



# Cumulative average dietary pattern scores in young adulthood and risk of incident type 2 diabetes: the CARDIA study

Kristin M. Hirahatake<sup>1</sup> · David R. Jacobs Jr<sup>2</sup> · James M. Shikany<sup>3</sup> · Luohua Jiang<sup>1</sup> · Nathan D. Wong<sup>1</sup> · Andrew O. Odegaard<sup>1</sup>

Received: 5 February 2019 / Accepted: 25 July 2019 / Published online: 2 September 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

**Aims/hypothesis** The evidence for the role of contemporary dietary patterns, trends and predominant aspects of energy intake in a typical American diet and in type 2 diabetes risk is limited. Therefore, we examined the association between dietary pattern scores created to reflect the 2015 Dietary Guidelines for Americans (DGA) Scientific Report, a Palaeolithic (Palaeo) diet, a diet high in ‘empty calories’, and the A Priori Diet Quality Score (APDQS) (cohort reference) and type 2 diabetes risk over time.

**Methods** We carried out a prospective analysis of 4719 young adult black and white men and women from the Coronary Artery Risk Development in Young Adults (CARDIA) study with repeated dietary histories collected at study years 0, 7 and 20. Using multivariable Cox proportional hazards regression models, we examined the association between time-dependent cumulative average dietary pattern scores and incident type 2 diabetes.

**Results** During the 30 year follow-up period, 680 (14.4%) incident cases of type 2 diabetes occurred. There was no association between the 2015 DGA, Palaeo or empty calorie scores and type 2 diabetes risk in the overall population. Participants in the fourth quartile of the APDQS, reflecting a more healthful dietary pattern, had a 45% lower risk of type 2 diabetes compared with those in the lowest quartile (HR 0.55 [95% CI 0.41, 0.74]). In stratified analyses there was an inverse association for the 2015 DGA in non-smokers per SD (HR 0.86 [95% CI 0.74, 0.99]) and an inverse association for the empty calorie score in white women (HR 0.76 [95% CI 0.60, 0.96]) as well as in a subgroup analysis of the Palaeo index of participants who maintained a high score over 20 years (per SD, HR 0.59 [95% CI 0.39, 0.88]).

**Conclusions/interpretation** Higher levels of the APDQS, which largely aligns with the 2015 DGA, were strongly inversely associated with 30 year type 2 diabetes risk in the CARDIA cohort; the results from the other patterns were nuanced and need to be considered in the context of the study and potential biases.

**Keywords** Dietary guidelines · Dietary patterns · Epidemiology · Palaeolithic · Type 2 diabetes

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00125-019-04989-5>) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

✉ Andrew O. Odegaard  
aodegaar@uci.edu

<sup>1</sup> Department of Epidemiology, School of Medicine, University of California, Irvine, CA 92697-7550, USA

<sup>2</sup> Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA

<sup>3</sup> Division of Preventive Medicine, School of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA

## Abbreviations

APDQS	A Priori Diet Quality Score
CARDIA	Coronary Artery Risk Development in Young Adults
DGA	Dietary Guidelines for Americans
EC	Empty calorie
EU	Exercise units
NCC	Nutrition Coordinating Center
Palaeo	Palaeolithic

## Introduction

The prevalence of type 2 diabetes has reached epidemic levels, with major economic and public health consequences [1]. If current trends continue, the prevalence of type 2 diabetes in

## Research in context

### What is already known about this subject?

- Dietary patterns incorporating higher intakes of fruits, vegetables, whole grains, fish, fermented dairy products, legumes, nuts and seeds and lower intakes of red meat, refined grains, sugar-sweetened beverages and processed meats are associated with lower type 2 diabetes risk

### What is the key question?

- What are the associations of dietary patterns reflecting the most recent Dietary Guidelines for Americans (DGA) and contemporary diet trends in the USA with incident type 2 diabetes risk in young adults?

### What are the new findings?

- A dietary pattern aligned with the 2015 DGA Scientific Report was inversely associated with type 2 diabetes risk over 30 years in non-smoking young adults
- The A Priori Diet Quality Score, which closely aligns with the 2015 DGA, was strongly inversely associated with type 2 diabetes risk in this cohort
- There was no association between a Palaeolithic dietary pattern and type 2 diabetes except in a subgroup of participants who maintained a high Palaeolithic score over 20 years (inverse association); there were also no clear associations between 'empty calorie' score and diabetes

### How might this impact on clinical practice in the foreseeable future?

- These results provide evidence on the association of contemporary dietary patterns and guidelines with type 2 diabetes risk and provides a reference for clinical practice in this area

the USA is projected to increase from approximately one in eight to one in three adults by 2050 [2]. Dietary intake is a major, modifiable risk factor for type 2 diabetes, and the current body of evidence suggests that diets incorporating higher intakes of fruits, vegetables, whole grains, legumes, nuts and seeds and lower intakes of red meat and processed meat, refined grains and sugar-sweetened beverages are associated with lower type 2 diabetes risk [3, 4].

The Scientific Report of the 2015 Dietary Guidelines Advisory Committee, the basis for the 2015 Dietary Guidelines for Americans (DGA) and related policy, reflects this evidence base but there is minimal evidence of a direct test of how a dietary pattern reflecting these guidelines links to type 2 diabetes risk. The report also recommends limiting consumption of foods high in 'empty calories' (ECs) (i.e. food or beverage that contribute few or no nutrients) primarily composed of solid fats and added sugars and also refined starches and alcohol. Such foods have been identified as leading contributors to excess energy intake in the USA [5]. Additionally, there is limited evidence on the relationship between other dietary patterns increasingly adopted by Americans (such as a Palaeolithic [Palaeo] diet) and type 2 diabetes risk.

To address this knowledge gap, we examined the association between dietary pattern scores constructed to reflect the 2015 DGA Scientific Report, a modern-day Palaeo diet and a high EC intake, and risk of type 2 diabetes in a cohort of young black and white men and women with repeated

assessments of diet over 30 years. We also examined the relationship between the A Priori Diet Quality Score (APDQS, cohort reference) and type 2 diabetes risk, as this score largely aligns with the 2015 DGA and has been shown to inversely associate with cardiovascular risk [6–9].

## Methods

**Study population** The Coronary Artery Risk Development in Young Adults (CARDIA) study is a multicentre, longitudinal investigation of the evolution of cardiovascular disease risk starting in young adulthood. Briefly, 5115 black and white men and women, aged 18–30 years, were recruited in 1985–1986 from four cities in the USA: Birmingham, Alabama; Chicago, Illinois; Minneapolis, Minnesota and Oakland, California. Details of participant enrolment and examination have been published elsewhere [10]. Re-examination occurred 2, 5, 7, 10, 15, 20, 25 and 30 years after baseline. The CARDIA study was approved by the institutional review board at each field centre and informed consent was obtained from all participants prior to enrolment [10].

Participants with a diagnosis of diabetes at baseline ( $n = 34$ ), missing baseline diabetes status ( $n = 88$ ) or baseline dietary data ( $n = 4$ ) were excluded from this analysis. Individuals without follow-up data were also excluded ( $n = 152$ ). Individuals who reported extreme energy intakes ( $<2510$  kJ/

day or >25,104 kJ/day for women [ $n = 53$ ] and <3347 kJ/day or >33,472 kJ/day for men [ $n = 64$ ]) were excluded. The proportion of missing data for other pertinent covariates was low (<1%). Missing values were imputed by multiple imputation, which had no impact on the results. The final study sample for this analysis comprised 4719 young adults.

**Dietary intake assessment** Dietary intake was assessed at study years 0, 7 and 20 by participant self-report to the interviewer-administered validated CARDIA Diet History, as previously described [11, 12]. Briefly, interviewers asked participants open-ended questions about dietary consumption during the past month within 100 food categories, referencing 1609 separate food items in years 0 and 7 and several thousand food items in year 20. Follow-up questions addressed serving size, frequency of consumption and common additions to foods. Diet history data used codes of the University of Minnesota Nutrition Coordinating Center (NCC) and foods were placed into 166 food groups using the food grouping system developed by the NCC. Food group intake was calculated as the total number of servings per day [11].

The creation of dietary pattern scores was modelled after the APDQS used in previous CARDIA studies [6, 13]. From the 166 food groups created by the NCC system, 44 condensed groups were formed based on similar nutrient characteristics and comparability with food groups previously defined [6, 13–16]. The study population was ranked into quintiles of intake (or, when the non-consumer group was large, a group of non-consumers and quartiles among consumers) for each of the 44 food groups. Dietary pattern scores were created by classifying each food group as beneficial (+), adverse (–) or neutral (not scored due to a lack of strong, conflicting or neutral evidence) (0) in terms of the recommendations of the specific diet (see electronic supplementary materials [ESM] Table 1). For the 2015 DGA and Palaeo scores, a moderate (+/–) classification for foods with recommended moderate consumption was incorporated. The scores for all food groups were relative to the distribution of consumption in the study population (not absolute quantities), as described below.

Food groups considered beneficial by the specified dietary pattern were scored 0–4 (the highest quintile of intake received a score of 4 and the lowest a score of 0) and those considered adverse were reverse scored (highest quintile scored 0, lowest 4). Food groups with recommended moderate consumption were scored 0, 2, 4, 2, 0, with the middle quintile receiving the highest score. Dietary pattern scores were calculated by summing an individual's scores for each food group. Neutral foods were not included in the final score. Higher scores reflect higher dietary alignment with the pre-specified patterns. The APDQS was calculated as described in previous CARDIA studies (ESM Table 2) [6, 8].

Previous investigators created the APDQS historically used in CARDIA dietary studies by classifying food groups in

terms of hypothesised health effects and relationship with disease to characterise diet quality [6, 15]. The 2015 Scientific Report concluded that diets associated with better health are characterised as follows: higher levels of vegetables, fruits, whole grains, low- or non-fat dairy products, seafood, legumes and nuts; moderate intake of alcohol (among adults); lower intake of red and processed meat and low intake of sugar-sweetened foods and drinks, and refined grains [5]. Foods and food groups not explicitly mentioned in the Advisory Committee's Consensus statement were not included in the score. The modern-day Palaeo dietary pattern was modelled after indices used in previous studies [17, 18] and was based on foods assumed to have been available to humans prior to the establishment of agriculture (mainly wild-animal and uncultivated-plant sources of foods), including meats, fish, vegetables, roots, eggs and nuts and excluding grains, legumes, dairy, salt, refined sugar and processed oils [19]. To test the concept of the Scientific Report definition of EC foods, we created an EC score based on 13 major food groups predominantly comprised of ECs [5], including alcohol, butter, margarine, chocolate, dairy dessert, fried foods, fried potatoes, fruit juice, grain dessert, refined grains, salty snacks, sugar-sweetened beverages and a sweet extra category. Intake of these food groups was measured as servings per 4184 kJ (1000 kcal). For comparison purposes, a score was also calculated using absolute servings of the 13 food groups. Higher EC scores reflect higher EC intake.

To examine the dietary patterns with type 2 diabetes risk, we calculated cumulative average scores for each pattern for each participant. For participants with follow-up time  $\leq 7$  years, baseline scores were used. For participants with follow-up time  $> 7$  and  $\leq 20$  years, the average of scores from year 0 and year 7 was used. For participants with repeated measures of diet and follow-up time  $> 20$  years, the average of dietary pattern scores at year 0, 7 and 20 was used. Cumulative averages were calculated based on available data; individuals without repeated measures of diet were assigned their baseline dietary pattern score.

**Type 2 diabetes status** Diabetes status was assessed clinically at examination years 0, 7, 10, 15, 20, 25 and 30. Incidence of type 2 diabetes was defined as use of diabetes medication, a fasting blood glucose level of  $\geq 7$  mmol/l (126 mg/dl), 2 h post-challenge glucose  $\geq 11.1$  mmol/l (200 mg/dl) and/or  $\text{HbA}_{1c} \geq 48$  mmol/mol (6.5%). The 2 h glucose was done at years 10, 20 and 25, while  $\text{HbA}_{1c}$  was done at years 20, 25 and 30. Details on blood collection and laboratory procedures can be found in [10]. In the CARDIA study there was no differentiation between type 1 and type 2 diabetes; however, it is likely that most incident cases identified during follow-up are type 2 diabetes given the age of the cohort. To avoid potential misclassification, a sensitivity analysis was performed excluding participants who developed diabetes before the age of 30 years.

**Covariates** At all CARDIA examinations, participants completed questionnaires on behaviours and sociodemographic, psychosocial and medical background [10]. Physical activity was assessed using the CARDIA physical activity questionnaire, a validated interviewer-based self-report of duration and intensity of participation in 13 categories of exercise over the past year [20]. Physical activity was reported in exercise units (EU), where 300 EU is approximately equal to 150 min of moderate-intensity physical activity per week or 30 min of moderate-intensity activity 5 days/week [21]. Body weight was measured with light clothing to the nearest 0.2 kg and height was measured without shoes to the nearest 0.5 cm. BMI was calculated as  $\text{kg/m}^2$  [10, 22]. Total energy intake was calculated from the CARDIA Diet History. Smoking status was assessed at all years using an interviewer-administered questionnaire. Participants were classified as current, former or never smokers. Participants who reported regular cigarette smoking (at least five cigarettes per week almost every week for at least 3 months) at the time of an examination were classified as current smokers. Former smokers were those who reported previously using cigarettes but not currently smoking. Never smokers reported no history of cigarette smoking. Self-report of cigarette smoking in CARDIA was validated at baseline against a biochemical marker of nicotine uptake (serum cotinine) and misclassification was found to be low [23].

**Statistical analysis** Participants' demographic, lifestyle and clinical characteristics were described by quartile of cumulative average dietary pattern score, using means with SD for continuous variables and frequencies with percentages for categorical variables. To compare characteristics between quartiles, ANOVA and  $\chi^2$  tests were performed for continuous variables and categorical variables, respectively. Survival analysis using multivariable Cox proportional hazards models was used to estimate the HRs and corresponding 95% CIs for incident diabetes. Separate models were fit for the 2015 DGA, Palaeo, EC and APDQS scores. Follow-up time was calculated as the time (years) from baseline to the first CARDIA study examination where diabetes was identified. Otherwise, participants were censored at the last diabetes status examination before death, loss to follow-up or end of cohort surveillance, whichever came first.

For the analyses, participants were ranked into quartiles of cumulative average dietary pattern score; the lowest quartile for each score served as the reference group. Multivariable models adjusted for preselected potential demographic and lifestyle confounders were used. The base model adjusted for age, race, sex and CARDIA study field centre. Model 2 included model 1 covariates plus sociodemographic and lifestyle confounders (smoking, education, energy intake and physical activity). We used the most recent smoking status and education level prior to diabetes diagnosis and the

cumulative average of years 0, 7 and 20 energy intake and physical activity data to account for repeated measures these variables. A third model adjusted for model 2 covariates plus cumulative average BMI. Cumulative averages were calculated for energy intake, physical activity and BMI in the same way as dietary pattern scores, using data from years 0, 7 and 20 or until censoring. Since family history of diabetes was missing for 698 participants (15%), we repeated the analysis using both models in the subset of the population who had this data ( $n = 4021$ ). Prior to adjustment for BMI, EC models were additionally adjusted for cumulative average fruit, vegetable, whole grain, red meat, fish and dairy intake. Dietary patterns scores were also examined individually as continuous variables to test for linear trend, with the HRs calculated per SD of the score.

