0.69 (0.67, 0.71)
	MV	1	0.78 (0.76, 0.80)
DASH	Cases	23,708	7,857
	Person-year	2,292,086	972,141
	Age	1	0.66 (0.65, 0.68)
	MV	1	0.78 (0.76, 0.80)
DRRD	Cases	23,708	7,463
	Person-year	2,292,086	972,819
	Age	1	0.64 (0.63, 0.66)
	MV	1	0.74 (0.72, 0.76)
WCRF/AICR	Cases	23,708	8,841
	Person-year	2,292,086	970,161
	Age	1	0.75 (0.73, 0.77)
	MV	1	0.86 (0.84, 0.88)
rEDIH	Cases	23,708	7,164
	Person-year	2,292,086	969,987
	Age	1	0.62 (0.60, 0.63)
	MV	1	0.67 (0.65, 0.69)
rEDIP	Cases	23,708	6,690
	Person-year	2,292,086	971,184
	Age	1	0.64 (0.62, 0.66)
	MV	1	0.67 (0.65, 0.69)

Participants who were in the highest quintile of each dietary pattern were compared with those who were never in the highest quintile of any pattern. Analyses were stratified by age (in month), calendar year and cohort. MV model: adjusted for family history of diabetes (yes or no), family history of cancer (yes or no), family history of CVD (yes or no), physical activity (<3.0, 3.0–8.9, 9.0–17.9, 18.0–26.9, 27.0–41.9 or ≥42 MET-h per week), cigarette smoking status (never, former quitting ≥10 years, former quitting <10 years, current), cigarette smoking pack-years (0, 1–4, 5–14, 15–24 or ≥25 pack-years), multivitamin use (yes or no), regular aspirin use (yes or no), regular NSAID use (yes or no), postmenopausal hormone use (premenopausal, never, former or current use) for women and total energy intake (quintiles). For DASH, hPDI and DRRD, alcohol consumption (<5.0, 5.0–14.9 or ≥15.0 g d^{-1}) was additionally adjusted for.

nature portfolio

Corresponding author(s): Peilu Wang

Last updated by author(s): Jan 16, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Statistical analyses were conducted using SAS (version 9.4). The analysis programs are publicly available through https://github.com/pwangepi/DPs-and-chronic-disease.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Because of participant confidentiality and privacy concerns, data are available upon written request. According to standard controlled access procedure, applications to use the Nurses' Health Studies and Health Professionals Follow-up Study resources will be reviewed by our External Collaborators Committee for scientific aims, evaluation of the fit of the data for the proposed methodology, and verification that the proposed use meets the guidelines of the Ethics and

Governance Framework and the consent that was provided by the participants. Investigators wishing to use the Nurses' Health Studies and Health Professionals Follow-up Study data are asked to submit a brief description of the proposed project. Further information including the procedures to obtain and access data from the Nurses' Health Studies and Health Professionals Follow-up Study is described at https://www.nurseshealthstudy.org/researchers (contact email: nhsaccess@channing.harvard.edu) and https://sites.sph.harvard.edu/hpfs/for-collaborators/.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	We conducted subgroup analysis defined by sex (self-reported) and the findings apply to both men and women. The effect estimates and number of participants are described in the text and presented in Extended Data Table 2 and Fig.4.
Population characteristics	The study included 162,667 women (from the Nurses' Health Study and the Nurses' Health Study II) and 43,185 men (from the Health Professionals Follow-up Study) free of major chronic diseases with detailed dietary assessments. Details of the population characteristics can be found in the Results and Methods sections and Table 2 and Extended Data Fig.1. The Nurses' Health Study was established in 1976 and enrolled 121,700 female nurses ages 30 to 55 at baseline. As a younger cohort, the Nurses' Health Study II enrolled 116,429 female nurses ages 25 to 42 in 1989. The Health Professionals Follow-up Study enrolled 51,529 male health professionals ages 40 to 75 in 1986.
Recruitment	The study population was recruited from the three prospective cohorts: the Nurses' Health Study, the Nurses' Health Study II, and the Health Professionals Follow-up Study. Questionnaires were sent to participants in these cohorts to collect and update their lifestyle and medical history every two years. Participants provided data on their dietary intake during the preceding year using validated semi-quantitative food frequency questionnaires every four years. The follow-up rates were around 90% in all three cohorts.
Ethics oversight	This is an observational study. The study protocol was approved by the institutional review boards (IRBs) of the Brigham and Women's Hospital, Harvard T.H. Chan School of Public Health, and participating registries (IRB Protocol number: 2001P001945/BWH and 10372). The IRBs allowed participants' completion of questionnaires to be considered as implied consent for participation in these studies of health professionals. Written informed consent was required for biomarker collection and for medical record acquisition. The study was performed in accordance with the Declaration of Helsinki.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The study population is preexisting. We documented 44,975 major chronic disease events (12,962 major CVD, 18,615 diabetes, and 17,909 total cancer), allowing us to detect the major findings of the present study with a statistical power of >90% (a post hoc analysis).
Data exclusions	We excluded participants who were missing data on dietary pattern scores, those with implausible energy intake, those with a baseline history of CVD, diabetes, or cancer (except non-melanoma skin cancer), and those with extreme body mass index (BMI) (<15 or > 50 kg/ sq m). To reduce potential reverse causation, participants were censored when they were 80 years old.
Replication	The validity/reproducibility of questionnaires and outcome ascertainment have been reported previously and described in the Methods. Moderate correlations were observed for foods, nutrients, and dietary patterns comparing the estimates from food frequency questionnaires with those from multiple one-week diet records. Participants who reported a new diagnosis of major chronic disease were asked for permission to obtain their medical records and pathological reports. Deaths were identified through the next-of-kin or postal office when questionnaires were mailed and through searches of the National Death Index. Death ascertainment using National Death Index was reported to have a high sensitivity (98%) and specificity (100%).
Randomization	Not applicable. This study is an observational study. There are no experimental groups or interventions in the study.
Blinding	Not applicable because this is an observational study without any interventions. Physicians who reviewed the medical records were unaware of the exposure information.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging