Associations Between Dietary Patterns and Incident Type 2 Diabetes: Prospective Cohort Study of 120,343 UK Biobank Participants

Min Gao,[1] Susan A. Jebb,[1,2] Paul Aveyard,[1,2] Gina L. Ambrosini,[3] Aurora Perez-Cornago,[4] Keren Papier,[4] Jennifer Carter,[5] and Carmen Piernas[1]

Diabetes Care 2022;45:1315–1325 | https://doi.org/10.2337/dc21-2258

OBJECTIVE

To identify dietary patterns (DPs) characterized by a set of nutrients of concern and their association with incident type 2 diabetes (T2D).

RESEARCH DESIGN AND METHODS

A total of 120,343 participants from the U.K. Biobank study with at least two 24 h dietary assessments were studied. Reduced rank regression was used to derive DPs explaining variability in energy density, free sugars, saturated fat, and fiber intakes. We investigated prospective associations with T2D using Cox proportional hazard models.

RESULTS

Over 8.4 years of follow-up from the latest dietary assessment, 2,878 participants developed T2D. Two DPs were identified that jointly explained a total of 63% variation in four nutrients. DP1 was characterized by high intakes of chocolate and confectionery, butter, low-fiber bread, and sugars and preserves, and low intakes of fruits and vegetables. DP1 was linearly associated with T2D in multivariable models without BMI adjustment (per *z* score, hazard ratio [HR] 1.11 [95% CI 1.08–1.14]) and after BMI adjustment (HR 1.09 [95% CI 1.06–1.12]). DP2 was characterized by high intakes of sugar-sweetened beverages, fruit juice, table sugars and preserves, and low intakes of high-fat cheese and butter, but showed no clear association with T2D. There were significant interactions between both DPs and age, with increased risks among younger people in DP1 (HR 1.13 [95% CI 1.09–1.18]) and DP2 (HR 1.10 [95% CI 1.05–1.15]), as well as with DP1 and BMI, with increased risks among people with obesity (HR 1.11 [95% CI 1.07–1.16]).

CONCLUSIONS

A DP characterized by high intakes of chocolate and confectionery, butter, low-fiber bread, and added sugars, and low in fresh fruits and vegetables intake is associated with a higher incidence of T2D, particularly among younger people and those with obesity.

The prevalence of diabetes, predominantly type 2 diabetes, has increased from 463 million in 2019 and is projected to reach 700 million by 2045 worldwide (1). Type 2 diabetes

[1]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, U.K.
[2]National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospitals, Oxford, U.K.
[3]School of Population and Global Health, University of Western Australia, Perth, Western Australia, Australia
[4]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, U.K.
[5]Nuffield Department of Population Health, University of Oxford, Oxford, U.K.

Corresponding author: Carmen Piernas, carmen.piernas-sanchez@phc.ox.ac.uk

Received 30 October 2021 and accepted 21 February 2022

This article contains supplementary material online at https://doi.org/10.2337/figshare.19209750.

is associated with higher morbidity and mortality risks from other noncommunicable diseases such as cardiovascular disease (2). In addition, observations during the COVID-19 pandemic have revealed that people with diabetes have at least twice the risk of severe COVID-19 outcomes (3).

The risk of type 2 diabetes can be reduced by modifying unhealthy behaviors, such as a poor diet (4). According to the 2019 Global Burden of Diseases, 29.7% of disability-adjusted life-years of diabetes are attributable to dietary factors (5). In previous research focused on nutrients, associations were reported between poor diet quality and the incidence of type 2 diabetes and its comorbidities, particularly for diets high in saturated fat (6) and low in fiber (7). Evidence of the role of free sugars is less consistent (8); although high intakes have been associated with increased risk of type 2 diabetes, it is unclear whether the detrimental effects are caused by free sugars per se or by their contribution to excess energy intake. Increased free sugars intake is associated with reduced insulin sensitivity (9,10), but it has been reported that a high intake of free sugars without excess energy may not have any detrimental impact on health (11). There is also some evidence for positive associations between single foods and an increased risk of developing type 2 diabetes, such as higher intakes of unprocessed and processed meat (12,13), fruit juices, sugar-sweetened beverages (SSBs), refined grains, sweets, and desserts (8,14) and low intakes of fresh fruit and vegetables (15).

However, the impact of individual nutrients or foods on health outcomes may not adequately reflect the health effects of a whole diet in which foods are eaten in combination. Therefore, the observed associations could be driven by one or more nutrients, foods or other aspects of the diet that co-occur simultaneously. Reduced rank regression (RRR) is a data-driven approach that derives dietary patterns (DPs) by using a priori knowledge to select nutrients hypothesized to be on the causal pathway to examine how specific combinations of foods are associated with disease outcomes (16,17). This method allows specific investigations into certain diet-disease pathways (18), as opposed to the more exploratory approaches of principal components analysis or factor analysis (19). RRR has previously been used to study the effect of DPs on pathways of glucose homeostasis (20), inflammatory biomarkers (21), plasma circulating fatty acids (22), and blood lipids (23) as biomarkers of type 2 diabetes risk, although this evidence is not particularly strong, because of smaller sample sizes and low explained variance for the intermediate biomarkers. Here, we use detailed dietary data from the UK Biobank participants to identify DPs explaining high variability in known dietary risk factors, such as energy density, free sugars, saturated fat, and low fiber intakes and to assess the association between DPs and incident type 2 diabetes.

RESEARCH DESIGN AND METHODS

Study Population

UK Biobank is a population-based cohort comprising 502,536 participants (aged 37 to 73 years) from 22 sites across England, Wales, and Scotland, with baseline measures collected between 2006 and 2010 and data linked to hospital and mortality records. Extensive information on sociodemographics, health behavior, and medical history was obtained via a touch-screen questionnaire and face-to-face interviews. Physical measures (e.g., height, weight) and blood and urine samples were collected by trained personnel conforming to a standardized protocol. All participants provided written informed consent (24,25).

Measurement of Dietary Intake

A web-based 24 h dietary instrument (Oxford WebQ) was used to collect detailed dietary data. The Oxford WebQ was validated against an interviewer-administered 24 h recall questionnaire (26) and collected information on the quantities of up to 206 widely consumed food items and 32 types of drinks consumed over the previous day (27). Participants with valid e-mail addresses were invited to complete a dietary questionnaire at baseline and were followed up to four times between April 2009 and June 2012 (cycle 1: February 2011 to April 2011; cycle 2: June 2011 to September 2011; cycle 3: October 2011 to December 2011; cycle 4: April 2012 to June 2012). Participants with at least two assessments were retained for analysis to better reflect usual intakes (28), and their mean dietary intake was calculated (17).

Using methods described previously (17,27), food intake data were aggregated into 50 main groups aligned to the U.K. National Diet and Nutrition Survey and according to the similarity of their nutritional composition and culinary use. Total energy and nutrient intake data were automatically estimated by multiplying the number of portions consumed by the set quantity of each food portion size and its nutrient composition according to the UK Nutrient Databank food composition tables (2012–2013 and 2013–2014) (29). Energy density, saturated fatty acid (SFA), free sugars, and fiber density were selected because of their significant roles in the development of obesity and type 2 diabetes and their high frequency of intake in daily life (4,9,30). Energy density (kJ/g) was calculated by dividing total food energy (in kilojoules) by total food weight (grams); all beverages were excluded because of their disproportionate influence on total energy density value (31). The percent total energies of SFA and free sugars (% total energy) were calculated by dividing daily energy from saturated fat or free sugars by total daily energy intake. Fiber density (g/MJ) was calculated by absolute intake of fiber (grams per day) divided by total daily energy intake (in megajoules), then multiplying by 1,000. Free sugars were defined as such on the basis of the U.K. Scientific Advisory Committee on Nutrition definitions; fiber was calculated using the Englyst method (29).

Individual estimated energy requirements (EERs) were calculated by the Schofield equation from the 1985 Food and Alcohol Organization/World Health Organization/United Nations University Expert Consultation Report on Human Energy Requirements. Dietary misreporting was calculated by using the ratio of energy intake (EI) to EER and its 95% CI (32). Dietary underreporters (EI:EER $<$ 95% CI EI:EER) and overreporter (EI:EER $>$95% CI EI:EER) also were excluded.

Ascertainment of Outcomes

The outcomes for this study were primary or secondary diabetes events (excluding type 1 diabetes) ascertained from hospital episode statistics and death registry data linked to the UK Biobank. We defined incident diabetes events as a hospital admission or death with the

following *International Classification of Diseases*, 10th Revision, codes in the hospital or death records: E11–E14. Hospital admission data were available up to 31 January 2021 in England and Scotland and 28 February 2018 in Wales. Deaths were ascertained via linkage to the death certificates and were available starting from baseline up to 31 January 2021 in England, Wales, and Scotland. Therefore, we censored diabetes analysis at the date of first incident diabetes or death or 31 January 2021, whichever occurred first.

Statistical Analyses

Identification of DPs

RRR was used to identify DPs that explained the maximum variation in a set of nutrient-response variables hypothesized to be on the causal pathway between food groups and incident type 2 diabetes. Energy density, SFA, free sugars, and fiber density were all used as response variables in the RRR model. The number of extracted patterns is in accordance with the number of response variables in the model. Respondents were scored for each DP (*z* score) representing the degree to which their dietary intake reflected each DP relative to that of other respondents. Increasing intakes of foods with positive factor loadings increase the DP *z* score; increasing intakes of foods with negative factor loadings decrease the DP *z* score. A larger factor loading value indicates that the food group makes a greater contribution to the DP. The associations between DPs and nutrient-response variables were evaluated by correlation coefficients (Supplementary Table 2). DPs that individually explained >20% of variation in response variables were retained for subsequent analyses.

Prospective Association of DPs With Incident Type 2 Diabetes

We used multivariable Cox proportional hazard models stratified for sex and regions (England, Scotland, and Wales) with age (years) as a timescale variable to obtain hazard ratios (HRs) with 95% CIs per unit increase in DP *z* scores, with sequential adjustment for race or ethnicity (White race, non-White race, missing data), Townsend Index of Deprivation (quintiles 1–5, with higher scores representing greater deprivation), education (higher degree [college or university degree, or professional qualifications], any U.K. school degree [advanced levels, advanced subsidiary levels, ordinary levels, General Certificate of Secondary Education, Certificate of Secondary Education, or equivalent], vocational qualifications [National Vocational Qualification, Higher National Diploma, Higher National Certificate, or equivalent], other [none of the listed qualifications], missing data), smoking status (never, current, previous, missing data), physical activity (low [<600 metabolic equivalent (MET)-minutes/week], moderate [≥600 and <3,000 MET-minutes/week], high [≥3,000 MET-minutes/week], missing data), log-transformed energy intake, family history of diabetes (yes, no), hypertension (yes, no), cardiovascular disease (yes, no), high cholesterol (yes, no), menopause in women (yes, no, not applicable [i.e., men]) and BMI group (underweight [<18.5 kg/m^2], healthy weight [18.5 to <25 kg/m^2], overweight [25 to <30 kg/m^2], obese [≥30 kg/m^2], missing data). Detailed information on variables are provided in Supplementary Table 1. The proportional hazards assumption was assessed by Schoenfeld residuals. Trend tests were performed by including the median score of each pattern quintile as a continuous variable in the models; the lowest quintile was used as the referent. Restricted cubic splines models with the same covariate specification were computed with 5 knots to examine nonlinear associations between DP *z* scores and incident type 2 diabetes.

Likelihood ratio tests were used to examine the heterogeneity of the associations of the DPs with risk of incident type 2 diabetes by age group (<60, ≥60 years), sex (female, male), smoking status (never, previous, current), physical activity (low, moderate, high), BMI group (underweight/healthy [<25], overweight [25 to <30], obesity [≥30]).

Sensitivity and Exploratory Analyses

To investigate potential bias in relation to random variation in individual intakes, we repeated the RRR analysis to derive DPs among participants providing at least three (*n* = 75,003), at least four (*n* = 34,644), or five (*n* = 5,504) online dietary assessments, which may better reflect usual intakes. A second sensitivity analysis excluded participants who had a diabetes event within 2 years after completing their last 24 h online dietary assessment.

RRR analysis was performed using SAS statistical software (version 9.4; SAS Institute). The rest of the analyses were performed in Stata/MP 14.0.

RESULTS

Of the 502,536 UK Biobank participants recruited at baseline, participants were excluded for the following reasons: a diabetes event occurred before baseline (*n* = 26,845); pregnancy (*n* = 375); did not complete any validated dietary assessment (*n* = 273,434) or completed only one 24 h online dietary assessment (*n* = 79,995); diabetes occurred before the latest dietary questionnaire (*n* = 172); missing nutrient data (*n* = 21); implausible energy intake (overreporters: *n* = 130; underreporters: *n* = 996; or missing BMI data (*n* = 225) (Supplementary Fig. 1). In this analysis, 120,343 participants who provided complete data using the 24 h online dietary assessments on at least two occasions were included.

Four DPs were derived (Supplementary Table 2). DP1 (43%) and DP2 (20%) (Fig. 1) jointly explained 63% of variation in all response variables; DP3 and DP4 only contributed to 10% and 4%, respectively, which were retained in subsequent analyses. DP1 was characterized by high consumption of chocolate and confectionery, butter and other animal-fat spreads, and low-fiber bread, and low consumption of fresh fruit, vegetables, and high-fiber breakfast cereals (Fig. 1). DP2 was characterized by high consumption of SSBs, fruit juice, table sugars, and preserves, and low consumption of high-fat cheese and butter and other animal-fat spreads (Fig. 1). In sensitivity analyses, we derived new DPs from RRR using subsamples of people who provided three, four, or five dietary assessments or excluding those who experienced the event within 2 years of completing their latest assessment (Supplementary Figs. 2–5), yielding consistent results.

During 1,350,644 person-years of follow-up (11.2 years of median follow-up from baseline, 8.4 years of follow-up after the latest dietary assessment), 2,878 participants developed incident type 2 diabetes (including 83 fatal cases). Baseline characteristics including demographics, socioeconomic status, behavior risk factors, health history and conditions, dietary intake, main food groups across quintiles of DP scores are given in Table 1. Higher quintiles of DP1 *z* scores included people with the following characteristics: a higher proportion of men, younger age, higher

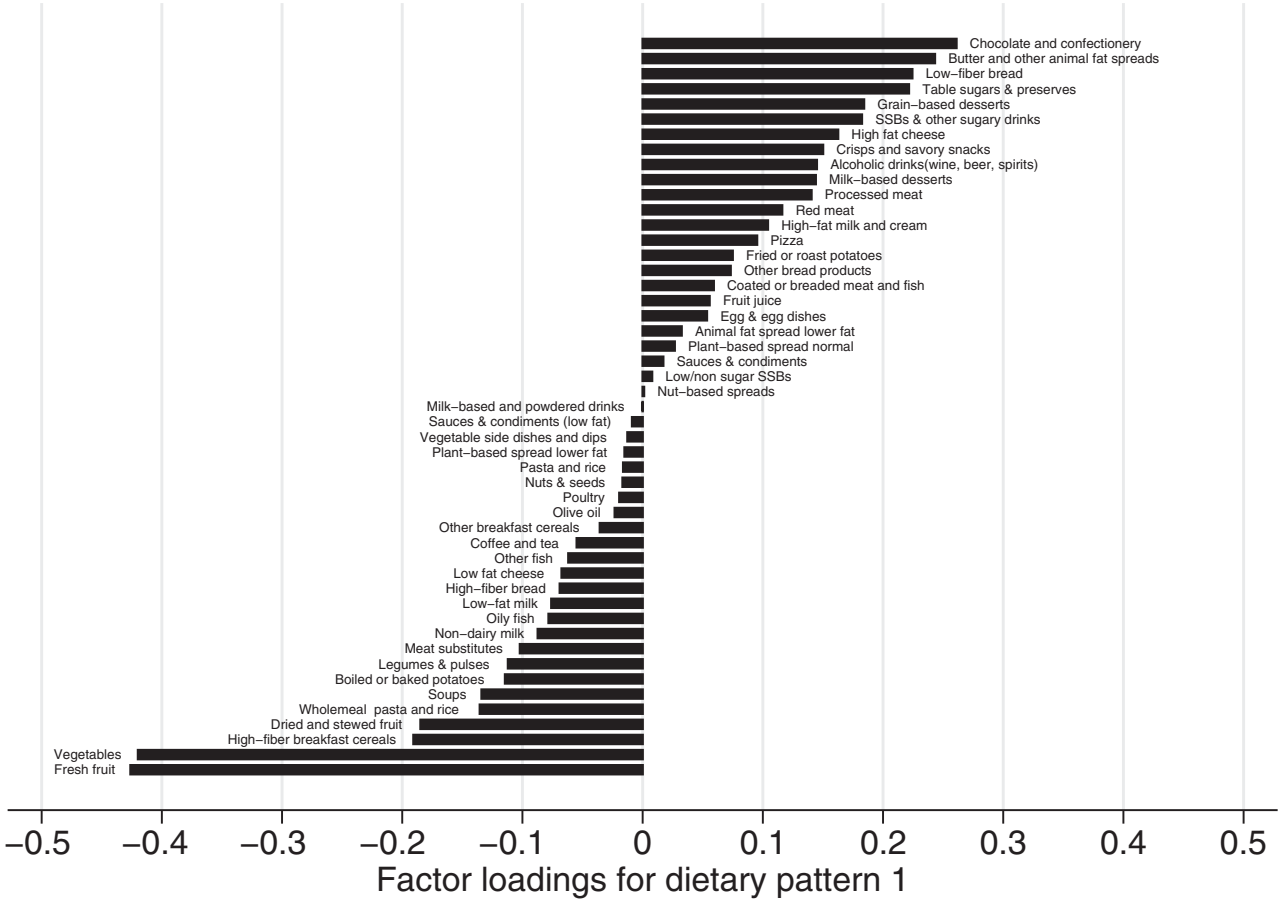

Figure 1—Factor loadings for food groups in each DP.

Townsend index, lower attained education level, current smokers, less physically active, and higher prevalence of obesity or hypertension. Higher quintiles of DP2 *z* scores included a higher proportion of younger participants, lower attained education level, never smokers, more physically active, higher prevalence of hypertension, cardiovascular disease, and high cholesterol.

Intakes of chocolate and confectionery, butter and other animal-fat spreads, table sugars and preserves, SSBs and other sugary drinks, low-fiber bread, processed meat, and high-fat cheese were higher across quintiles of DP1 *z* scores, along with lower intakes of high-fiber bread, vegetables, and fresh fruit. In DP2, intakes of chocolate and confectionery, added sugars and preserves, SSBs and other sugary drinks, fruit juice, and fresh fruit were higher across quintiles of DP2 *z* scores, with lower intakes of low-fiber bread, processed meat, high-fat cheese, and vegetables.

There was positive association between adherence to DP1 and incident type 2 diabetes after sequential adjustment for demographics, sociodemographics, behavior risk factors, and health history and conditions ($P_{trend} < 0.001$), and remained statistically significant after adjustment for BMI ($P_{trend} < 0.001$; Table 2, Supplementary Fig. 6). There was no evidence of an association between adherence to DP2 and the risk of type 2 diabetes. The exclusion of participants who had a diabetes event within 2 years of completing their last 24 h online dietary assessment showed that the associations between DPs and the risk of diabetes events were largely unchanged (Supplementary Table 4).

There was a significant interaction between both DPs and age, with significantly higher risks among younger people (<60 years; DP1: HR 1.13 [95% CI 1.09–1.18], DP2: HR 1.10 [95% CI 1.05–1.15]; Fig. 2). A significant interaction was also found for DP1 and BMI, with higher risks among people with obesity at baseline (HR 1.11; 95% CI 1.07–1.16). There was no evidence of any effect modification by sex, smoking status, or level of physical activity.

CONCLUSIONS

In this large cohort of British adults, we identified two main DPs. The main DP (DP1) explained almost half of the variance in nutrient intakes (43%) and was characterized by high intakes of chocolate and confectionery, butter and other animal-fat spreads, and low-fiber bread, and low intakes of fresh fruit, vegetables, and high-fiber breakfast cereals, contributing to excess energy intake, SFAs, free sugars, and low intake of fiber. There was a positive linear association between this DP and type 2 diabetes events, with stronger associations found among people aged <60 years and those who had obesity at baseline. A second DP (DP2), explaining much less of the variance in nutrient intakes (20%) and characterized by a very high

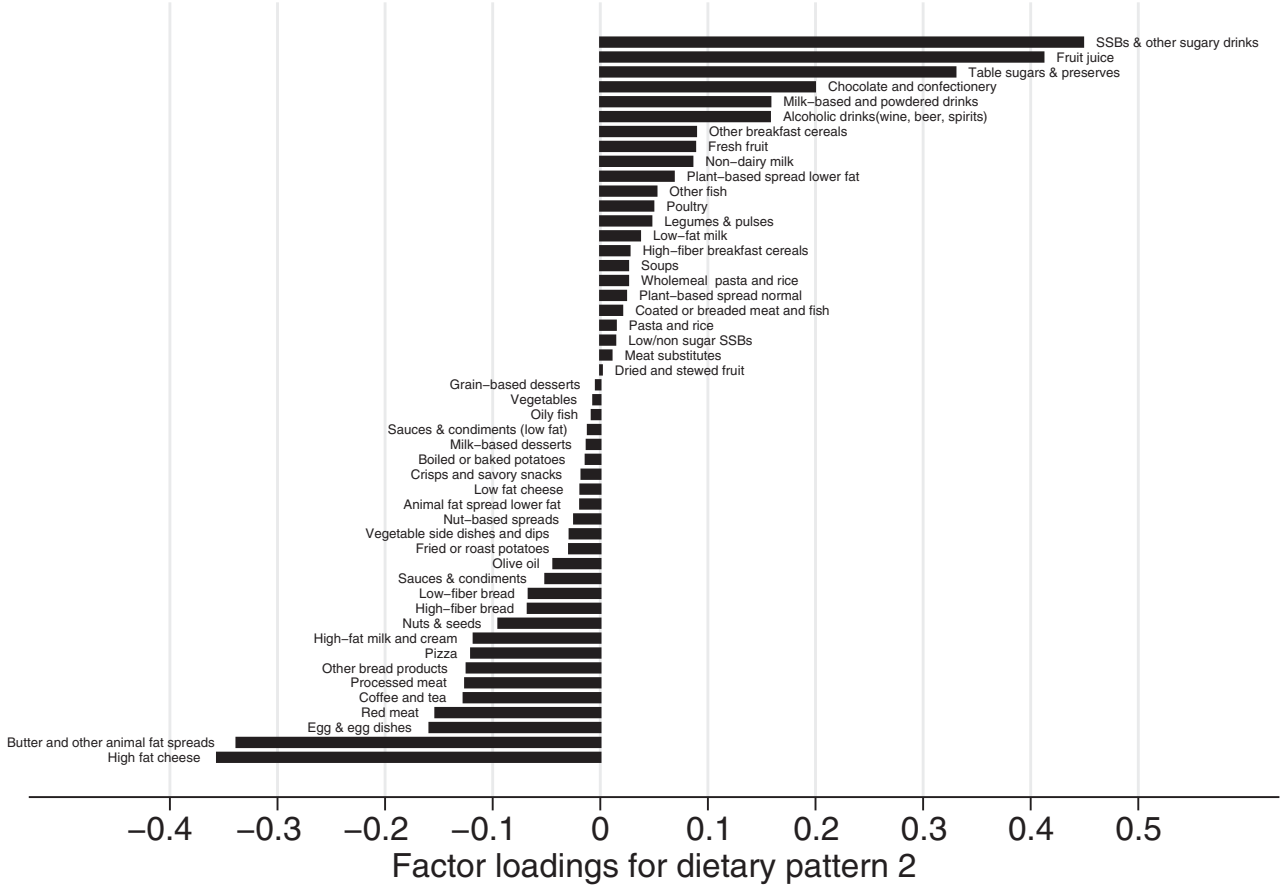

Figure 1—Continued.

intake of free sugars from higher intakes of SSBs, fruit juice, table sugars, and preserves, together with lower intakes of high-fat cheese and butter and other animal-fat spreads, was not associated with the incidence of type 2 diabetes.

This study contributes new observational evidence to help understand the associations of a whole diet with incident type 2 diabetes. In practice, nutrients are not consumed separately and they may exert effects on health in a synergistic way. Analyses of DPs through RRR, which use nutrients of concern as response variables in the causal pathway, may help detect stronger associations due to the cumulative effects of different dietary components. The main DP derived in this study is characterized by food groups (and underlying nutrients) for which associations with type 2 diabetes are consistent with findings from previous studies (19,33). For example, excessive consumption of chocolate and confectionery (34), red and processed meat (35), added sugars and preserves (36), and low consumption of vegetables, fresh fruit (37), and high-fiber bread (38) have been associated with adverse metabolic effects and higher risk of diabetes. In a recent systematic literature review and meta-analysis of 16 cohorts, authors found three patterns derived by RRR showing refined grains, SSBs, and processed meat were all significantly associated with diabetes risk independent of the biomarkers used as response variables and study populations (39). A recent clinical trial in overweight men provided evidence that compared with a diet enriched in free sugars, a diet enriched in SFAs led to increased intrahepatic triacylglycerol levels and postprandial glycemia, but a diet enriched in free sugars did not influence intrahepatic triacylglycerol levels and led to only minor metabolic changes (40).

The second DP was characterized by foods that contributed high intake of free sugars (e.g., SSBs, fruit juice) but in the context of low SFA levels and energy density and adequate amounts of fiber due to high intake of fruits and vegetables. People in the highest quintile of this pattern consumed, on average, 17.4% of dietary energy from free sugars (more than three times the U.K. dietary guideline) but with only 10% from SFAs (the recommended maximum level) with no significant increase in energy intake. Here there was no evidence of an overall association with incidence of diabetes. Accordingly, this analysis, which controls for differences in BMI, supports the hypothesis that free sugars increase the risk for type 2 diabetes because of their contribution to excess energy intake rather than having an direct, independent effect (8,41). In short- or medium-term isoenergetic intervention trials, free sugars showed no effect on body weight or blood pressure and the effects on blood lipid profile were not conclusive (11). Our findings are also consistent with the evidence that diets low in saturated fat reduce the risk of type 2 diabetes (42).

Table 1—Baseline characteristics of the study population and average dietary intake

	Total (N = 120,343)	DP1 quintiles				DP2 quintiles			
		1 (n = 24,069)	3 (n = 24,068)	5 (n = 24,068)	P value*	1 (n = 24,069)	3 (n = 24,068)	5 (n = 24,068)	P value*
Demographics									
Female sex (%)	56.5	70.5	59.0	35.9	<0.001	56.2	60.8	47.5	<0.001
Age (years)†	56.1 (7.8)	57.4 (7.3)	56.3 (7.7)	54.1 (8.1)	<0.001	56.0 (7.7)	56.3 (7.8)	55.5 (8.1)	<0.001
White race (%)	96.8	96.6	96.9	96.7	0.089	97.7	97.2	94.7	<0.001
Socioeconomic status: Townsend index quintile				<0.001					<0.001
1	20.0	20.3	20.6	18.4		19.3	20.5	19.2	
3	20.0	19.8	20.7	19.9		19.8	20.4	19.8	
5	19.9	19.2	18.6	22.5		20.9	18.6	21.4	
Missing data	0.1	0.1	0.1	0.1		0.1	0.1	0.1	
Education group (%)									
Vocational qualification (NVQ, HND, HNC, or equivalent)	12.4	10.4	12.4	16.1		11.7	11.9	14.0	
Any school degree (A level, AS level, O level, GCSE, CSE, or equivalent)	29.2	27.0	28.9	31.6		28.8	29.9	28.8	
Higher degree (college, university, or professional degree or qualification)	51.8	56.5	52.5	44.6		53.6	51.6	50.0	
None of the preceding groups	6.3	5.8	5.9	7.3	<0.001	5.6	6.4	6.7	<0.001
Missing data	0.3	0.3	0.3	0.3		0.2	0.3	0.3	
Behavior risk factors									
Smoking status (%)					<0.001				<0.001
Never	57.6	60.3	59.3	52.2		54.3	58.3	59.4	
Previous	35.3	35.7	35.0	34.9		37.0	35.5	33.0	
Current	6.9	3.8	5.5	12.8		8.5	6.0	7.3	
Missing	0.2	0.2	0.2	0.2		0.2	0.2	0.2	
Physical activity (IPAQ) (%)					<0.001				<0.001
Low	19.1	14.0	19.4	23.8		19.4	19.6	18.1	
Moderate	45.0	43.0	45.9	44.2		45.8	45.4	43.4	
High	35.8	43.0	34.6	32.0		34.8	34.9	38.5	
Missing data	<0.1	<0.1	<0.1	<0.1		<0.1	<0.1	<0.1	
Health history and conditions									
Family history of diabetes (%)	15.7	15.4	15.7	16.1	0.32	15.5	15.9	16.0	0.45
Hypertension (%)	23.3	22.8	23.1	24.3	<0.001	22.3	23.1	24.6	<0.001
Cardiovascular disease (%)	5.2	5.3	5.3	5.2	0.86	4.4	5.0	6.1	<0.001
High cholesterol (%)	13.5	13.3	13.6	13.3	0.76	11.4	13.7	15.2	<0.001
Menopause in women (%)	23.2	24.1	24.3	19.4	<0.001	22.8	24.7	20.8	<0.001
BMI (kg/m^2)	26.5(4.4)	26.0(4.4)	26.4(4.3)	27.3(4.6)	<0.001	26.7(4.7)	26.5(4.4)	26.5(4.4)	<0.001

Continued on p. 1321

Table 1—Continued

	Total (*N* = 120,343)	DP1 quintiles				DP2 quintiles			
		1 (*n* = 24,069)	3 (*n* = 24,068)	5 (*n* = 24,068)	*P* value*	1 (*n* = 24,069)	3 (*n* = 24,068)	5 (*n* = 24,068)	*P* value*
BMI group (%)					<0.001				<0.001
Underweight (<18.5)	0.6	0.9	0.5	0.5		0.7	0.5	0.5	
Healthy weight (18.5 to <25)	40.1	45.4	41.3	32.1		39.3	40.6	39.1	
Overweight (25 to <30)	41.1	38.1	41.2	44.0		39.9	41.0	42.7	
Obese (≥30)	18.2	15.5	17.0	23.3		20.0	17.8	17.6	
Missing data	<0.1	0.1	0.0	<0.1		<0.1	<0.1	<0.1	
Dietary intake									
Energy intake (MJ/day)	8.7 (2.2)	8.2 (2.1)	8.3 (2.0)	10.0 (2.5)	<0.001	9.3 (2.4)	8.3 (2.1)	9.0 (2.3)	<0.001
Dietary energy density (kJ/g)	6.5 (1.6)	4.8 (0.8)	6.4 (0.8)	8.4 (1.4)	<0.001	7.1 (1.6)	6.3 (1.5)	6.4 (1.6)	<0.001
Saturated fat (% E)	11.7 (3.2)	9.7 (2.6)	11.8 (2.8)	13.4 (3.3)	<0.001	14.4 (2.9)	11.3 (2.7)	10.0 (2.8)	<0.001
Free sugars (% E)	11.5 (5.2)	8.9 (4.1)	11.4 (4.5)	14.7 (6.1)	<0.001	7.7 (3.3)	10.6 (3.6)	17.4 (5.3)	<0.001
Fiber (g/day)	18.1 (6.2)	23.3 (6.7)	17.1 (5.0)	15.1 (5.3)	<0.001	18.2 (6.2)	17.8 (6.0)	18.5 (6.8)	<0.001
Fiber density (g/MJ)	2.1 (0.6)	2.9 (0.6)	2.1 (0.4)	1.5 (0.4)	<0.001	2.0 (0.6)	2.2 (0.7)	2.1 (0.7)	<0.001
Main food groups (g/day)									
Chocolate and confectionery	11.9 (21.4)	6.0 (12.4)	9.7 (16.0)	23.2 (33.5)	<0.001	8.3 (16.0)	10.4 (17.9)	18.9 (30.8)	<0.001
Butter and other normal animal-fat spreads	5.1 (8.7)	1.9 (4.8)	4.3 (7.2)	10.5 (12.4)	<0.001	12.1 (12.6)	3.5 (6.4)	2.0 (5.0)	<0.001
Added sugars and preserves	8.9 (13.0)	5.4 (8.4)	7.6 (10.2)	15.7 (19.4)	<0.001	6.0 (9.4)	7.4 (10.4)	15.1 (18.6)	<0.001
SSBs and other sugary drinks	87.3 (166.1)	45.1 (107.4)	74.1 (134.8)	163.1 (245.5)	<0.001	30.8 (77.6)	58.4 (105.1)	219.5 (267.7)	<0.001
Fruit juice	110.0 (137.6)	104.8 (139.9)	114.0 (135.4)	108.1 (144.3)	<0.001	52.7 (87.6)	92.7 (108.0)	205.7 (188.9)	<0.001
Low-fiber bread	28.1 (42.3)	9.2 (22.1)	22.2 (32.2)	61.0 (58.7)	<0.001	36.6 (50.4)	25.3 (38.2)	25.9 (41.0)	<0.001
Processed meat	18.2 (28.0)	10.3 (19.7)	16.8 (24.5)	29.4 (37.5)	<0.001	25.9 (35.6)	16.8 (25.1)	14.0 (24.3)	<0.001
High-fat cheese	14.8 (18.1)	10.9 (15.4)	14.4 (16.7)	19.4 (22.3)	<0.001	28.0 (24.3)	12.1 (14.1)	7.9 (12.4)	<0.001
High-fiber bread	49.0 (47.4)	53.8 (50.0)	51.2 (45.5)	39.3 (47.1)	<0.001	55.8 (52.6)	47.2 (44.2)	46.5 (47.7)	<0.001
Vegetables	191.0 (146.0)	325.6 (185.2)	170.7 (105.8)	107.1 (92.6)	<0.001	196.0 (147.7)	191.3 (140.6)	184.2 (154.9)	<0.001
Fresh fruit	197.0 (153.3)	342.7 (181.7)	177.9 (111.6)	97.4 (97.4)	<0.001	169.9 (142.6)	200.8 (148.5)	216.3 (170.0)	<0.001

A level, advanced level; AS, advanced subsidiary; CSE, Certificate of Secondary Education; E, energy; GCSE, General Certificate of Secondary Education; HNC, Higher National Certificate; HND, Higher National Diploma; IPAQ, International Physical Activity Questionnaire; NVQ, National Vocational Qualification; O, ordinary. *ANOVA or χ^2 test where appropriate.

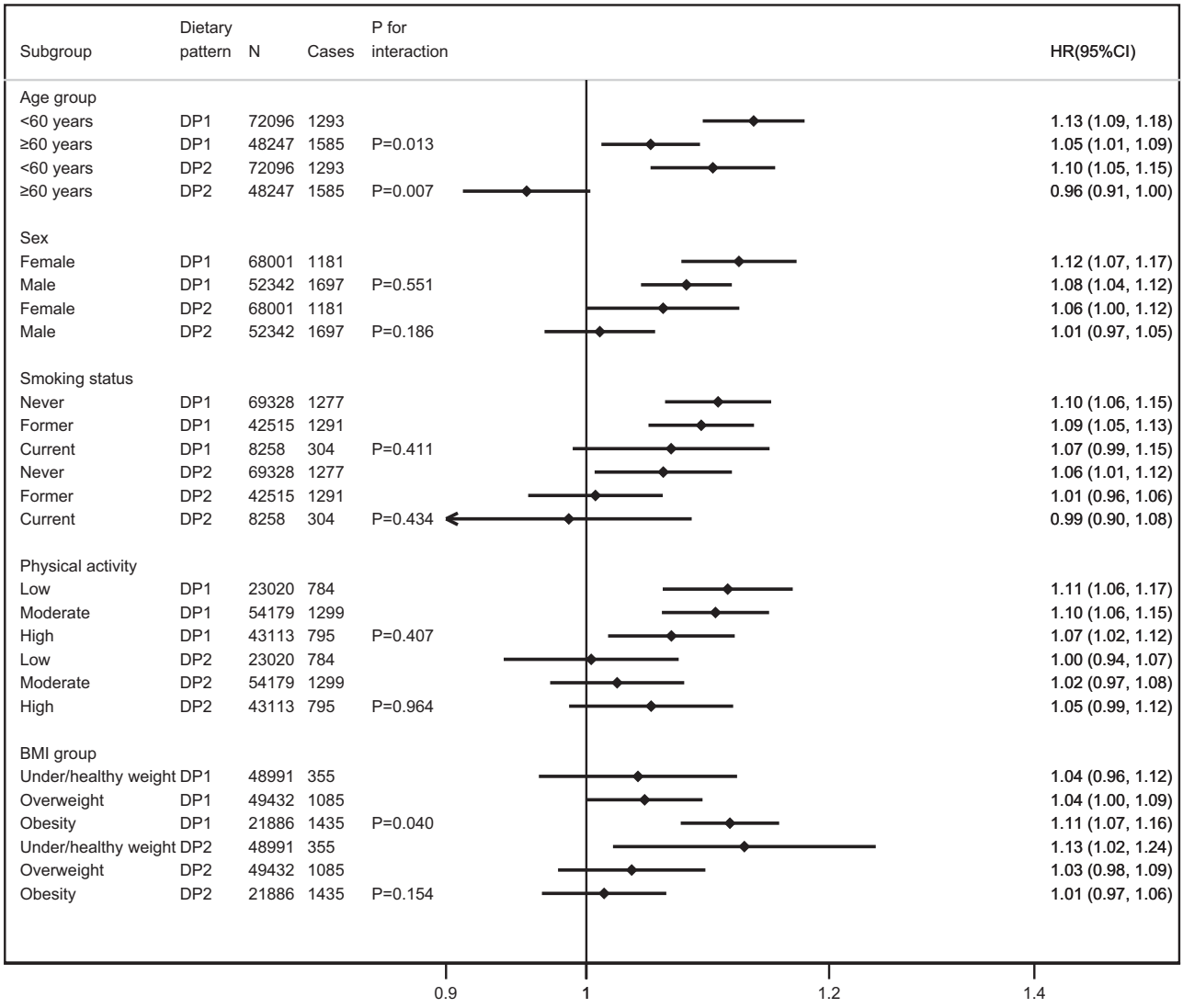

Figure 2—The association between DPs and risk of diabetes by age, sex, smoking status, and BMI group (*N* = 120,343).

We found that the positive associations between DP1 and risk of type 2 diabetes were slightly attenuated but remained statistically significant after adjusting for BMI, suggesting that BMI may only partly explain the variability in type 2 diabetes risk. However, in subsequent stratified analyses, DP1 was much more strongly associated with type 2 diabetes among people with obesity at baseline. Therefore, these results suggest that residual confounding or mediation by adiposity cannot be ruled out and might account for some of the observed associations.

Adherence to both DP1 and DP2 was a particular risk factor for the development of type 2 diabetes in younger people (<60 years), suggesting that when all other things are held constant, a poor diet could increase the risk of type 2 diabetes in younger people. Also, there are also age-related differences in healthy food consumption such as fruits and vegetable, with older adults consuming more fruit and vegetables than do younger adults—a pattern that has been observed in several U.K. national surveys. In the National Diet and Nutrition Survey (2016–2017 to 2018–2019), adults (aged 19–64 years) consumed, on average, 4.3 portions of fruit and vegetables per day and older adults (≥65 years) consumed 4.5 portions per day (43). Reports in the wider literature also have suggested that older people tend to consume more fruits and vegetables and other healthy foods (44), which may be the case among the UK Biobank participants. In our middle-aged population, 55% of type 2 diabetes events were observed in people aged ≥60 years. Our study did not show differential associations by sex, although a previous study using the RRR approach found that a high-fat, high glycemic index, low-fiber DP was associated with increased type 2 diabetes risk in middle-aged British women but not men (45), though the mechanistic basis for such differences is unclear. Also, differential associations across physical activity groups were not observed. DP1 was consistently associated with higher risk of type 2 diabetes regardless of the amount of physical activity reported; however, this association was slightly less strong in people with a high level of physical activity.

Strengths of this study include a large sample size and the empirical hypothesis-

Table 2—Sequentially adjusted HRs of total incident diabetes associated with each *z* score increase in DP

	No. of events	Adjusted for sex, age, and region	Plus sociodemographics*	Plus behavior risk factors†	Plus health history and conditions‡	Plus BMI group
DP1						
Total	2,878	1.16 (1.13–1.19)	1.14 (1.11–1.18)	1.11 (1.08–1.14)	1.11 (1.08–1.14)	1.09 (1.06–1.12)
Q1	452	1.00	1.00	1.00	1.00	1.00
Q2	513	1.15 (1.05–1.25)	1.15 (1.05–1.26)	1.11 (1.02–1.21)	1.10 (1.01–1.20)	1.13 (1.04–1.23)
Q3	549	1.23 (1.13–1.33)	1.22 (1.12–1.33)	1.16 (1.06–1.26)	1.15 (1.06–1.25)	1.19 (1.10–1.30)
Q4	621	1.41 (1.31–1.53)	1.39 (1.29–1.51)	1.28 (1.18–1.38)	1.27 (1.17–1.37)	1.25 (1.16–1.35)
Q5	743	1.73 (1.60–1.86)	1.65 (1.53–1.78)	1.46 (1.35–1.59)	1.45 (1.34–1.57)	1.38 (1.27–1.49)
P_{trend}		<0.001	<0.001	<0.001	<0.001	<0.001
DP2						
Total	2,878	1.05 (1.01–1.08)	1.03 (0.99–1.06)	1.03 (1.00–1.07)	1.02 (0.98–1.05)	1.03 (0.99–1.06)
Q1	610	1.00	1.00	1.00	1.00	1.00
Q2	545	0.98 (0.90–1.06)	0.97 (0.89–1.06)	0.96 (0.89–1.05)	0.96 (0.88–1.04)	0.97 (0.89–1.05)
Q3	567	1.02 (0.94–1.11)	1.02 (0.94–1.10)	1.01 (0.93–1.10)	0.99 (0.91–1.07)	1.01 (0.93–1.10)
Q4	524	0.93 (0.85–1.01)	0.91 (0.83–0.99)	0.91 (0.84–1.00)	0.88 (0.81–0.96)	0.90 (0.82–0.98)
Q5	632	1.09 (1.01–1.18)	1.04 (0.96–1.13)	1.06 (0.98–1.14)	1.01 (0.93–1.09)	1.04 (0.96–1.12)
P_{trend}		0.234	0.735	0.476	0.840	0.818

Adjusted HRs and 95% CIs of total DP scores obtained using Cox proportional hazard regression. Age at risk adjusted for by using age during the study as the underlying timescale for Cox regression. All the models were stratified by sex and regions (England, Scotland, and Wales). Q, quintile. *Sociodemographic characteristics: Race or ethnicity (White, others, missing), Townsend Index of Deprivation (Q1–Q5, with higher scores representing greater deprivation), education (higher degree [college or university degree, or professional qualifications], any school degree [advanced levels, advanced subsidiary levels, ordinary levels, General Certificate of Secondary Education, Certificate of Secondary Education], vocational qualifications [National Vocational Qualification, Higher National Diploma, Higher National Certificate], other [none of the listed qualifications], missing data). †Behavior risk factors: smoking status (never, current, previous, missing data), physical activity level (low [<600 metabolic equivalent (MET)-minutes/week], moderate [≥600 and <3,000 MET-minutes/week], high [≥3,000 MET-minutes/week], missing data), log-transformed energy intake. ‡Health history and conditions: family history of diabetes, menopause in women, hypertension, cardiovascular disease, high cholesterol. Full model was stratified by sex and regions (England, Scotland, and Wales) and adjusted for demographics, socioeconomic status, behavior risk factors, and health history and conditions with BMI group included.

based DP analysis that was used to identify DPs in this cohort of middle-aged British adults and their association with type 2 diabetes. Some limitations of this study should be acknowledged. First, DPs were identified and validated from at least two 24 h, online, self-reported dietary assessments, which may be affected by recall bias or misreporting (e.g., people at higher risk of diabetes or aware of their diabetes biomarkers may report foods differentially), and DPs were derived using data from people with higher willingness to report their diet (which may be affected by selection bias). However, our results were reassured by the sensitivity analyses, which used at least 2, 3, 4, or 5 repeated assessments to better capture usual intakes, showing consistent results. Second, some uncommon foods and other important covariates were not captured by the questionnaires, hence some residual confounding may remain. Third, although the identified DPs are specific to the U.K. population, assuming that the relationships between dietary factors and health outcomes would be similar among individuals, outcome-dependent methods, such as RRR, are expected to be relatively reproducible across different study populations (46,47). Fourth, incident diabetes cases were ascertained based on hospital records, but this may not fully capture all cases of type 2 diabetes, because many of these are reported in primary care. Finally, as usually happens with the RRR approach, the emerging DPs did not explain all of the variability in the nutrient-response variables that were included in the RRR model, and any remaining variability may be explained by other nutrients that may also be involved in the disease pathway. In exploratory mediation analyses (Supplementary Table 4), we investigated the role of the nutrient-response variables in the association between DPs and the risk of diabetes. After adjustment for the nutrients with the highest correlations with each DP, the associations for DP1 and DP2 were largely unchanged, which means that the derived DPs were associated with diabetes independent of those nutrient-response variables. This can be partly explained by the fact that both the DP and the nutrient-response variables may explain the same variability. But this also suggests that there might be other nutrients in the causal pathway explaining the observed risk with type 2 diabetes, for which future

research is warranted. Residual confounding due to unmeasured variables (e.g., blood glucose (48), impaired fasting glycemia (49,50) is another limitation of observational research that can also explain the observed results.

Current dietary guidelines in the U.K. and many other countries are still based predominately on nutrient recommendations. Although this reflects the underlying evidence base, especially experimental and mechanistic research, it does not reflect the way people eat. Food-based dietary guidelines may help accelerate behavior change compared with nutrient-based recommendations by providing targeted and simpler advice on foods, which make it more, or less, likely to achieve an overall healthy diet (51). Moreover, this may help reduce the risk of conflicting messages regarding the relative importance of one or another nutrient, particularly saturated fat and free sugars, and recognize that there are many foods that are important sources of both.

In conclusion, this large population-based cohort study has shown clear evidence that diets high in chocolate and confectionery, butter and other animal-fat spreads, and low-fiber bread, and low intakes of fresh fruit, vegetables, and high-fiber breakfast cereals are associated with a higher risk of developing type 2 diabetes in this cohort of middle-aged British adults. The effects of a poor diet are especially pronounced in younger people and those living with obesity.

Acknowledgments. This research was conducted using the UK Biobank resource under application number 14990.

Funding. C.P. received a British Nutrition Foundation pump-priming award that paid for the access to the data. S.A.J. and C.P. are funded by the Oxford and Thames Valley National Institute of Health Research (NIHR) Applied Research Centre. S.A.J. and P.A. are NIHR senior investigators. S.A.J. and J.C. are funded by NIHR Oxford Biomedical Research Centre. J.C. is supported by core grants to Clinical Trial Service Unit from the Medical Research Council and the British Heart Foundation (grant CH/1996001/9454).

The funders had no role in designing the study, the analysis, or the decision to submit the paper. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. M.G. and C.P. conceived and designed the research question and wrote the first draft of the manuscript. M.G. analyzed the data and G.L.A. contributed to dietary pattern analysis. G.L.A., A.P.C., K.P., J.C., and C.P. contributed to the data analysis. All authors contributed to interpretation of the data, revised the manuscript critically for important intellectual content, and read and approved the final manuscript. M.G. and C.P. are guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Chan JCN, Lim L-L, Wareham NJ, et al. The Lancet Commission on diabetes: using data to transform diabetes care and patient lives. Lancet 2021;396:2019–2082
2. Shah AD, Langenberg C, Rapsomaniki E, et al. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1·9 million people. Lancet Diabetes Endocrinol 2015;3:105–113
3. Barron E, Bakhai C, Kar P, et al. Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study. Lancet Diabetes Endocrinol 2020;8:813–822
4. Neuenschwander M, Ballon A, Weber KS, et al. Role of diet in type 2 diabetes incidence: umbrella review of meta-analyses of prospective observational studies. BMJ 2019;366:l2368
5. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020;396:1223–1249
6. van Dam RM, Willett WC, Rimm EB, Stampfer MJ, Hu FB. Dietary fat and meat intake in relation to risk of type 2 diabetes in men. Diabetes Care 2002;25:417–424
7. Reynolds AN, Akerman AP, Mann J. Dietary fibre and whole grains in diabetes management: systematic review and meta-analyses. PLoS Med 2020;17:e1003053
8. Choo VL, Viguiliouk E, Blanco Mejia S, et al. Food sources of fructose-containing sugars and glycaemic control: systematic review and meta-analysis of controlled intervention studies. BMJ 2018;363:k4644
9. Macdonald IA. A review of recent evidence relating to sugars, insulin resistance and diabetes. Eur J Nutr 2016;55(suppl 2):17–23
10. Johnson RK, Lichtenstein AH, Anderson CAM, et al.; American Heart Association Nutrition Committee of the Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; Council on Quality of Care and Outcomes Research; and Stroke Council. Low-calorie sweetened beverages and cardiometabolic health: a science advisory from the American Heart Association. Circulation 2018;138:e126–e140
11. Fattore E, Botta F, Agostoni C, Bosetti C. Effects of free sugars on blood pressure and lipids: a systematic review and meta-analysis of nutritional isoenergetic intervention trials. Am J Clin Nutr 2017;105:42–56
12. Sluik D, Brouwer-Brolsma EM, Berendsen AAM, et al. Protein intake and the incidence of pre-diabetes and diabetes in 4 population-based studies: the PREVIEW project. Am J Clin Nutr 2019;109:1310–1318
13. Chen Z, Franco OH, Lamballais S, et al. Associations of specific dietary protein with longitudinal insulin resistance, prediabetes and type 2 diabetes: the Rotterdam Study. Clin Nutr 2020;39:242–249
14. Imamura F, O'Connor L, Ye Z, et al. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction. BMJ 2015;351:h3576
15. Zheng J-S, Sharp SJ, Imamura F, et al. Association of plasma biomarkers of fruit and vegetable intake with incident type 2 diabetes: EPIC-InterAct case-cohort study in eight European countries. BMJ 2020;370:m2194
16. Afshin A, Sur PJ, Fay KA, et al.; 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 2019;393:1958–1972
17. Gao M, Jebb SA, Aveyard P, et al. Associations between dietary patterns and the incidence of total and fatal cardiovascular disease and all-cause mortality in 116,806 individuals from the UK Biobank: a prospective cohort study. BMC Med 2021;19:83
18. Weikert C, Schulze MB. Evaluating dietary patterns: the role of reduced rank regression. Curr Opin Clin Nutr Metab Care 2016;19:341–346
19. Forouhi NG, Misra A, Mohan V, Taylor R, Yancy W. Dietary and nutritional approaches for prevention and management of type 2 diabetes. BMJ 2018;361:k2234
20. McNaughton SA, Mishra GD, Brunner EJ. Dietary patterns, insulin resistance, and incidence of type 2 diabetes in the Whitehall II Study. Diabetes Care 2008;31:1343–1348
21. Liese AD, Weis KE, Schulz M, Tooze JA. Food intake patterns associated with incident type 2 diabetes: the Insulin Resistance Atherosclerosis Study. Diabetes Care 2009;32:263–268
22. Seah JYH, Ong CN, Koh W-P, Yuan J-M, van Dam RM. A dietary pattern derived from reduced rank regression and fatty acid biomarkers is associated with lower risk of type 2 diabetes and coronary artery disease in Chinese adults. J Nutr 2019;149:2001–2010
23. Duan M-J, Dekker LH, Carrero J-J, Navis G. Blood lipids-related dietary patterns derived from reduced rank regression are associated with incident type 2 diabetes. Clin Nutr 2021;40:4712–4719
24. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779
25. Collins R. What makes UK Biobank special? Lancet 2012;379:1173–1174
26. Liu B, Young H, Crowe FL, et al. Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. Public Health Nutr 2011;14:1998–2005
27. Piernas C, Perez-Cornago A, Gao M, et al. Describing a new food group classification system for UK biobank: analysis of food groups and

sources of macro- and micronutrients in 208,200 participants. Eur J Nutr 2021;60:2879–2890
28. Carter JL, Lewington S, Piernas C, et al. Reproducibility of dietary intakes of macronutrients, specific food groups, and dietary patterns in 211 050 adults in the UK Biobank study. J Nutr Sci 2019;8:e34
29. Perez-Cornago A, Pollard Z, Young H, et al. Description of the updated nutrition calculation of the Oxford WebQ questionnaire and comparison with the previous version among 207,144 participants in UK Biobank. Eur J Nutr 2021;60:4019–4030
30. Thom G, Lean M. Is there an optimal diet for weight management and metabolic health? Gastroenterology 2017;152:1739–1751
31. Johnson L, Mander AP, Jones LR, Emmett PM, Jebb SA. A prospective analysis of dietary energy density at age 5 and 7 years and fatness at 9 years among UK children. Int J Obes 2008;32:586–593
32. Schofield WN, Schofield C, James WPT. Basal metabolic rate: review and prediction, together with an annotated bibliography of source material. Hum Nutr 1985;39:5–41
33. Magliano DJ, Islam RM, Barr ELM, et al. Trends in incidence of total or type 2 diabetes: systematic review. BMJ 2019;366:l5003
34. Curtis PJ, Sampson M, Potter J, Dhatariya K, Kroon PA, Cassidy A. Chronic ingestion of flavan-3-ols and isoflavones improves insulin sensitivity and lipoprotein status and attenuates estimated 10-year CVD risk in medicated postmenopausal women with type 2 diabetes: a 1-year, double-blind, randomized, controlled trial. Diabetes Care 2012;35:226–232
35. de Koning L, Fung TT, Liao X, et al. Low-carbohydrate diet scores and risk of type 2 diabetes in men. Am J Clin Nutr 2011;93:844–850
36. Kmietowicz Z. Type 2 diabetes: sweetened drinks pose greater risk than other sugary foods. BMJ 2018;363:k4943
37. Bazzano LA, Li TY, Joshipura KJ, Hu FB. Intake of fruit, vegetables, and fruit juices and risk of diabetes in women. Diabetes Care 2008;31:1311–1317
38. Schulze MB, Schulz M, Heidemann C, Schienkiewitz A, Hoffmann K, Boeing H. Fiber and magnesium intake and incidence of type 2 diabetes: a prospective study and meta-analysis. Arch Intern Med 2007;167:956–965
39. Jannasch F, Kröger J, Schulze MB. Dietary patterns and type 2 diabetes: a systematic literature review and meta-analysis of prospective studies. J Nutr 2017;147:1174–1182
40. Parry SA, Rosqvist F, Mozes FE, et al. Intrahepatic fat and postprandial glycemia increase after consumption of a diet enriched in saturated fat compared with free sugars. Diabetes Care 2020;43:1134–1141
41. Te Morenga L, Mallard S, Mann J. Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies. BMJ 2012;346:e7492
42. Morio B, Fardet A, Legrand P, Lecerf JM. Involvement of dietary saturated fats, from all sources or of dairy origin only, in insulin resistance and type 2 diabetes. Nutr Rev 2016;74:33–47
43. Bates B, Collins D, Jones K, et al. National Diet and Nutrition Survey Rolling Programme Years 9 to 11 (2016/2017 to 2018/2019). 2020. Available from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943114/NDNS_UK_Y9-11_report.pdf
44. Reininger B, Lee M, Jennings R, Evans A, Vidoni M. Healthy eating patterns associated with acculturation, sex and BMI among Mexican Americans. Public Health Nutr 2017;20:1267–1278
45. Pastorino S, Richards M, Pierce M, Ambrosini GL. A high-fat, high-glycaemic index, low-fibre dietary pattern is prospectively associated with type 2 diabetes in a British birth cohort. Br J Nutr 2016;115:1632–1642
46. Huybrechts I, Lioret S, Mouratidou T, et al. Using reduced rank regression methods to identify dietary patterns associated with obesity: a cross-country study among European and Australian adolescents. Br J Nutr 2017;117:295–305
47. Krebs-Smith SM, Subar AF, Reedy J. Examining dietary patterns in relation to chronic disease: matching measures and methods to questions of interest. Circulation 2015;132:790–793
48. Russell WR, Baka A, Björck I, et al. Impact of diet composition on blood glucose regulation. Crit Rev Food Sci Nutr 2016;56:541–590
49. de Hoogh IM, Oosterman JE, Otten W, et al. The effect of a lifestyle intervention on type 2 diabetes pathophysiology and remission: the Stevenshof Pilot Study. Nutrients 2021;13:2913
50. Kim MK, Han K, Koh ES, et al. Cumulative exposure to impaired fasting glucose and future risk of type 2 diabetes mellitus. Diabetes Res Clin Pract 2021;175:108799
51. Van Horn L, Carson JAS, Appel LJ, et al.; American Heart Association Nutrition Committee of the Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular Disease in the Young; Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; and Stroke Council. Recommended dietary pattern to achieve adherence to the American Heart Association/American College of Cardiology (AHA/ACC) guidelines: a scientific statement from the American Heart Association. Circulation 2016;134:e505–e529