We tested for effect modification by sex, race, education, BMI and smoking using models that included an interaction term for the variable of interest and each dietary pattern score separately and by stratification. To evaluate how a change in diet quality over time might impact type 2 diabetes risk, we conducted a sensitivity analysis in individuals with baseline and year 20 diet data who were free of diabetes through to year 20. An individual's diet was categorised as stable low or stable high if their baseline score was less than or greater than, respectively, the median and the difference between their year 20 and baseline score was within 1 SD of the population baseline score. An increase or decrease in diet quality was categorised as a difference between year 20 and baseline scores greater than or less than 1 SD. As a secondary analysis to inform the interpretation of the EC score, we examined the relationship between quartiles of per cent of total energy intake from added sugar and saturated fats from the year 20 dietary assessment in a subgroup of individuals with available dietary data who were free of type 2 diabetes through year 20 ( $n = 2436$ ). To avoid potential misclassification, a sensitivity analysis was performed excluding participants who developed diabetes before the age of 30 years ( $n = 8$ ). The proportional hazards assumption was tested by including an interaction term with  $\log_e$ -transformed time for each covariate. There was no evidence that the assumption was violated in any of the models. All analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA).

## Results

Demographic, socioeconomic, lifestyle and clinical cardiometabolic characteristics by quartiles of dietary pattern scores are presented in Tables 1, 2, 3 and 4. The 2015 DGA, APDQS and Palaeo scores were positively correlated and the EC score was negatively correlated with all other dietary pattern scores (Table 5). Descriptive dietary characteristics of the scores are presented in ESM Tables 3–6. ESM Table 7 displays 20 year

**Table 1** Baseline characteristics of CARDIA study participants according to year 0, 7 and 20 cumulative average 2015 DGA scientific report diet score quartile

Characteristic	Quartile of DGA 2015 score				<i>p</i> value <sup>a</sup>
	Q1	Q2	Q3	Q4	
<i>n</i>	1167	1182	1199	1171	
DGA 2015 score <sup>b</sup>	38.3 (4.0)	47.0 (2.1)	54.3 (2.3)	65.2 (5.2)	<0.0001
Score range	20.0–43.3	43.5–50.5	50.7–58.3	58.5–86.5	
Age, years	23.5 (3.7)	24.6 (3.7)	25.3 (3.5)	25.9 (3.2)	<0.0001
Race, % white	25.1	40.1	55.8	79.8	<0.0001
Sex, % male	60.5	48.9	41.9	30.1	<0.0001
Education, years	13.0 (4.7)	13.6 (4.1)	14.2 (3.2)	15.1 (3.3)	<0.0001
Smoking status, %					
Never	55.2	56.4	59.8	58.8	
Former	6.4	10.2	15.8	21.4	<0.0001
Current	38.4	33.4	24.4	19.8	
Alcohol intake (ml/day)	13.8 (24.0)	12.3 (22.3)	10.0 (15.4)	10.2 (15.2)	<0.0001
Baseline physical activity, EU/week <sup>c</sup>	386.5 (296.1)	376.4 (283.7)	422.9 (292.6)	475.2 (288.5)	<0.0001
Cumulative average physical activity (EU/week) <sup>d</sup>	344.1 (244.8)	337.8 (232.0)	382.1 (238.8)	429.7 (231.5)	<0.0001
BMI, kg/m <sup>2</sup>	24.4 (5.3)	25.1 (5.4)	24.8 (5.0)	23.5 (4.1)	<0.0001
Cumulative average BMI (kg/m <sup>2</sup> ) <sup>d</sup>	26.4 (5.7)	27.1 (5.8)	26.7 (5.5)	25.0 (4.8)	<0.0001
Family history of diabetes, % <sup>e</sup>	17.7	19.5	15.5	12.1	<0.0001
Fasting serum glucose, mmol/l	4.5 (0.6)	4.5 (0.5)	4.6 (0.5)	4.6 (0.4)	0.46
Serum insulin, pmol/l	79.9 (56.9)	82.0 (64.6)	73.6 (50.7)	64.5 (43.1)	<0.0001
Blood pressure, mmHg					
Systolic	111.7 (10.8)	111.2 (10.8)	110.2 (10.9)	108.2 (10.7)	<0.0001
Diastolic	68.7 (10.0)	68.8 (9.8)	68.8 (9.6)	68.0 (8.9)	0.11
Plasma triacylglycerols, mmol/l	0.8 (0.5)	0.8 (0.6)	0.8 (0.5)	0.8 (0.6)	0.4
Plasma cholesterol, mmol/l					
Total cholesterol	4.5 (0.9)	4.6 (0.9)	4.7 (0.9)	4.6 (0.8)	0.001
LDL-cholesterol	2.8 (0.8)	2.9 (0.8)	2.9 (0.8)	2.8 (0.8)	0.0001
HDL-cholesterol	1.3 (0.3)	1.3 (0.3)	1.4 (0.3)	1.4 (0.3)	<0.0001

Data are presented as unadjusted mean (SD) unless noted as percentage

<sup>a</sup> *p* value was determined by ANOVA (age, education, physical activity, alcohol, BMI, glucose, blood pressure, lipids) or  $\chi^2$  test (race, sex, smoking, family history) of association between dietary pattern quartile and characteristic

<sup>b</sup> Mean (SD) 2015 DGA score for the study population was 51.2 (10.5) with a range of 20–86.5

<sup>c</sup> Physical activity score derived from the CARDIA physical activity history, where 300 EU is approximately equal to 150 min of moderate-intensity physical activity per week

<sup>d</sup> Cumulative average reflects the mean of data from CARDIA study examinations at year 0, 7 and 20

<sup>e</sup> Family history data were unavailable for 698 participants

changes in the 2015 DGA and APDQS score in participants who had baseline and year 20 diet data and who were free from diabetes through year 20.

A total of 680 incident cases of diabetes occurred during follow-up (mean [SD] 25.3 [8.3] years). As presented in Table 6, there was no association between the 2015 DGA, Palaeo and EC scores and type 2 diabetes risk in either multivariable model but there was a strong inverse association between higher APDQS score and type 2 diabetes risk.

The results for the EC score were consistent using both the score calculated per 4184 kJ of total energy intake and absolute total servings of foods high in ECs (data not presented). In a secondary analysis, a higher percentage of total energy intake from added sugar and saturated fat at year 20 was associated with an increased risk for type 2 diabetes (HR 1.16 [95% CI 1.04, 1.30]; *p* < 0.01) but the association was attenuated by adjustment for lifestyle factors, diet quality and BMI (ESM Table 8).

