Protein intake and the incidence of pre-diabetes and diabetes in 4 population-based studies: the PREVIEW project

Diewertje Sluik,[1] Elske M Brouwer-Brolsma,[1] Agnes AM Berendsen,[1] Vera Mikkilä,[2,3] Sally D Poppitt,[4] Marta P Silvestre,[4] Angelo Tremblay, Louis Pérusse,[5,6] Claude Bouchard,[7] Anne Raben,[8] and Edith JM Feskens[1]

[1]Division of Human Nutrition and Health, Wageningen University and Research, Wageningen, The Netherlands; [2]Division of Nutrition, Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland; [3]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland; [4]Human Nutrition Unit, Department of Medicine and School of Biological Sciences, University of Auckland, Auckland, New Zealand; [5]Department of Kinesiology; [6]Institute of Nutrition and Functional Foods, Université Laval, Québec City, Canada; [7]Human Genomics Laboratory, Pennington Biomedical Research Center, Baton Rouge, LA; and [8]Department of Nutrition, Exercise and Sports, Faculty of Science, University of Copenhagen, Denmark

ABSTRACT

Background: Data on the relationship between protein intake and the risk of type 2 diabetes are conflicting.

Objectives: We studied prospective associations between the intake of total, plant-based, and animal protein and the risk of pre-diabetes and diabetes in 4 population-based studies included in the PREVIEW project.

Methods: Analyses were conducted with the use of data from 3 European cohorts and 1 Canadian cohort, including 78,851 participants. Protein intake was assessed through the use of harmonized data from food-frequency questionnaires or 3-d dietary records. Cohort-specific incidence ratios (IRs) were estimated for pre-diabetes and diabetes, adjusting for general characteristics, lifestyle and dietary factors, disease history, and body mass index (BMI) and waist circumference; results were pooled based on a random-effects meta-analysis.

Results: Higher total protein intake (g · kg^{-1} · d^{-1}) was associated with lower incidences of pre-diabetes and diabetes (pooled IRs: 0.84; 95% CI: 0.82, 0.87 and 0.49; 95% CI: 0.28, 0.83, respectively); plant-based protein intake was the main determinant (pooled IRs: 0.83; 95% CI: 0.81, 0.86 and 0.53; 95% CI: 0.36, 0.76, respectively). Substituting 2 energy percentage (E%) protein at the expense of carbohydrates revealed increased risks of pre-diabetes and diabetes (pooled IRs: 1.04; 95% CI: 1.01, 1.07 and 1.09; 95% CI: 1.01, 1.18, respectively). Except for the associations between intakes of total protein and plant-based protein (g · kg^{-1} · d^{-1}) and diabetes, all other associations became nonsignificant after adjustment for BMI and waist circumference.

Conclusions: Higher protein intake (g · kg^{-1} · d^{-1}) was associated with a lower risk of pre-diabetes and diabetes. Associations were substantially attenuated after adjustments for BMI and waist circumference, which demonstrates a crucial role for adiposity and may account for previous conflicting findings. This study was registered at ISRCTN as ISRCTN31174892. *Am J Clin Nutr* 2019;109:1310–1318.

Keywords: diabetes, impaired glucose metabolism, protein intake, observational studies, epidemiology

Introduction

A healthy diet is a key modifiable lifestyle factor in a type 2 diabetes (T2D) prevention strategy (1). Over the past decades, numerous observational studies and clinical trials have investigated the role of nutrition in the prevention of T2D. With respect to macronutrients, it has been shown that the quality rather than the quantity of fat and carbohydrates, as reflected by the types of fatty acids and glycemic index (GI), is associated with increased T2D risk (1). The role of protein has not yet been studied as extensively as the other macronutrients, but several studies have been published more recently (2–4).

Dietary proteins have an insulinotropic effect and promote insulin secretion, which leads to increased rate of glucose clearance from the blood (5). Furthermore, both energy-restricted and ad libitum higher-protein diets have shown a greater weight loss and fat loss than achieved with lower-protein diets (6, 7). Because overweight and obesity are the strongest risk factors for T2D, dietary proteins may play a role in diabetes prevention. However, results from clinical trials and observational studies have been mixed. A meta-analysis of 74 randomized controlled trials (RCTs) with durations ranging between 28 d and 12 mo showed beneficial effects of a high-protein diet on several obesity and cardiometabolic parameters, including weight loss and fasting insulin (8). Another meta-analysis of 15 RCTs of ≥12 mo comparing higher-protein with lower-protein diets on health outcomes showed favorable effects on fasting insulin, but not on weight, waist circumference, serum lipids, or blood pressure (9).

Conversely, several large prospective cohort studies have shown detrimental associations between protein intake and T2D risk (3, 4, 10). For instance, within the European Prospective Investigation into Cancer and Nutrition (EPIC)-InterAct study, a case-cohort study including 12,403 incident T2D cases from 8 European countries, a 10-g increase in protein intake was associated with an increased T2D risk (HR: 1.06; 95% CI: 1.02, 1.09) (4). Furthermore, in all studies the associations attenuated or even disappeared after adjustment for BMI (3, 4, 10). Moreover, an individual's response to dietary protein may

be dependent on the degree of insulin sensitivity, which is partly determined by their adiposity level (5).

The objective of the current study was therefore to investigate associations of the intake of total protein, plant-based protein, and animal protein with the incidence of pre-diabetes and diabetes in 4 population studies from the Netherlands, Finland, and Canada, taking into account the role of BMI and waist circumference.

Methods

Study design and population

The PREVention of diabetes through lifestyle Intervention and population studies in Europe and around the World (PREVIEW) project aims to determine the extent to which a high-protein, low-glycemic diet in combination with moderate or high intensity physical activity can reduce the incidence of T2D in pre-diabetic overweight children, adults, and elderly (www.previewstudy.com). Here we report on the analysis of longitudinal population studies based on the use of cohort studies in Europe

and Canada. The present study describes longitudinal analyses from 4 population studies included in PREVIEW: Lifelines (the Netherlands), the Nutrition Questionnaires plus (NQplus) study (the Netherlands), the Cardiovascular Risk in Young Finns Study (YFS; Finland), and the Quebec Family Study (QFS; Canada). Ethical approval was provided by local ethics committees. All participants gave written informed consent before participating. The population studies have been registered in the ISRCTN registry as ISRCTN31174892. **Table 1** gives an overview of the 4 study populations, and **Figure 1** shows the population flowchart.

Lifelines.

Lifelines is a multidisciplinary prospective population-based cohort study examining in a unique 3-generation design the health and health-related behaviors of 167,729 individuals living in the north of the Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, sociodemographic, behavioral, physical, and psychological factors that contribute to the health and disease of the general population, with a special focus on multimorbidities and complex genetics (11). Between 2006 and 2013, a total of 167,729 participants, aged from 6 mo to 93 y were included. At the time of the current analysis, baseline data for 152,180 participants were available. After exclusion of those with missing data for baseline glycated hemoglobin (HbA1c) or fasting plasma glucose (FPG) ($n = 4836$), diet ($n = 8085$), and unreliable dietary data (men with energy intakes <800 or >4200 kcal/d or women with energy intakes <500 or >3500 kcal/d) ($n = 15,064$), the baseline analytical sample comprised 128,290 individuals. Follow-up data for HbA1c or FPG was available for 83,326 participants. After excluding participants with diabetes at baseline [i.e., self-reported ($n = 1881$); HbA1c ≥ 48 mmol/l or FPG ≥ 7 mmol/L ($n = 2147$); or reported use of diabetes medication ($n = 1435$)], $n = 75,778$ remained for the final analyses on incidence diabetes. Analyses on incidence pre-diabetes were conducted after also excluding those with pre-diabetes at baseline ($n = 23,632$) and diabetes at follow-up ($n = 777$ of which $n = 93$ did not have pre-diabetes at baseline), resulting in a sample of $n = 52,053$.

The NQplus study.

The NQplus study is an ongoing longitudinal study on diet and health within Wageningen and surrounding cities, all located in the central part of the Netherlands (12, 13). The main aims of NQplus are to develop a national dietary assessment reference database to generate and validate food-frequency questionnaires (FFQs) for etiologic research and nutrition monitoring, and to establish a database of longitudinal dietary factors and intermediate health outcomes. A total of 2048 adult men and women aged 20–77 y were recruited between May 2011 and February 2013 in the NQplus study, of whom 1638 provided reliable dietary data. After exclusion of those with missing data for HbA1c or FPG at follow-up ($n = 239$), $n = 1399$ remained for the final analyses on incidence diabetes. After excluding those with diabetes at baseline ($n = 83$), the analytic sample on diabetes incidence included 1316 participants. For analyses on incidence pre-diabetes, participants with pre-diabetes at baseline

TABLE 1 Study characteristics of the 4 population studies in PREVIEW[1]

Population study	Country	Baseline	Follow-up	Dietary assessment method	Diabetes ascertainment	n	Age range, y	Men, %
Lifelines	The Netherlands	2006–2013	2014–2017	FFQ	Self-reported diagnosis, diabetes medication use, FPG (mmol/L), HbA1c (mmol/mol)	75,778	18–90	41
NQplus	The Netherlands	2011–2013	2013–2015	FFQ	Self-reported diagnosis, diabetes medication use, FPG (mmol/L), HbA1c (mmol/mol)	1316	20–76	52
(Cardiovascular Risk in) Young Finns Study	Finland	2007–2008	2010–2012	FFQ	Self-reported diagnosis, diabetes medication use, FPG (mmol/L)	1502	30–45	43
Quebec Family Study	Canada	1992–1994	1997	3-d dietary record	Self-reported diagnosis, FPG (mmol/L)	255	18–78	43

[1]FFQ, food-frequency questionnaire; FPG, fasting plasma glucose; HbA1c, glycated hemoglobin; NQplus, Nutrition Questionnaires plus.

and diabetes at follow-up were excluded as well, resulting in an analytic sample of $n = 795$.

Young Finns Study.

The Cardiovascular Risk in Young Finns Study is a multicenter follow-up study in samples from communities surrounding 5 university towns in Finland with medical schools: Helsinki, Kuopio, Oulu, Tampere, and Turku (14). The main aim of the study was to determine the contribution made by childhood lifestyle, biological, and psychological measures to the risk of cardiovascular disease (CVD) in adulthood. In 1980, 4320 boys and girls in 6 age cohorts were randomly chosen from the national register; 83.2% of those invited participated in the study. For PREVIEW, the 27-y follow-up from 2007 to 2008 was used as baseline and the 30-y follow-up in 2010–2012 as follow-up, based on the quality of the dietary and physical activity assessment. In 2007–2008, data were available for $n = 3600$ participants. After exclusion of those with missing baseline data for FPG ($n = 2207$), missing or unreliable dietary intake data ($n = 2066$), and missing follow-up data for FPG ($n = 363$), $n = 1542$ remained for analyses. After excluding

	Lifelines	NQplus	YFS	QFS	Total
Data available of					
	152,180	2,048	3,600	952	158,780
Missing baseline HbA1c/FPG					
	4,836	74	2,207	109	
Missing or unreliable dietary intake data					
	23,149	410	2,066	119	
Complete baseline data					
	128,290	1,564	1,905	745	132,504
Missing follow-up HbA1c/FPG					
	44,964	239	363	462	
Complete follow-up data					
	83,326	1,399	1,542	283	86,550
No baseline DM	75,778	1,316	1,502	255	78,851

Flowchart of the four included cohort studies.
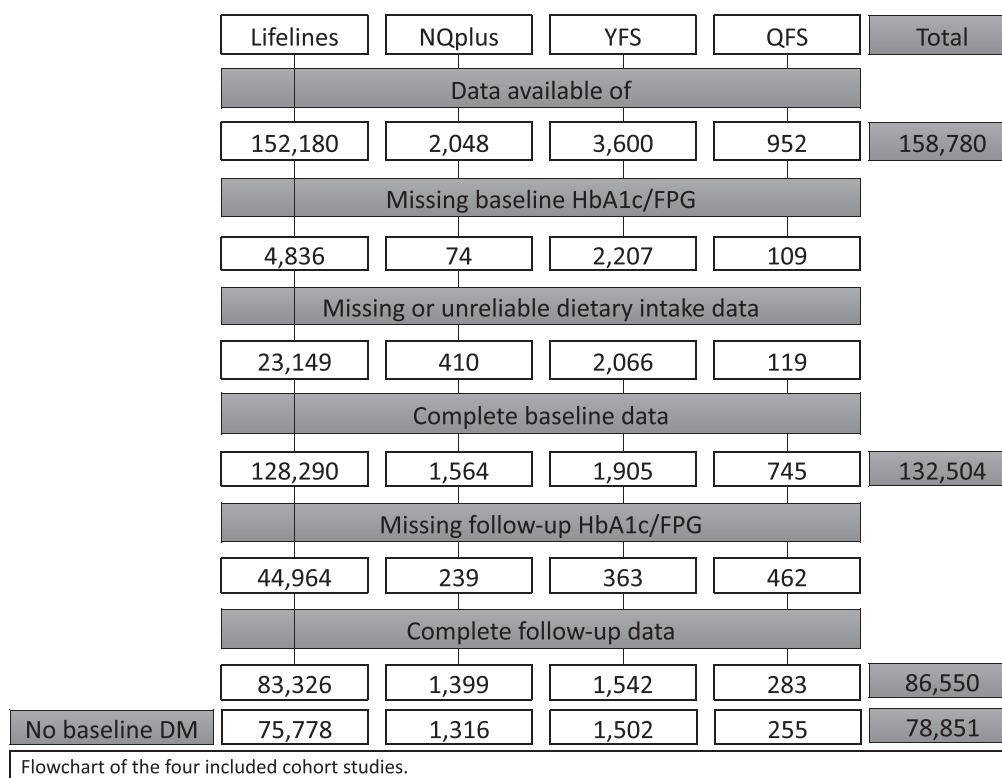
FIGURE 1 Participant flow of the 4 population-based studies included in the PREVIEW project. DM, diabetes mellitus; FPG, fasting plasma glucose; HbA1c, glycated hemoglobin; NQplus, Nutrition Questionnaires plus; PREVIEW, PREVention of diabetes through lifestyle Intervention and population studies in Europe and around the World; QFS, Quebec Family Study; YFS, Young Finns Study.

those with diabetes at baseline, the analytical sample comprised 1502 participants for diabetes incidence. Additionally excluding participants with baseline pre-diabetes and diabetes at follow-up resulted in an analytic sample for incidence pre-diabetes of $n = 1204$.

Quebec Family Study.

The Quebec Family Study is a population-based family study of French-Canadians (15, 16). The aim of the study was to investigate the role of genetics in the etiology of obesity, fitness, and cardiovascular and diabetes risk factors. Participants are Caucasian and of French descent living within ∼80 km of Quebec City, with a socioeconomic status representative of the general French-Canadian population. Participants were recruited through the media in 1978–1982. The study consisted of 3 phases: phase 1 (1978–1982), phase 2 (1992 –1994), and phase 3 (1997–2002). For PREVIEW, phase 2 data acted as baseline and phase 3 data as follow-up, based on the quality of the dietary and physical activity assessment. In total 303 participants provided data both on diet and diabetes incidence. After excluding those with diabetes at baseline ($n = 28$), missing baseline data on diabetes ($n = 2$), unreliable dietary data ($n = 6$), and those aged <18 y ($n = 12$), the analytic sample for diabetes incidence was $n = 255$. The analytic sample for incidence pre-diabetes was $n = 219$ after excluding those with pre-diabetes at baseline and diabetes at follow-up.

Data assessment and harmonization

Data in the PREVIEW study have been collected within the framework of independent population studies, with different protocols for data collection and distinct original research foci. Therefore, data harmonization was a major task of the study. Harmonized variables were created for all parameters of interest for the PREVIEW data analysis.

Dietary assessment

In Lifelines, a newly developed FFQ was used (13). This FFQ was based on a new approach consisting of a combination of 1 main and 3 complementary questionnaires to avoid a long and burdensome FFQ. The main FFQ consists of 110 items and estimates intakes of energy, fat, carbohydrates, protein, and alcohol, and was used for the dietary calculations in PREVIEW. Average daily nutrient intakes were calculated by multiplying frequency of consumption by portion size and nutrient content per gram according to the 2011 Dutch food composition table (17). The validation study of the FFQ used in Lifelines is ongoing, but preliminary analyses showed satisfactory results when comparing Flower-FFQ against a 180-item FFQ (18–20)—validated for energy intake, macronutrients including total protein, animal-based protein, and plant-based protein, dietary fiber, and selected vitamins—as well as urinary nitrogen. Within NQplus, participants completed a validated 180-item semiquantitative FFQ at baseline (18–20). Average daily nutrient intakes were calculated by multiplying frequency of consumption by portion size and nutrient content per gram according to the 2011 Dutch food composition table (17). In the Young Finns Study, a 131-item validated quantitative FFQ was used, which was developed and validated by the Finnish National Institute for Health and Welfare (21). The QFS applied a 3-d dietary record,

including 2 week days and 1 weekend day. Participants were shown how to complete this record by a dietician who provided instruction about measuring the quantities of ingested foods. This method has been shown to provide a reliable measure of diet in this population (22). Mean daily intake was estimated from the computerized version of the Canadian Nutrient File.

Diabetes ascertainment

According to the diagnostic criteria of the American Diabetes Association (23), an individual was classified with diabetes type 1 or 2 with a self-reported diagnosis, self-reported use of diabetes medication, HbA1c ≥48 mmol/mol, or FPG ≥7.0 mmol/L. Pre-diabetes was defined as a self-reported diagnosis, self-reported use of diabetes medication, HbA1c between ≥39 and <48 mmol/mol, or FPG between ≥5.6 and <7.0 mmol/L. The diabetes diagnoses for the different studies are shown in Table 1.

Covariate assessment

Age, sex, educational attainment, smoking status, sleep duration, and medical history were assessed with self- or interviewer-administered questionnaires. Anthropometric measurements including weight, height, and waist circumference were taken by trained personnel. Educational attainment was categorized into less than secondary school qualification (low), secondary school diploma up to university classes but no bachelor's degree (medium), and bachelor's, master's, or PhD degree (high). Smoking status was categorized into never, former, currently smoking <10 cigarettes/d, and currently smoking ≥10 cigarettes/d. The prevalence of CVD, hypertension, and high cholesterol was assessed by self-report as well as the use of medication against hypertension or high cholesterol. In Lifelines and NQplus, physical activity was assessed according to the Short Questionnaire to Assess Health (SQUASH) and the Activity Questionnaire for Adults and Adolescents (AQuAA). In the QFS, physical activity was assessed with a 3-d activity record in which participants recorded a score on a 1–9 scale for every 15-min period over 24 h. The score was based on the estimated representative energy expenditure for each category: 1 corresponding to activities of very low energy expenditure such as sleeping, 9 to activities of very high energy expenditure such as running. Dietary glycemic load (GL) was calculated by multiplying the carbohydrate content of the reported food by the daily quantity consumed and the respective GI value. The GL was summed for all reported foods, after which dietary GI was calculated as the dietary GL divided by the total carbohydrate intake. Within the population studies, dietary GI and GL values were assigned through the use of country-specific GI data based on the raw dietary data on food code level, as far as possible.

Statistical analysis

All statistical analyses were performed with SAS 9.3 (SAS Institute) and IBM SPSS Statistics 22. For NQplus, the Quebec Family Study, and the Young Finns Study, missing values of co-variates were imputed through the use of the multiple imputation method, where all variables included in the statistical models were included in the procedure (24). Five duplicate datasets were produced. After statistical inference on the duplicate datasets, pooled estimates were calculated. Because of the large dataset,

TABLE 2 General characteristics on complete cases of the 4 population studies in PREVIEW[1]

Variable	Lifelines	NQplus	Young Finns Study	Quebec Family Study
N	75,778	1316	1502	255
Mean age, y	45.7 (± 12.6)	53.5 (± 11.3)	38.0 (±5.0)	40.4 (± 15.2)
Median age, y	46.0 (37.0, 54.0)	56.0 (47.0, 63.0)	39.0 (33.0, 42.0)	39.4 (26.2, 53.8)
Males, n (%)	31,080 (41)	684 (52)	642 (43)	110 (43)
Education, n (%)				
Low	1624 (2)	10 (1)	530 (35)	44 (17)
Intermediate	50,331 (67)	574 (43)	607 (41)	139 (55)
High	23,526 (31)	732 (56)	365 (24)	72 (28)
Smoking status, n (%)				
Never	36,078 (48)	693 (53)	762 (51)	216 (85)
Former	25,436 (34)	526 (40)	355 (24)	—
Current <10 cig/d	6772 (9)	43 (3)	258 (17)	11 (4)
Current ≥10 cig/d	7166 (9)	54 (4)	127 (8)	28 (11)
BMI, kg/m^2	25.9 (± 4.1)	25.7 (± 3.9)	25.7 (± 4.5)	25.6 (± 5.0)
BMI categories, n (%)				
Normal weight (<25 kg/m^2)	34,526 (46)	624 (47)	726 (48)	134 (52)
Overweight (25–29.9 kg/m^2)	30,699 (40)	523 (40)	538 (36)	83 (33)
Obese (≥30 kg/m^2)	10,537 (14)	169 (13)	238 (16)	38 (15)
Waist circumference, cm	89.8 (± 11.9)	91.0 (± 12.1)	87.9 (± 12.9)	83.4 (± 13.9)
Sleep duration, h/d	7.5 (± 1.0)	7.3 (± 1.1)	7.4 (± 0.8)	7.7 (± 1.1)
Physical activity, MET-min/wk				
Moderate	1673 (816, 2958)	811 (211, 1681)	223 (33, 667)	2772 (1155, 5544)
Intense	0 (0, 630)	361 (1, 1409)	104 (8, 548)	210 (0, 1680)
Sedentary activities, min/wk	840 (630, 1260)	1861 (1261, 2701)	2101 (1261, 2521)	3185 (2520, 3990)
Medical history, n (%)				
CVD	1645 (2)	32 (2)	0 (0)	—
Hypertension	16,201 (21)	310 (24)	80 (5)	26 (10)
High cholesterol	9886 (13)	236 (18)	176 (12)	27 (11)
Medication use, n (%)				
Hypertension	8741 (12)	157 (12)	97 (6)	25 (10)
High cholesterol	3912 (5)	97 (7)	24 (2)	—

[1]Values are means ± SDs, median (25th, 75th percentiles), or n (%). Missing Lifelines: education (n = 297), smoking (n = 326), BMI (n = 16), waist (n = 16), sleep duration (n = 273), moderate and intense physical activity (n = 5220), sedentary activities (n = 132). cig, cigarettes; CVD, cardiovascular disease; HbA1c, glycated hemoglobin; MET, metabolic task equivalent; NQplus, Nutrient Questionnaires plus; PREVIEW, PREVention of diabetes through lifestyle Intervention and population studies in Europe and around the World.

Lifelines analyses were conducted on complete cases. Before the analyses, protein intake and all other dietary variables were adjusted for energy intake with the residual method (25). Total, plant-based, and animal protein intake was analyzed as grams per kilogram bodyweight, because this is the unit of protein requirements. Incidence ratios (IRs) and accompanying 95% CIs of pre-diabetes and diabetes were calculated by Cox proportional hazard regression, setting time to a constant value and with the use of robust variance estimation (26). Adjusted estimates were generated with the use of 4 adjustment models. Model 1 was adjusted for age, sex, and education (3 categories). Model 2 was additionally adjusted for light, moderate, and intense physical activity (MET-min/wk), sedentary activity (min/wk), alcohol consumption (0, >0–6, 7–12, or >12 g/d), smoking status (categories), sleep duration (h/night), total fat energy percentage (E%), total energy intake (kcal/d), and GI. Model 3 added self-reported prevalence of hypertension (yes/no), high cholesterol (yes/no), and CVD (yes/no), medication use for hypertension and high cholesterol (yes/no). Finally, model 4 also included BMI (kg/m^2) and waist circumference (cm). Subsequently, IRs were pooled by random-effects meta-analysis. To assess the robustness of the results, analyses were stratified for sex and BMI categories, and tested for interaction by adding a product term to the

model. Moreover, a sensitivity analysis was performed within the Lifelines cohort excluding all individuals who reported prevalent diseases, i.e., hypertension, high cholesterol, or CVD, at baseline. Finally, multivariate nutrient density models were analyzed to examine associations of substitutions of protein by carbohydrates and fat with pre-diabetes and diabetes risk. Protein (as nutrient density, E%) was included as the exposure variable, and total energy and other energy-yielding nutrients were included as covariates (as E%, except the nutrient to be replaced).

Results

The mean ± SD age of the 4 study populations ranged from 38.0 ± 5.0 y in YFS up to 53.5 ± 11.3 y in NQplus (**Table 2**). In most cohorts, the majority of participants were educated at intermediate level, except NQplus, where 56% had a higher education. Most individuals were nonsmokers and the proportion of overweight and obese people ranged from 48% in QFS to 54% in Lifelines. Energy intake from protein ranged from 14.9 E% in Lifelines and NQplus to 17.4 E% in YFS. Whereas in NQPlus the intake of plant-based protein and animal protein was rather balanced (mean intake 34.4 compared with 40.2 g/d), the intake of plant-based protein in Lifelines was about two-thirds of that

TABLE 3 Energy and macronutrient intake in the 4 population studies of PREVIEW[1]

Variable	Lifelines	NQplus	Young Finns Study	Quebec Family Study
n	75,778	1316	1502	255
Energy, kcal/d	2058 ($\pm$ 584)	2056 ($\pm$ 568)	2309 ($\pm$ 643)	2274 ($\pm$ 630)
Total protein, E%	14.9 ($\pm$ 2.2)	14.7 ($\pm$ 2.1)	17.5 ($\pm$ 2.4)	15.5 ($\pm$ 2.9)
Total protein, g/d	75.3 ($\pm$ 19.7)	74.5 ($\pm$ 19.1)	100.2 ($\pm$ 28.9)	87.2 ($\pm$ 25.7)
Plant-based protein, g/d	23.8 ($\pm$ 5.9)	34.4 ($\pm$ 10.9)	21.8 ($\pm$ 7.8)	—
Animal protein, g/d	37.5 ($\pm$ 10.7)	40.2 ($\pm$ 13.3)	71.2 ($\pm$ 23.7)	—
Total protein, $g \cdot kg^{-1} \cdot d^{-1}$	0.98 ($\pm$ 0.21)	1.0 ($\pm$ 0.2)	1.4 ($\pm$ 0.3)	1.3 ($\pm$ 0.3)
Total fat, E%	35.3 ($\pm$ 4.9)	35.4 ($\pm$ 5.3)	32.8 ($\pm$ 4.7)	33.0 ($\pm$ 6.0)
Total fat, g/d	81.5 ($\pm$ 28.1)	81.7 ($\pm$ 28.0)	84.1 ($\pm$ 27.0)	84.0 ($\pm$ 30.8)
Saturated fat	NA	27.9 ($\pm$ 10.3)	30.2 ($\pm$ 10.8)	30.2 ($\pm$ 13.0)
MUFA	NA	29.6 ($\pm$ 11.0)	28.4 ($\pm$ 9.7)	29.5 ($\pm$ 11.8)
PUFA	NA	17.3 ($\pm$ 6.8)	13.6 ($\pm$ 4.8)	14.6 ($\pm$ 6.5)
Total carbohydrates, E%	44.9 ($\pm$ 5.6)	43.1 ($\pm$ 5.7)	46.0 ($\pm$ 5.6)	49.4 ($\pm$ 6.7)
Total carbohydrates, g/d	231.0 ($\pm$ 70.7)	221.4 ($\pm$ 67.3)	266.1 ($\pm$ 81.9)	279.4 ($\pm$ 81.3)
Polysaccharides	131.8 ($\pm$ 42.5)	124.5 ($\pm$ 44.2)	129.1 ($\pm$ 48.6)	—
Mono/disaccharides	99.2 ($\pm$ 39.7)	96.8 ($\pm$ 33.8)	135.3 ($\pm$ 47.8)	—
Dietary fiber, g/d	NA	24.2 ($\pm$ 7.2)	23.7 ($\pm$ 9.5)	—
Alcohol intake, E%	1.5 (0.4, 3.6)	2.8 (0.7, 5.6)	1.5 (0.5, 3.3)	0 (0, 3.3)
Alcohol intake, g/d	4.0 (0.9, 10.3)	7.9 (1.9, 16.4)	4.7 (1.7, 10.5)	0 (0, 9.1)
GI	55.9 ($\pm$ 3.3)	52.8 ($\pm$ 4.2)	55.4 ($\pm$ 4.0)	55.7 ($\pm$ 3.5)
GL	129.7 ($\pm$ 42.1)	118.0 ($\pm$ 39.8)	147.9 ($\pm$ 48.0)	144.5 ($\pm$ 42.8)

[1]Values are means $\pm$ SDs, median (25th, 75th percentiles). Missing Lifelines: total protein, $g \cdot kg^{-1} \cdot d^{-1}$ ($n = 16$). GI, glycemic index; GL, glycemic load; NA, not available; NQplus, Nutrient Questionnaires plus; PREVIEW, PREVention of diabetes through lifestyle Intervention and population studies in Europe and around the World.

of animal protein (mean intake 23.8 compared with 37.5 g/d) and even less in the Young Finns study (mean intake 21.8 compared with 71.2 g/d) (**Table 3**).

Table 4 shows the prospective associations between 1 energy-adjusted standard deviation per kilogram of body weight increase in total protein intake (e.g., 0.20 $g \cdot kg^{-1} \cdot d^{-1}$ in NQplus), plant-based protein intake (e.g., 0.12 $g \cdot kg^{-1} \cdot d^{-1}$ in NQplus), and animal protein intake (e.g., 0.16 $g \cdot kg^{-1} \cdot d^{-1}$ in NQplus) and the IR of prediabetes and diabetes in the 4 population studies. After adjustment for general characteristics, lifestyle and dietary factors, a higher protein intake was prospectively associated with a lower risk of prediabetes (pooled IR: 0.84; 95% CI: 0.82, 0.87; $I^2 = 0\%$; $P = 0.81$) and diabetes (IR: 0.49; 95% CI: 0.28, 0.83; $I^2 = 60\%$; $P = 0.08$) (model 2). Analyzing the intake of plant-based protein and animal protein separately indicated that these associations were largely related to the intake of plant-based protein—pooled IR: 0.83 (95% CI: 0.81, 0.86) for pre-diabetes and 0.53 (95% CI: 0.36, 0.76) for diabetes (model 2). These associations were attenuated after additional adjustment for waist circumference (pooled IR for total protein: 1.01; 95% CI: 0.97, 1.04; $I^2 = 0\%$; $P = 0.98$ for prediabetes, and 0.85; 95% CI: 0.75, 0.95; $I^2 = 0\%$; $P = 0.37$ for diabetes) (model 4). Substituting 2 E% protein at the expense of carbohydrates, but not fat, significantly increased the risks of prediabetes and diabetes—pooled IR: 1.04; 95% CI: 1.01, 1.07, and 1.09; 95% CI: 1.01, 1.18, respectively. However, also this association became nonsignificant after adjustment for BMI and waist circumference (**Supplemental Table 1**).

Overall, no interactions were observed between protein intake and sex or BMI categories (data not shown). The stratified analyses by sex and BMI did also not show substantially different associations (data not shown). Excluding persons with

prevalent diseases, while maintaining sufficient power, was only possible in Lifelines and did not alter the findings (**Supplemental Table 2**).

Discussion

The current study aimed to analyze prospective associations between protein intake and different diabetes indicators across 4 studies from Europe and Canada. After pooling the 4 population studies, total protein intake expressed as $g \cdot kg^{-1} \cdot d^{-1}$ was inversely associated with both prediabetes and diabetes incidence. Separate analyses for plant-based protein and animal protein indicated that these associations were largely due to the impact of plant-based protein intake. All these associations were attenuated after additional adjustment for BMI or waist circumference, but remained significant for total protein and plant-based protein intake (grams per kilogram per day) and diabetes.

Previously, Malik et al. (3) reported on the associations between total protein and diabetes incidence in 3 large US cohorts, i.e., the Nurses' Health Study I and II and the Health Professionals Follow-up Study. The pooled HR was 1.07 (95% CI: 1.01, 1.17). A 10-g higher protein intake was associated with an HR of 1.06 (95% CI: 1.02, 1.09) in EPIC-InterAct (4) and with an HR of 1.16 (95% CI: 1.06, 1.26) in EPIC-NL (10). Additionally, a higher protein intake was also related to a higher diabetes incidence in the Malmö Diet and Cancer cohort (27) and the Women's Health Initiative (28). In contrast, analyses within 2332 Finnish middle-aged men from the Kuopio Ischaemic Heart Disease Risk Factor Study did not show associations between total protein and T2D risk (29). A recent dose-response meta-analysis for a 5% energy increment from protein supports these

TABLE 4 Prospective associations between a 1-SD increase in energy-adjusted total protein intake (g · kg^{-1} · d^{-1}) and incidence ratio (IR) of pre-diabetes and diabetes in 4 population studies in PREVIEW[1]

Variable	Lifelines	NQplus	Young Finns Study	Quebec Family Study	Meta-analyses Pooled	I^2, %	P
Incidence ratios of pre-diabetes[2]							
Cases/n	6018/51,878	113/795	157/1204	30/219	—	—	
Total protein							
Model 1	0.84 (0.81, 0.86)	0.79 (0.65, 0.96)	0.84 (0.72, 0.98)	0.70 (0.48, 1.03)	0.84 (0.81, 0.86)	0	0.77
Model 2	0.84 (0.82, 0.87)	0.79 (0.65, 0.97)	0.85 (0.73, 0.99)	0.71 (0.47, 1.09)	0.84 (0.82, 0.87)	0	0.81
Model 3	0.85 (0.82, 0.88)	0.79 (0.65, 0.97)	0.84 (0.72, 0.99)	0.67 (0.43, 1.04)	0.85 (0.82, 0.87)	0	0.66
Model 4	1.00 (0.97, 1.04)	1.05 (0.85, 1.30)	1.00 (0.81, 1.24)	0.99 (0.63, 1.55)	1.01 (0.97, 1.04)	0	0.98
Plant-based protein							
Model 1	0.83 (0.81, 0.85)	0.77 (0.63, 0.94)	0.89 (0.76, 1.06)	NA	0.83 (0.81, 0.85)	0	0.53
Model 2	0.83 (0.81, 0.86)	0.77 (0.62, 0.95)	0.88 (0.74, 1.06)	NA	0.83 (0.81, 0.86)	0	0.62
Model 3	0.84 (0.81, 0.87)	0.78 (0.62, 0.97)	0.88 (0.73, 1.05)	NA	0.84 (0.81, 0.86)	0	0.70
Model 4	0.98 (0.94, 1.02)	1.02 (0.82, 1.27)	1.03 (0.85, 1.25)	NA	0.99 (0.95, 1.02)	0	0.85
Animal-based protein							
Model 1	0.92 (0.89, 0.94)	0.92 (0.79, 1.08)	0.86 (0.75, 1.00)	NA	0.91 (0.89, 0.94)	0	0.73
Model 2	0.91 (0.88, 0.94)	0.92 (0.78, 1.07)	0.87 (0.74, 1.01)	NA	0.91 (0.88, 0.94)	0	0.81
Model 3	0.91 (0.89, 0.94)	0.91 (0.78, 1.06)	0.86 (0.74, 1.00)	NA	0.91 (0.89, 0.94)	0	0.74
Model 4	1.01 (0.97, 1.05)	1.03 (0.88, 1.21)	0.97 (0.81, 1.16)	NA	1.01 (0.98, 1.04)	0	0.88
Incidence ratios of diabetes[3]							
Cases/n	768/75,465	11/1316	11/1502	3/255	—	—	
Total protein							
Model 1	0.47 (0.43, 0.52)	0.22 (0.10, 0.48)	0.98 (0.40, 2.38)	Too few cases	0.46 (0.25, 0.84)	69	0.04
Model 2	0.48 (0.43, 0.53)	0.26 (0.12, 0.58)	1.04 (0.42, 2.59)	Too few cases	0.49 (0.28, 0.83)	60	0.08
Model 3	0.50 (0.45, 0.55)	0.26 (0.11, 0.63)	1.06 (0.41, 2.70)	Too few cases	0.51 (0.29, 0.87)	56	0.10
Model 4	0.85 (0.75, 0.96)	0.47 (0.17, 1.32)	1.43 (0.44, 4.72)	Too few cases	0.85 (0.75, 0.95)	0	0.37
Plant-based protein							
Model 1	0.50 (0.45, 0.54)	0.34 (0.13, 0.91)	0.95 (0.46, 1.95)	NA	0.54 (0.35, 0.84)	46	0.08
Model 2	0.47 (0.43, 0.52)	0.42 (0.16, 1.13)	0.91 (0.43, 1.90)	NA	0.53 (0.36, 0.76)	34	0.22
Model 3	0.50 (0.45, 0.55)	0.46 (0.17, 1.22)	0.90 (0.44, 1.87)	NA	0.54 (0.40, 0.71)	21	0.28
Model 4	0.84 (0.74, 0.95)	0.95 (0.35, 2.60)	1.04 (0.44, 2.50)	NA	0.84 (0.75, 0.95)	0	0.87
Animal-based protein							
Model 1	0.66 (0.60, 0.71)	0.46 (0.25, 0.87)	1.08 (0.46, 2.51)	NA	0.65 (0.51, 0.83)	21	0.28
Model 2	0.64 (0.58, 0.70)	0.41 (0.20, 0.83)	1.16 (0.44, 3.04)	NA	0.63 (0.45, 0.89)	33	0.23
Model 3	0.66 (0.60, 0.72)	0.41 (0.17, 0.99)	1.18 (0.43, 3.21)	NA	0.65 (0.49, 0.88)	18	0.30
Model 4	0.93 (0.84, 1.03)	0.56 (0.19, 1.65)	1.47 (0.47, 4.58)	NA	0.93 (0.84, 1.03)	0	0.48

[1] Values are IR (95% CI). Model 1: adjusted for age, sex, and education. Model 2: Model 1 + light, moderate, and intense physical activity, sedentary activity, alcohol consumption, smoking status, sleep duration, total fat intake (E%), GI, and total energy intake. Model 3: Model 2 + prevalence of hypertension, high cholesterol, and CVD, medication use for hypertension and high cholesterol. Model 4: Model 3 + waist circumference. CVD, cardiovascular disease; GI, glycemic index; IR, incidence ratio; NA, not available; NQplus, Nutrient Questionnaire plus; PREVIEW, PREVention of diabetes through lifestyle Intervention and population studies in Europe and around the World.

[2] Prevalent cases of baseline pre-diabetes and diabetes were excluded.

[3] Prevalent cases of baseline diabetes were excluded. Results were obtained by Cox regression analyses.

results, i.e., showing a higher risk of T2D with higher total protein intakes (RR: 1.27; 95% CI: 1.23, 1.31), largely due to the intake of protein of animal origin (RR: 1.32; 95% CI: 1.27, 1.36). Plant protein showed a borderline inverse association with T2D risk (RR: 0.89; 95% CI: 0.71, 1.11) (30). In most studies, additional adjustment for BMI attenuated, e.g., weakened, the association between protein intake (g/d) and diabetes risk. None of these studies expressed protein intake in g · kg^{-1} · d^{-1}, limiting the possibility of comparing our findings with previous studies.

The role of dietary protein and amino acids on glucose homeostasis is complex and involves multiple mechanisms (31). Higher-protein diets acutely increase energy expenditure by elevating postprandial thermogenesis and resting metabolic rate. During weight loss, high-protein diets attenuate the decrease in resting energy expenditure (6). Although high-quality evidence from RCTs is lacking, high-protein diets also seem to translate into better weight-loss maintenance and improvement of cardiometabolic risk factors in the long term (8, 9). On the other hand, a higher protein intake makes a greater demand for insulin. This increased demand may place too great a burden

on vulnerable persons, such as pre-diabetic individuals, and can adversely affect their insulin secretion. This may explain why the increased risk in the substitution model was more pronounced for diabetes than for pre-diabetes.

It should be noted that most, but not all, RCTs are based on energy-restricted high-protein diets or follow a weight-loss period, whereas participants in observational studies can generally be assumed to be in energy balance—or potentially even in positive energy balance. To alleviate this potential problem, protein intake was adjusted for energy with the residual method in the current study. By regressing total protein intake upon total energy intake and using the residuals in the analyses, variation dependent on total energy intake is adjusted for (25). In general, persons with higher-protein diets were more likely to be older, have a higher BMI, and display an unhealthier lifestyle in observational studies (4, 10). Although the associations were also adjusted for multiple socioeconomic and lifestyle factors, residual confounding due to unmeasured or imperfectly measured confounding factors cannot be ruled out as is typically the case in observational studies.

Furthermore, when protein was analyzed in terms of g · kg^{-1} · d^{-1}, a strong inverse association with diabetes was observed. This may reflect a true inverse association between protein intake and diabetes risk with a prevailing negative energy balance. These findings might mirror the beneficial effects of protein on weight loss as reported by intervention studies (32). However, it also appeared that the association was driven by the protective effect of a lower weight and BMI on diabetes risk, rather than a high protein intake: a higher protein intake per kilogram bodyweight did not only reflect a higher protein intake, but also a lower bodyweight and subsequent BMI in all cohorts.

Indeed, overweight and adiposity are the strongest risk factors for T2D (23). T2D risk increases with increasing body fat, and this association is already visible in the normal range of BMI and waist circumference (1). In some studies, the risk associated with a higher waist circumference, as a marker of central adiposity, is stronger than the risk associated with a higher BMI (33). Therefore, in the present study, associations were adjusted for both BMI and waist circumference. In line with other observational studies (3, 4, 10), the associations were attenuated after these adiposity-related adjustments, revealing that adiposity is likely a mediator on the pathway from protein intake to diabetes. In most population studies, associations even became inverse or nonsignificant upon similar adjustments. Malik et al. (3) speculated that the positive association between protein and diabetes risk is more pronounced in individuals who are more insulin sensitive, as indicated by their adiposity. A modified observational study, minimizing differences between an RCT and observational data, demonstrated that, compared with a low-protein diet, a high-protein diet was associated with better weight maintenance when individuals with greater BMI and waist circumference were analyzed (34). This indicates that the role of adiposity in the association between protein and diabetes is complex, but does not fully account for the potentially detrimental association.

The PREVIEW project covers a large-scale, multinational collaboration of observational studies with reasonably large sample sizes. Pooled analyses of individual participant data within each study is a cost-efficient analytic approach. However, here we had to rely on data that were collected at different time points in the context of variable study objectives and protocols. Nevertheless, the results are based on harmonized and individual participant data which should yield a higher level of evidence than a simple meta-analysis of published studies. Furthermore, it can be considered a strength that the PREVIEW population studies comprise large samples of people with different cultural and ethnic backgrounds and dietary habits.

In conclusion, prospective analyses of individual participant data from 4 population studies, comprising >75,000 subjects, showed that a higher protein intake (g · kg^{-1} · d^{-1}) is associated with a lower risk of both pre-diabetes and diabetes. These associations were attenuated after additional adjustment for BMI, waist circumference, or both. This demonstrates that BMI and waist circumference play a crucial role in the interpretation of the association between protein intake and diabetes occurrence and incidence. Future studies should focus on elucidating the apparent discrepancies between intervention and observational studies on the role of protein in weight loss and diabetes. Moreover, studies on the different food sources of protein and diabetes risk are warranted.