**Table 2** Baseline characteristics of CARDIA study participants according to year 0, 7 and 20 cumulative average Palaeo diet score quartile

Characteristic	Quartile of Palaeo score				<i>p</i> value <sup>a</sup>
	Q1	Q2	Q3	Q4	
<i>n</i>	1183	1200	1160	1176	
Palaeo score <sup>b</sup>	62.5 (4.0)	70.4 (1.6)	75.7 (1.5)	83.4 (4.2)	<0.0001
Score range	44.0–67.3	67.5–73.0	73.3–78.3	78.5–105.0	
Age, years	23.7 (3.7)	24.5 (3.7)	25.2 (3.5)	25.9 (3.2)	<0.0001
Race, % white	39.6	45.0	53.9	62.6	<0.0001
Sex, % male	65.4	47.8	39.0	28.8	<0.0001
Education, years	13.5 (5.4)	13.8 (4.7)	14.1 (2.2)	14.6 (2.3)	<0.0001
Smoking status, %					
Never	57.1	58.8	56.6	57.6	
Former	9.6	11.2	15.2	18.0	<0.0001
Current	33.3	30.0	28.2	24.4	
Alcohol intake, ml/day	13.9 (23.3)	11.2 (19.3)	10.8 (18.7)	10.3 (16.7)	<0.0001
Physical activity, EU/week <sup>c</sup>	446.1 (311.4)	390.8 (277.4)	410.9 (292.6)	418.4 (286.1)	<0.0001
Cumulative average physical activity (EU/week) <sup>d</sup>	401.1 (254.5)	347.4 (226.7)	366.4 (239.5)	378.0 (233.8)	<0.0001
BMI, kg/m <sup>2</sup>	24.3 (4.9)	24.4 (5.1)	24.5 (4.9)	24.7 (5.0)	0.28
Cumulative average BMI (kg/m <sup>2</sup> ) <sup>d</sup>	26.1 (5.4)	26.3 (5.6)	26.4 (5.4)	26.3 (5.7)	0.65
Family history of diabetes, % <sup>e</sup>	17.0	14.6	16.7	16.0	0.49
Fasting serum glucose, mmol/l	4.5 (0.5)	4.6 (0.6)	4.5 (0.5)	4.5 (0.4)	0.74
Serum insulin, pmol/l	75.0 (53.5)	77.1 (59.0)	75.0 (56.9)	72.2 (50.0)	0.14
Blood pressure, mmHg					
Systolic	111.9 (10.8)	110.7 (10.6)	109.6 (10.8)	109.0 (11.2)	<0.0001
Diastolic	69.0 (10.1)	68.5 (9.6)	68.2 (9.5)	68.5 (9.2)	0.32
Plasma triacylglycerols, mmol/l	0.8 (0.5)	0.8 (0.5)	0.8 (0.5)	0.8 (0.6)	0.86
Plasma cholesterol, mmol/l					
Total cholesterol	4.5 (0.8)	4.6 (0.9)	4.6 (0.9)	4.7 (0.9)	<0.0001
LDL-cholesterol	2.8 (0.8)	2.8 (0.8)	2.8 (0.8)	2.9 (0.8)	0.005
HDL-cholesterol	1.3 (0.3)	1.4 (0.3)	1.4 (0.3)	1.4 (0.4)	<0.0001

Data are presented as unadjusted mean (SD) unless noted as percentage

<sup>a</sup> *p* value was determined by ANOVA (age, education, physical activity, alcohol, BMI, glucose, BP, lipids) or  $\chi^2$  test (race, sex, smoking, family history) of association between dietary pattern quartile and characteristic

<sup>b</sup> Mean (SD) Palaeo score for the study population was 73.0 (8.2) with a range of 44–105

<sup>c</sup> Physical activity score derived from the CARDIA physical activity history, where 300 EU is approximately equal to 150 min of moderate-intensity physical activity per week

<sup>d</sup> Cumulative average reflects the mean of data from CARDIA study examinations at year 0, 7 and 20

<sup>e</sup> Family history data were unavailable for 698 participants

Repeated measures of the dietary pattern scores were correlated over time ( $r = 0.34$ – $0.65$ ) (ESM Table 9). The association between 20 year change in dietary pattern score and type 2 diabetes risk is presented in ESM Table 10. Individuals whose DGA score increased between year 0 and year 20 by more than one SD of year 0 scores (12.0) had a 44% lower risk of type 2 diabetes compared with those with stable DGA scores below the population median. Individuals with stable high (>median) Palaeo scores had a 41% lower risk of type 2 diabetes than those with stable low scores (per SD, HR 0.59 [95% CI 0.39, 0.88]). There was no association between

changes in APDQS and EC scores as formulated in this analysis and type 2 diabetes risk.

Sensitivity analyses provided evidence for potential effect measure modification by smoking status in the DGA analysis ( $p_{\text{interaction}} = 0.007$ ). When the DGA score was stratified on smoking status (current smokers vs non-smokers) with HRs calculated per SD of diet score, the 2015 DGA score was inversely associated with type 2 diabetes risk in non-smokers ( $n = 2716$ ) (HR 0.86 [95% CI 0.74, 0.99];  $p = 0.03$ ) but no association was observed for current smokers ( $n = 1368$ ) (HR 0.92 [95% CI 0.77, 1.10];  $p = 0.36$ ). In stratified

**Table 3** Baseline characteristics of CARDIA study participants according to year 0, 7 and 20 cumulative average APDQS quartile

Characteristic	Quartile of APDQS score				<i>p</i> value <sup>a</sup>
	Q1	Q2	Q3	Q4	
<i>n</i>	1188	1170	1188	1173	
APDQS score <sup>b</sup>	50.6 (4.6)	60.7 (2.3)	69.0 (2.6)	80.8 (5.2)	<0.0001
Score range	31.0–56.5	56.7–64.5	64.7–73.7	74.0–100.5	
Age, years	23.3 (3.7)	24.5 (3.6)	25.3 (3.4)	26.1 (3.1)	<0.0001
Race, % white	20.8	36.2	58.5	85.7	<0.0001
Sex, % male	53.4	46.8	45.7	35.2	<0.0001
Education, years	13.0 (5.3)	13.4 (3.1)	14.2 (2.2)	15.4 (4.1)	<0.0001
Smoking status, %					
Never	57.7	55.8	59.5	57.1	
Former	6.3	11.0	14.1	22.5	<0.0001
Current	35.9	33.2	26.4	20.4	
Alcohol intake, ml/day	9.3 (19.6)	11.9 (22.8)	11.6 (18.2)	13.4 (17.6)	<0.0001
Physical activity, EU/week <sup>c</sup>	363.4 (288.6)	372.0 (284.3)	424.9 (281.1)	501.1 (296.4)	<0.0001
Cumulative average physical activity (EU/week) <sup>d</sup>	325.3 (242.3)	338.3 (231.0)	380.6 (231.0)	448.9 (233.7)	<0.0001
BMI, kg/m <sup>2</sup>	24.7 (5.6)	25.1 (5.4)	24.5 (4.7)	23.5 (4.0)	<0.0001
Cumulative average BMI (kg/m <sup>2</sup> ) <sup>d</sup>	26.6 (6.0)	27.2 (5.8)	26.5 (5.2)	25.0 (4.7)	<0.0001
Family history of diabetes, % <sup>e</sup>	19.4	18.4	14.1	13.0	<0.0001
Fasting serum glucose, mmol/l	4.5 (0.6)	4.5 (0.5)	4.5 (0.4)	4.6 (0.4)	0.12
Serum insulin, pmol/l	85.4 (65.3)	80.6 (59.7)	71.5 (48.6)	61.8 (39.6)	<0.0001
Blood pressure, mmHg					
Systolic	111.3 (11.0)	111.2 (10.8)	110.3 (10.7)	108.5 (10.8)	<0.0001
Diastolic	68.4 (10.2)	69.0 (9.6)	68.6 (9.7)	68.2 (8.8)	0.32
Plasma triacylglycerols, mmol/l	0.8 (0.5)	0.8 (0.5)	0.8 (0.6)	0.8 (0.5)	0.15
Plasma cholesterol, mmol/l					
Total cholesterol	4.5 (0.9)	4.6 (0.9)	4.6 (0.9)	4.6 (0.8)	0.001
LDL-cholesterol	2.8 (0.8)	2.9 (0.8)	2.9 (0.8)	2.8 (0.8)	0.004
HDL-cholesterol	1.3 (0.3)	1.4 (0.3)	1.4 (0.3)	1.4 (0.3)	<0.0001

Data are presented as unadjusted mean (SD) unless noted as percentage

<sup>a</sup> *p* value was determined by ANOVA (age, education, physical activity, alcohol, BMI, glucose, BP, lipids) or  $\chi^2$  test (race, sex, smoking, family history) of association between dietary pattern quartile and characteristic

<sup>b</sup> Mean (SD) APDQS score for the study population was 65.2 (11.7) with a range of 31–100.5

<sup>c</sup> Physical activity score derived from the CARDIA physical activity history, where 300 EU is approximately equal to 150 min of moderate-intensity physical activity per week

<sup>d</sup> Cumulative average reflects the mean of data from CARDIA study examinations at year 0, 7 and 20

<sup>e</sup> Family history data were unavailable for 698 participants

analysis by education level, we observed an inverse association in participants with a college degree or higher (per SD, HR 0.75 [95% CI 0.61, 0.92]) but no association in those without a degree. There was no evidence that the association between cumulative average Palaeo, APDQS or EC scores and type 2 diabetes risk differed by smoking status. Stratified analyses, as well as formal tests for interaction between dietary pattern scores and race, sex, BMI and family history of diabetes, provided no evidence of effect modification by these factors. For the EC score, there was an inverse association in white women (HR 0.76 [95% CI 0.60, 0.96])

but null associations in all other groups. Results excluding incident diabetes cases documented before age 30 years ( $n = 8$ ) were not materially different from the main analysis.

## Discussion

In the CARDIA study, a cumulative average DGA 2015 score was not associated with type 2 diabetes in the main models but in pre-specified stratified analyses there was an inverse association in participants who did not smoke and in participants

**Table 4** Characteristics of CARDIA study participants according to year 0, 7 and 20 cumulative average empty calorie score quartile

Characteristic	Quartile of EC score				<i>p</i> value <sup>a</sup>
	Q1	Q2	Q3	Q4	
<i>n</i>	1119	1244	1194	1162	
EC score <sup>b</sup>	16.6 (2.6)	21.6 (1.0)	24.8 (0.9)	29.1 (2.1)	<0.0001
Score range	2.0–19.7	20.0–23.0	23.3–26.5	26.7–38.0	
Age, years	25.3 (3.5)	25.0 (3.6)	24.6 (3.6)	24.3 (3.7)	<0.0001
Race, % white	60.4	52.7	46.6	41.5	<0.0001
Sex, % male	32.5	39.6	47.5	61.6	<0.0001
Education, years	14.2 (3.4)	14.1 (4.1)	13.9 (4.1)	13.8 (4.2)	0.06
Smoking status, %					
Never	54.9	56.8	58.2	60.2	
Former	18.0	12.6	13.0	10.5	<0.0001
Current	27.2	30.6	28.8	29.3	
Alcohol intake, ml/day	10.8 (20.6)	11.7 (20.9)	11.0 (19.5)	12.7 (17.6)	0.08
Physical activity, EU/week <sup>c</sup>	407.9 (292.3)	407.5 (290.4)	399.1 (286.8)	447.1 (308.7)	0.0003
Cumulative average physical activity (EU/week) <sup>d</sup>	375.2 (239.3)	363.6 (229.2)	358.5 (234.3)	396.6 (254.0)	0.0005
BMI, kg/m <sup>2</sup>	24.7 (5.1)	24.5 (5.0)	24.3 (5.0)	24.3 (4.8)	0.18
Cumulative average BMI (kg/m <sup>2</sup> ) <sup>d</sup>	26.4 (5.7)	26.4 (5.6)	26.2 (5.4)	26.2 (5.3)	0.77
Family history of diabetes, % <sup>e</sup>	15.4	17.0	15.2	16.5	0.64
Fasting serum glucose, mmol/l	4.5 (0.4)	4.5 (0.6)	4.5 (0.5)	4.5 (0.5)	0.82
Serum insulin, pmol/l	73.1 (52.1)	75.5 (57.4)	76.4 (55.5)	74.6 (54.7)	0.54
Blood pressure, mmHg					
Systolic	109.0 (11.0)	110.5 (11.3)	110.4 (10.5)	111.4 (10.6)	<0.0001
Diastolic	68.0 (9.6)	68.7 (9.5)	68.5 (9.2)	68.9 (10.0)	0.18
Plasma triacylglycerols, mmol/l	0.8 (0.5)	0.8 (0.6)	0.8 (0.5)	0.8 (0.6)	0.48
Plasma cholesterol, mmol/l					
Total cholesterol	4.6 (0.9)	4.6 (0.9)	4.6 (0.8)	4.6 (0.9)	0.70
LDL-cholesterol	2.8 (0.8)	2.8 (0.8)	2.8 (0.8)	2.8 (0.8)	0.87
HDL-cholesterol	1.4 (0.3)	1.4 (0.3)	1.4 (0.3)	1.3 (0.3)	0.0003

Data are presented as unadjusted mean (SD) for all characteristics unless noted as a percentage

<sup>a</sup> *p* value was determined by ANOVA (age, education, physical activity, alcohol, BMI, glucose, BP, lipids) or  $\chi^2$  test (race, sex, smoking, family history) of association between dietary pattern quartile and characteristic

<sup>b</sup> Mean (SD) EC score for the study population was 23.1 (4.9) with a range of 2–38

<sup>c</sup> Physical activity score derived from the CARDIA physical activity history, where 300 EU is approximately equal to 150 min of moderate-intensity physical activity per week

<sup>d</sup> Cumulative average reflects the mean of data from CARDIA study examinations at year 0, 7 and 20

<sup>e</sup> Family history data were unavailable for 698 participants

**Table 5** Pearson's correlations between dietary pattern scores

Dietary pattern	APDQS	DGA 2015	Palaeolithic	Empty calorie
APDQS	1.00	0.89	0.57	−0.39
DGA 2015	0.89	1.00	0.67	−0.53
Palaeo	0.57	0.67	1.00	−0.66
Empty calorie	−0.39	−0.53	−0.66	1.00

*p*<0.0001 for all correlations

with a college degree, as well as in a subgroup of participants who increased their score over 20 years. The Palaeo score was not associated with type 2 diabetes in the main models but in a sensitivity analysis participants with a consistently high score over 20 years had a lower risk of type 2 diabetes. The EC score was not associated with type 2 diabetes in the main models but we did observe an inverse association between higher EC scores and type 2 diabetes risk in white women in the CARDIA study, contrary to any hypothesis. Last, the cohort reference APDQS score was strongly inversely associated with type 2 diabetes.



**Table 6** Association between year 0, 7 and 20 cumulative average dietary patterns and 30 year diabetes risk in young adult men and women from the CARDIA study, 1985–2015

Score	Dietary pattern score quartile <sup>a</sup>				Per SD	<i>p</i> value <sup>b</sup>
	Q1	Q2	Q3	Q4		
<b>2015 DGA</b>						
Type 2 diabetes incidence, <i>n</i> /person-years	212/26,780	196/28,314	161/29,748	111/29,881		
Model 1 <sup>c</sup>	1.0	0.91 (0.75, 1.11)	0.76 (0.61, 0.95)	0.61 (0.46, 0.80)	0.82 (0.74, 0.90)	<0.0001
Model 2 <sup>d</sup>	1.0	0.99 (0.80, 1.21)	0.92 (0.73, 1.17)	0.82 (0.62, 1.10)	0.93 (0.83, 1.03)	0.16
Model 3 <sup>e</sup>	1.0	0.96 (0.78, 1.17)	0.90 (0.71, 1.13)	0.88 (0.66, 1.17)	0.94 (0.84, 1.05)	0.28
<b>Palaeo</b>						
Type 2 diabetes incidence, <i>n</i> /person-years	183/27,946	173/29,822	164/28,571	160/28,384		
Model 1 <sup>c</sup>	1.0	0.90 (0.72, 1.11)	0.93 (0.74, 1.15)	0.97 (0.77, 1.22)	0.99 (0.91, 1.07)	0.74
Model 2 <sup>d</sup>	1.0	0.93 (0.75, 1.15)	1.00 (0.79, 1.26)	1.10 (0.85, 1.41)	1.04 (0.94, 1.14)	0.49
Model 3 <sup>e</sup>	1.0	0.92 (0.74, 1.15)	1.00 (0.80, 1.26)	1.07 (0.83, 1.37)	1.02 (0.93, 1.12)	0.72
<b>APDQS</b>						
Type 2 diabetes incidence, <i>n</i> /person-years	230/25,788	195/28,366	161/29,866	94/30,703		
Model 1 <sup>c</sup>	1.0	0.74 (0.61, 0.90)	0.63 (0.50, 0.78)	0.39 (0.29, 0.52)	0.69 (0.62, 0.76)	<0.0001
Model 2 <sup>d</sup>	1.0	0.79 (0.65, 0.97)	0.74 (0.59, 0.92)	0.52 (0.38, 0.70)	0.76 (0.68, 0.85)	<0.0001
Model 3 <sup>e</sup>	1.0	0.77 (0.63, 0.94)	0.74 (0.59, 0.92)	0.55 (0.41, 0.74)	0.78 (0.70, 0.86)	<0.0001
<b>EC</b>						
Type 2 diabetes incidence, <i>n</i> /person-years	149/26,336	192/30,097	174/29,394	165/28,896		
Model 1 <sup>c</sup>	1.0	1.05 (0.85, 1.30)	0.91 (0.73, 1.14)	0.84 (0.67, 1.06)	0.94 (0.87, 1.02)	0.13
Model 2 <sup>d</sup>	1.0	1.05 (0.85, 1.31)	0.90 (0.72, 1.13)	0.83 (0.65, 1.05)	0.93 (0.86, 1.02)	0.10
Model 3 <sup>f</sup>	1.0	1.04 (0.83, 1.29)	0.89 (0.71, 1.12)	0.83 (0.65, 1.06)	0.93 (0.85, 1.02)	0.13
Model 4 <sup>g</sup>	1.0	1.08 (0.87, 1.34)	0.93 (0.74–1.17)	0.86 (0.68, 1.11)	0.95 (0.87, 1.04)	0.28

<sup>a</sup> Higher 2015 DGA, Palaeo, and APDQS scores (upper quartiles) theoretically reflect higher diet quality, whereas higher empty calorie scores suggest less healthful dietary components are more abundant in the diet

<sup>b</sup> *p* values were determined by modelling the dietary pattern scores as continuous variables in Cox proportional hazards multivariable regression models and correspond to the HR (95% CI) per SD of the dietary pattern score

<sup>c</sup> Model 1: HR (95% CI) derived from Cox proportional hazards models adjusted for potential confounding by age, race, sex and CARDIA study centre

<sup>d</sup> Model 2: HR (95% CI) derived from Cox proportional hazards models adjusted for Model 1 covariates plus repeated measures of education and smoking status, cumulative average physical activity and cumulative average estimated total energy intake

<sup>e</sup> Model 3: HR (95% CI) derived from Cox proportional hazards models adjusted for Model 2 covariates and cumulative average BMI

<sup>f</sup> Model 3: HR (95% CI) derived from Cox proportional hazards models adjusted for EC Model 2 covariates and cumulative average fruit, vegetable, whole grain, red meat, fish and dairy intake

<sup>g</sup> Model 4: HR (95% CI) derived from Cox proportional hazards models adjusted for EC Model 3 covariates and cumulative average BMI

We are not aware of published research examining the relationship between a dietary pattern score based on the 2015 Scientific Report and incident type 2 diabetes. Previous observational studies of dietary pattern scores created to reflect earlier versions of the DGA have produced inconclusive results. Zamora et al. found no association between the 2005 Diet Quality Index and 20 year type 2 diabetes risk in the CARDIA study [24]. Other prospective studies also found no association between Healthy Eating Index (HEI) scores, reflecting both the 2005 and 2010 DGAs, and type 2 diabetes risk [25, 26]. However, the Alternate Healthy Eating Index (AHEI), which reflects the prior versions of the DGA and also includes foods and nutrients predictive of chronic diseases [27], has been associated with a decreased risk for type 2

diabetes in some study populations [25, 26, 28] but not others [29]. The heterogeneity in the overall body of evidence related to testing the DGA recommendations, including this study, suggests the need to account for the context of the population under study and the known limitations of observational dietary research.

Research examining the relationship between a modern-day Palaeo dietary pattern and incident type 2 diabetes is also lacking. Proponents of the Palaeo diet assert that humans are genetically adapted to foods available prior to changes in the food supply with the establishment of agriculture and that core metabolic processes central to common chronic disease aetiology are misaligned with these dietary changes [19, 30]. There was no association between a cumulative average modern-day Palaeo

dietary pattern score and type 2 diabetes risk in the main models. However, in a subgroup analysis, participants with scores >1 SD above the median at year 0 and 20 had a lower risk of type 2 diabetes compared with participants with stable low scores. This suggests that a diet more closely aligned with a Palaeo dietary pattern may have traction over time but it is necessary to consider whether this was due to potential selection biases. This pattern differs from most dietary recommendations by eschewing food groups with evidence of being protective for type 2 diabetes risk (legumes, whole grains, dairy) [31–33] and emphasising animal protein, which in the context of a typical western diet is largely considered a dietary factor that increases type 2 diabetes risk [34].

The objective of examining an EC score was to investigate a largely untested concept. We hypothesised that there would be a positive association between higher EC intake and diabetes risk but there was no association in the main models. Indeed, there was an un-hypothesised inverse association in white women in the CARDIA study, suggesting potential confounding. There is a scientific basis for the recommendation of avoidance of an EC dietary pattern but the definition is nebulous and contains heterogeneous foods/nutrients. Evidence informing the concept is largely reductionist in nature or extrapolated, which makes performing rigorous science around the concept difficult to conduct and interpret.

Although the 2015 DGA and APDQS were highly correlated, it is important to contextualise their differences since imprecision surrounding the components due to the limitations of self-reported dietary data may have impacted the strength of the associations detectable in this cohort. First, alcohol was scored moderately in the 2015 DGA score and positively in the APDQS. Foods scored positively in the APDQS were not included in the 2015 DGA score (e.g. oil, poultry, coffee and tea). In addition, the APDQS negatively scored butter, fried foods and whole-fat dairy products, which were not included in the 2015 DGA score, while fruit juice, margarine and refined grains were negatively scored in the 2015 DGA score but were not included in the APDQS. The similarity between the scores included positive scoring of fruits, vegetables, avocado, nuts and seeds, legumes, fish, low-fat dairy products and whole grains and negative scoring of red and processed meats, sugar-sweetened foods and beverages, fried potatoes and salty snacks. Individuals in the highest quartile of the APDQS consumed approximately 3 g fibre/day more than those in the highest quartile of the DGA score (ESM Tables 3, 5). The inverse association between dietary fibre intake and type 2 diabetes risk is well established [35]; however, the data used in this study are not suitable for further inference on this observation. The study of dietary patterns as a whole, as opposed to the reductionist approach of examining single foods or food groups in isolation, has been increasingly recognised as the optimal approach to study diet–disease relationships because patterns are better suited to account for the correlated nature of dietary data and theoretical synergistic effects of total diet on health [36].

Smoking has been identified as a strong, independent risk factor for diabetes [37] and previous studies suggest that smoking may blunt the association between diet and diabetes risk [38–40]. Multiple pathophysiological pathways have been implicated in the association between smoking and type 2 diabetes [41–43]. Smoking may also influence the detection of diabetes through its effect on oral and intravenous glucose tolerance tests [44]. The association in smokers may further be confounded by the clustering of unhealthful behaviours and socioeconomic status [37]. These potential mechanisms may help explain the differential association in non-smokers vs smokers for the 2015 DGA.

Although the statistical models were comprehensive, residual confounding occurs to some extent in all diet–disease observational studies. The use of self-reported dietary data are also subject to recall and other biases that may alter estimates, potentially resulting in false-positive or -negative associations. The validity and reliability of the CARDIA Diet History have been demonstrated [11, 12]. Any null findings were not due to a lack of statistical power as this study were adequately powered to detect an association between dietary intake and type 2 diabetes, though a larger sample size could improve the precision of the estimates given these limitations [45].

In conclusion, adherence to contemporary dietary recommendations (2015 DGA) were inversely associated with risk of type 2 diabetes in non-smokers and individuals with a college degree. A modern-day Palaeo dietary pattern score was not associated with risk of type 2 diabetes except in a subgroup of participants who maintained a high score over 20 years. A score created to reflect high EC intake was not associated with type 2 diabetes in main models but there was a surprising inverse association between empty calorie intake and type 2 diabetes in white women in the cohort. The cohort reference APDQS, which was highly correlated with the 2015 DGA and similarly scored, was strongly inversely associated with type 2 diabetes. Overall, examination of dietary recommendations and trends in the CARDIA cohort and their relationship with risk for type 2 diabetes produced nuanced results that should be considered in the context of different potential biases.

**Data availability** The data that support the findings of this study are available from the CARDIA Coordinating Center but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available (<https://www.cardia.dopm.uab.edu/publications-2>).

**Funding** The CARDIA study is supported by contracts HHSN268201800003I, HHSN268201800004I, HHSN268201800005I, HHSN268201800006I and HHSN268201800007I from the National Heart, Lung, and Blood Institute (NHLBI).

**Duality of interest** DJ declares consultancy fees and honoraria received from the California Walnut Commission. All other authors declare no duality of interest associated with their involvement with this manuscript.

**Contribution statement** KH and AO are responsible for the concept and design of the study, analysis and interpretation and writing of the

manuscript. KH, AO and DJ are responsible for data acquisition. DJ, LJ, NW, and JS contributed to the analysis and interpretation of the data and provided a critical review of the manuscript. All authors have read and approved the final version of the manuscript. KH and AO are the guarantors of this work and, as such, had full access to the data provided by the CARDIA study and take responsibility for the integrity and accuracy of the data analysis.

## References

- World Health Organization (2016) Global report on diabetes. WHO, Geneva
- Centers for Disease Control and Prevention (2017) National Diabetes Statistics Report, 2017. Centers for Disease Control and Prevention US Department of Health and Human Services, Atlanta
- Alhazmi A, Stojanovski E, McEvoy M, Garg ML (2014) The association between dietary patterns and type 2 diabetes: a systematic review and meta-analysis of cohort studies. *J Hum Nutr Diet* 27(3): 251–260. <https://doi.org/10.1111/jhn.12139>
- Jannasch F, Kröger J, Schulze MB (2017) Dietary patterns and type 2 diabetes: a systematic literature review and meta-analysis of prospective studies. *J Nutr* 147(6):1174–1182. <https://doi.org/10.3945/jn.116.242552>
- Dietary Guidelines Advisory Committee (2015) Scientific report of the 2015 Dietary Guidelines Advisory Committee: advisory report to the secretary of health and human services and the secretary of agriculture. U.S. Department of Agriculture, Agricultural Research Service, Washington, DC. <https://doi.org/10.1017/CBO9781107415324.004>
- Meyer KA, Sijtsma FP, Nettleton JA et al (2013) Dietary patterns are associated with plasma F(2)-isoprostanes in an observational cohort study of adults. *Free Radic Biol Med* 57:201–209. <https://doi.org/10.1016/j.freeradbiomed.2012.08.574>
- Shikany JM, Jacobs DR Jr, Lewis CE et al (2013) Associations between food groups, dietary patterns, and cardiorespiratory fitness in the Coronary Artery Risk Development in Young Adults study. *Am J Clin Nutr* 98(6):1402–1409. <https://doi.org/10.3945/ajcn.113.058826>
- Sijtsma FPC, Meyer KA, Steffen LM et al (2014) Diet quality and markers of endothelial function: the CARDIA study. *Nutr Metab Cardiovasc Dis* 24(6):632–638. <https://doi.org/10.1016/j.numecd.2013.12.010>
- Buijsse B, Jacobs DR Jr, Steffen LM, Kromhout D, Gross MD (2015) Plasma ascorbic acid, a priori diet quality score, and incident hypertension: a prospective cohort study. *PLoS One* 10(12): e0144920. <https://doi.org/10.1371/journal.pone.0144920>
- Friedman GD, Cutter GR, Donahue RP et al (1988) CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol* 41(11):1105–1116. [https://doi.org/10.1016/0895-4356\(88\)90080-7](https://doi.org/10.1016/0895-4356(88)90080-7)
- McDonald A, Van Horn L, Slattery M et al (1991) The CARDIA dietary history: development, implementation, and evaluation. *J Am Diet Assoc* 91(9):1104–1112
- Liu K, Slattery M, Jacobs D Jr et al (1994) A study of the reliability and comparative validity of the cardia dietary history. *Ethn Dis* 4(1): 15–27
- Sijtsma FP, Meyer KA, Steffen LM et al (2012) Longitudinal trends in diet and effects of sex, race, and education on dietary quality score change: the Coronary Artery Risk Development in Young Adults study. *Am J Clin Nutr* 95(3):580–586. <https://doi.org/10.3945/ajcn.111.020719>
- Jacobs DR Jr, Sluik D, Rokling-Andersen MH, Anderssen SA, Drevon CA (2009) Association of 1-y changes in diet pattern with cardiovascular disease risk factors and adipokines: results from the 1-y randomized Oslo Diet and Exercise Study. *Am J Clin Nutr* 89(2):509–517. <https://doi.org/10.3945/ajcn.2008.26371>
- Lockheart MS, Steffen LM, Rebnord HM et al (2007) Dietary patterns, food groups and myocardial infarction: a case-control study. *Br J Nutr* 98(2):380–387. <https://doi.org/10.1017/S0007114507701654>
- Nettleton JA, Schulze MB, Jiang R, Jenny NS, Burke GL, Jacobs DR Jr (2008) A priori-defined dietary patterns and markers of cardiovascular disease risk in the Multi-Ethnic Study of Atherosclerosis (MESA). *Am J Clin Nutr* 88(1):185–194. <https://doi.org/10.1093/ajcn/88.1.185>
- Jönsson T, Granfeldt Y, Åhrén B et al (2009) Beneficial effects of a Paleolithic diet on cardiovascular risk factors in type 2 diabetes: a randomized cross-over pilot study. *Cardiovasc Diabetol* 8(1):35. <https://doi.org/10.1186/1475-2840-8-35>
- Whalen KA, McCullough M, Flanders WD, Hartman TJ, Judd S, Bostick RM (2014) Paleolithic and Mediterranean diet pattern scores and risk of incident, sporadic colorectal adenomas. *Am J Epidemiol* 180(11):1088–1097. <https://doi.org/10.1093/aje/kwu235>
- Eaton SB, Konner M (1985) Paleolithic nutrition - a consideration of its nature and current implications. *N Engl J Med* 312(5):283–289. <https://doi.org/10.1056/NEJM198501313120505>
- Jacobs DR Jr, Hahn LP, Haskell WL, Pirie P, Sidney S (1989) Validity and reliability of a short physical activity history: CARDIA and the Minnesota Heart Health Program. *J Cardiopulm Rehabil* 9(12):448–459
- Parker ED, Schmitz KH, Jacobs DR Jr, Dengel DR, Schreiner PJ, Schreiner PJ (2007) Physical activity in young adults and incident hypertension over 15 years of follow-up: the CARDIA study. *Am J Public Health* 97(4):703–709. <https://doi.org/10.2105/AJPH.2004.055889>
- Cutter GR, Burke GL, Dyer AR et al (1991) Cardiovascular risk factors in young adults. The CARDIA baseline monograph. *Control Clin Trials* 12:1S–77S
- Wagenknecht LE, Burke GL, Perkins LL, Haley NJ, Friedman GD (1992) Misclassification of smoking status in the CARDIA study: a comparison of self-report with serum cotinine levels. *Am J Public Health* 82(1):33–36. <https://doi.org/10.2105/AJPH.82.1.33>
- Zamora D, Gordon-Larsen P, He K, Jacobs DR Jr, Shikany JM, Popkin BM (2011) Are the 2005 Dietary Guidelines for Americans associated with reduced risk of type 2 diabetes and cardiometabolic risk factors? Twenty-year findings from the CARDIA study. *Diabetes Care* 34(5):1183–1185. <https://doi.org/10.2337/dc10-2041>
- Jacobs S, Harmon BE, Boushey CJ et al (2015) A priori-defined diet quality indexes and risk of type 2 diabetes: the multiethnic cohort. *Diabetologia* 58(1):98–112. <https://doi.org/10.1007/s00125-014-3404-8>
- de Koning L, Chiuve SE, Fung TT, Willett WC, Rimm EB, Hu FB (2011) Diet-quality scores and the risk of type 2 diabetes in men. *Diabetes Care* 34(5):1150–1156. <https://doi.org/10.2337/dc10-2352>
- McCullough ML, Feskanich D, Stampfer MJ et al (2002) Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *Am J Clin Nutr* 76(6):1261–1271. <https://doi.org/10.1093/ajcn/76.6.1261>
- Chiuve SE, Fung TT, Rimm EB et al (2012) Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr* 142(6):1009–1018. <https://doi.org/10.3945/jn.111.157222>
- InterAct Consortium (2014) Adherence to predefined dietary patterns and incident type 2 diabetes in European populations: EPIC-

- InterAct Study. *Diabetologia* 57(2):321–333. <https://doi.org/10.1007/s00125-013-3092-9>
30. Eaton SB, Cordain L, Lindeberg S (2002) Evolutionary health promotion: a consideration of common counterarguments. *Prev Med (Baltim)* 34(2):119–123. <https://doi.org/10.1006/pmed.2001.0966>
  31. Hallfrisch J, Facn BKM (2000) Mechanisms of the effects of grains on insulin and glucose responses. *J Am Coll Nutr* 19(sup3):320S–325S. <https://doi.org/10.1080/07315724.2000.10718967>
  32. Higgins JA (2012) Whole grains, legumes, and the subsequent meal effect: implications for blood glucose control and the role of fermentation. *J Nutr Metab* 2012:829238–829237. <https://doi.org/10.1155/2012/829238>
  33. Bjørnshave A, Hermansen K (2014) Effects of dairy protein and fat on the metabolic syndrome and type 2 diabetes. *Rev Diabet Stud* 11(2):153–166. <https://doi.org/10.1900/RDS.2014.11.153>
  34. Barnard N, Levin S, Trapp C (2014) Meat consumption as a risk factor for type 2 diabetes. *Nutrients* 6(2):897–910. <https://doi.org/10.3390/nu6020897>
  35. McRae MP (2018) Dietary fiber intake and type 2 diabetes mellitus: an umbrella review of meta-analyses. *J Chiropr Med* 17(1):44–53. <https://doi.org/10.1016/j.jcm.2017.11.002>
  36. Jacobs DR Jr, Steffen LM (2003) Nutrients, foods, and dietary patterns as exposures in research: a framework for food synergy. *Am J Clin Nutr* 78(3):508S–513S. <https://doi.org/10.1093/ajcn/78.3.508S>
  37. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J (2007) Active smoking and the risk of type 2 diabetes. *JAMA* 298(22):2654–2664. <https://doi.org/10.1001/jama.298.22.2654>
  38. Sargeant LA, Khaw K-T, Bingham S et al (2001) Cigarette smoking and glycaemia: the EPIC-Norfolk Study. *Int J Epidemiol* 30(3):547–554. <https://doi.org/10.1093/ije/30.3.547>
  39. Odegaard AO, Koh WP, Butler LM et al (2011) Dietary patterns and incident type 2 diabetes in Chinese men and women: the Singapore Chinese health study. *Diabetes Care* 34(4):880–885. <https://doi.org/10.2337/dc10-2350>
  40. Hozawa A, Jacobs DR Jr, Steffes MW, Gross MD, Steffen LM, Lee DH (2006) Associations of serum carotenoid concentrations with the development of diabetes and with insulin concentration: interaction with smoking: the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Epidemiol* 163(10):929–937. <https://doi.org/10.1093/aje/kwj136>
  41. Dietrich M, Block G, Norkus EP et al (2003) Smoking and exposure to environmental tobacco smoke decrease some plasma antioxidants and increase  $\gamma$ -tocopherol in vivo after adjustment for dietary antioxidant intakes. *Am J Clin Nutr* 77(1):160–166. <https://doi.org/10.1093/ajcn/77.1.160>
  42. Li N, Frigerio F, Maechler P (2008) The sensitivity of pancreatic  $\beta$ -cells to mitochondrial injuries triggered by lipotoxicity and oxidative stress. *Biochem Soc Trans* 36(5):930–934. <https://doi.org/10.1042/BST0360930>
  43. Chiolero A, Faeh D, Paccaud F, Cornuz J (2008) Consequences of smoking for body weight, body fat distribution, and insulin resistance. *Am J Clin Nutr* 87(4):801–809. <https://doi.org/10.1093/ajcn/87.4.801>
  44. Janzon L, Berntorp K, Hanson M, Lindell SE, Trelle E (1983) Glucose tolerance and smoking: a population study of oral and intravenous glucose tolerance tests in middle-aged men. *Diabetologia* 25(2):86–88. <https://doi.org/10.1007/bf00250893>
  45. Willett WC (2013) *Nutritional epidemiology*, 3rd edn. Oxford University Press, New York

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.