References

1. Ley SH, Hamdy O, Mohan V, Hu FB. Prevention and management of type 2 diabetes: dietary components and nutritional strategies. Lancet 2014;383(9933):1999–2007.
2. Liu AY, Silvestre MP, Poppitt SD. Prevention of type 2 diabetes through lifestyle modification: is there a role for higher-protein diets? Adv Nutr 2015;6(6):665–73.
3. Malik VS, Li Y, Tobias DK, Pan A, Hu FB. Dietary protein intake and risk of type 2 diabetes in US men and women. Am J Epidemiol 2016;183(8):715–28.
4. van Nielen M, Feskens EJ, Mensink M, Sluijs I, Molina E, Amiano P, Ardanaz E, Balkau B, Beulens JW, Boeing H, et al. Dietary protein intake and incidence of type 2 diabetes in Europe: the EPIC-InterAct Case-Cohort Study. Diabetes Care 2014;37(7):1854–62.
5. Rietman A, Schwarz J, Tome D, Kok FJ, Mensink M. High dietary protein intake, reducing or eliciting insulin resistance? Eur J Clin Nutr 2014;68(9):973–9.
6. Leidy HJ, Clifton PM, Astrup A, Wycherley TP, Westerterp-Plantenga MS, Luscombe-Marsh ND, Woods SC, Mattes RD. The role of protein in weight loss and maintenance. Am J Clin Nutr 2015;101(6):1320S–9S.
7. Skov AR, Toubro S, Ronn B, Holm L, Astrup A. Randomized trial on protein vs carbohydrate in ad libitum fat reduced diet for the treatment of obesity. Int J Obes Relat Metab Disord 1999;23(5):528–36.
8. Santesso N, Akl EA, Bianchi M, Mente A, Mustafa R, Heels-Ansdell D, Schunemann HJ. Effects of higher- versus lower-protein diets on health outcomes: a systematic review and meta-analysis. Eur J Clin Nutr 2012;66(7):780–8.
9. Schwingshackl L, Hoffmann G. Long-term effects of low-fat diets either low or high in protein on cardiovascular and metabolic risk factors: a systematic review and meta-analysis. Nutr J 2013;12:48.
10. Sluijs I, Beulens JW, van der AD, Spijkerman AM, Grobbee DE, van der Schouw YT. Dietary intake of total, animal, and vegetable protein and risk of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-NL study. Diabetes Care 2010;33(1):43–8.
11. Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, van Dijk F, van Zon SK, Wijmenga C, Wolffenbuttel BH, et al. Cohort profile: Lifelines, a three-generation cohort study and biobank. Int J Epidemiol 2015;44(4):1172–80.
12. Brouwer-Brolsma EM, van Lee L, Streppel MT, Sluik D, van de Wiel AM, de Vries JHM, Geelen A, Feskens EJM. Nutrition Questionnaires plus (NQplus) study, a prospective study on dietary determinants and cardiometabolic health in Dutch adults. BMJ open 2018;8(7): e020228.
13. Brouwer-Brolsma EM, Streppel MT, van Lee L, Geelen A, Sluik D, van de Wiel AM, de Vries JHM, van't Veer P, Feskens EJM. A National Dietary Assessment Reference Database (NDARD) for the Dutch population: rationale behind the design. Nutrients 2017;9(10): 10–13. doi:10.3390/nu9101136.
14. Raitakari OT, Juonala M, Ronnemaa T, Keltikangas-Jarvinen L, Rasanen L, Pietikainen M, Hutri-Kahonen N, Taittonen L, Jokinen E, Marniemi J, et al. Cohort profile: the Cardiovascular Risk in Young Finns Study. Int J Epidemiol 2008;37(6):1220–6.
15. Chaput JP, Leblanc C, Perusse L, Despres JP, Bouchard C, Tremblay A. Risk factors for adult overweight and obesity in the Quebec Family Study: have we been barking up the wrong tree? Obesity (Silver Spring) 2009;17(10):1964–70.
16. Chaput JP, Perusse L, Despres JP, Tremblay A, Bouchard C. Findings from the quebec family study on the etiology of obesity: genetics and environmental highlights. Curr Obes Rep 2014;3:54–66.
17. NEVO-tabel. Dutch Food Composition Table 2011/version 3. Bilthoven: RIVM; 2011.

18. Streppel MT, de Vries JH, Meijboom S, Beekman M, de Craen AJ, Slagboom PE, Feskens EJ. Relative validity of the food frequency questionnaire used to assess dietary intake in the Leiden Longevity Study. Nutr J 2013;12:75.

19. Feunekes GI, Van Staveren WA, De Vries JH, Burema J, Hautvast JG. Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol. Am J Clin Nutr 1993;58(4):489–96.

20. Siebelink E, Geelen A, de Vries JH. Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. Br J Nutr 2011;106(2):274–81.

21. Paalanen L, Mannisto S, Virtanen MJ, Knekt P, Rasanen L, Montonen J, Pietinen P. Validity of a food frequency questionnaire varied by age and body mass index. J Clin Epidemiol 2006;59(9):994–1001.

22. Tremblay A, Sévigny J, Leblanc C, Bouchard C. The reproducibility of a three-day dietary record. Nutr Res 1983;3(6):819–30.

23. American Diabetes Association. Standards of medical care in diabetes—2013. Diabetes Care 2013;36 Suppl 1:S11–66.

24. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.

25. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65(4 Suppl):S1220–8; discussion S9–31.

26. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Method 2003;3:21.

27. Ericson U, Sonestedt E, Gullberg B, Hellstrand S, Hindy G, Wirfalt E, Orho-Melander M. High intakes of protein and processed meat associate with increased incidence of type 2 diabetes. Br J Nutr 2013;109(6):1143–53.

28. Tinker LF, Sarto GE, Howard BV, Huang Y, Neuhouser ML, Mossavar-Rahmani Y, Beasley JM, Margolis KL, Eaton CB, Phillips LS, et al. Biomarker-calibrated dietary energy and protein intake associations with diabetes risk among postmenopausal women from the Women's Health Initiative. Am J Clin Nutr 2011;94(6):1600–6.

29. Virtanen HEK, Koskinen TT, Voutilainen S, Mursu J, Tuomainen TP, Kokko P, Virtanen JK. Intake of different dietary proteins and risk of type 2 diabetes in men: the Kuopio Ischaemic Heart Disease Risk Factor Study. Br J Nutr 2017;117(6):882–93.

30. Zhao LG, Zhang QL, Liu XL, Wu H, Zheng JL, Xiang YB. Dietary protein intake and risk of type 2 diabetes: a dose-response meta-analysis of prospective studies. Eur J Nutr 2018. doi:10.1007/s00394-018-1737-7.

31. Azzout-Marniche D, Gaudichon C, Tome D. Dietary protein and blood glucose control. Curr Opin Clin Nutr Metab Care 2014;17(4):349–54.

32. Westerterp-Plantenga MS, Lemmens SG, Westerterp KR. Dietary protein—its role in satiety, energetics, weight loss and health. Br J Nutr 2012;108 Suppl 2:S105–12.

33. Vazquez G, Duval S, Jacobs DR, Jr., Silventoinen K. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. Epidemiol Rev 2007;29:115–28.

34. Ankarfeldt MZ, Angquist L, Stocks T, Jakobsen MU, Overvad K, Halkjaer J, Saris WH, Astrup A, Sorensen TI. Body characteristics, [corrected] dietary protein and body weight regulation. Reconciling conflicting results from intervention and observational studies? PloS one 2014;9(7):e101134.