

C. Heidemann · K. Hoffmann · J. Spranger ·
K. Klipstein-Grobusch · M. Möhlig · A. F. H. Pfeiffer ·
H. Boeing

A dietary pattern protective against type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)—Potsdam Study cohort

Received: 12 August 2004 / Accepted: 28 January 2005 / Published online: 12 May 2005
© Springer-Verlag 2005

Abstract *Aims/hypothesis:* The aim of this study was to identify a dietary pattern associated with diabetes-related biomarkers and to investigate whether this pattern is associated with the incidence of type 2 diabetes. *Methods:* A nested case–control study of 192 cases of incident type 2 diabetes and 382 control subjects matched for sex and age was conducted. All subjects were participants in the population-based European Prospective Investigation into Cancer and Nutrition (EPIC)—Potsdam Study. Dietary pattern score was derived using intake data on 48 food groups as exposure variables and the biomarkers HbA_{1c}, HDL cholesterol, C-reactive protein and adiponectin as response variables in reduced rank regression. The association of the score with diabetes risk was estimated by conditional logistic regression analysis. *Results:* A high score for the identified dietary pattern was characterised by a high intake of fresh fruit and a low intake of high-caloric soft drinks, beer, red meat, poultry, processed meat, legumes and bread (excluding wholegrain bread). Subjects with high scores had high plasma concentrations of HDL cholesterol and adiponectin and low plasma concentrations of HbA_{1c} and C-reactive protein. After multivariate adjustment, the odds ratios for type 2 diabetes across increasing quintiles of the dietary pattern score were 1.0, 0.59, 0.51,

0.26 and 0.27, respectively ($p=0.0006$ for trend). *Conclusions/interpretation:* A high score for the identified dietary pattern is associated with a more favourable biomarker profile and a substantially reduced incidence of type 2 diabetes.

Keywords Adiponectin · Biomarker · C-reactive protein · Dietary pattern · EPIC · Epidemiology · HbA_{1c} · HDL cholesterol · Type 2 diabetes

Abbreviations CRP: C-reactive protein · EPIC: European Prospective Investigation into Cancer and Nutrition · OR: odds ratio · RRR: reduced rank regression

Introduction

Type 2 diabetes is an increasingly common and costly chronic disease that predisposes affected individuals to substantial risk in terms of morbidity and mortality [1, 2]. A knowledge of the risk factors and protective factors associated with type 2 diabetes is essential for the development of prevention strategies. Diet is thought to have an important influence on the development of diabetes. The traditional approach used to investigate diet–disease associations focuses on single dietary components, such as single nutrients or foods. For example, it has been suggested that different types of fatty acids [3], cereal fibre [4], and magnesium [5] are associated with the risk of diabetes. However, diet is a complex entity, and therefore it is difficult to investigate the effects of individual dietary components separately. Thus, within the field of nutritional epidemiology, the use of dietary patterns has been proposed. A dietary pattern is a comprehensive variable of several foods or food groups that takes into consideration the interactions and cumulative effects of dietary components on disease risk. To date, patterns have been obtained using a hypothesis-oriented approach, such as diet indices, or an exploratory approach, such as cluster and factor analysis [6]. Factor analysis, the method most commonly used to identify patterns associated with diabetes [7–10], derives

C. Heidemann (✉) · K. Hoffmann · K. Klipstein-Grobusch ·
H. Boeing
Department of Epidemiology,
German Institute of Human Nutrition Potsdam-Rehbruecke,
Arthur-Scheunert-Allee 114-116,
14558 Nuthetal, Germany
e-mail: heidemann@mail.dife.de
Tel.: +49-33200-88731
Fax: +49-33200-88721

J. Spranger · M. Möhlig · A. F. H. Pfeiffer
Department of Clinical Nutrition,
German Institute of Human Nutrition Potsdam-Rehbruecke,
Nuthetal, Germany

J. Spranger · M. Möhlig · A. F. H. Pfeiffer
Department of Endocrinology, Diabetes and Nutrition,
Charité-University Medicine Berlin,
Berlin, Germany

the typical dietary patterns of a specific study population, regardless of their relevance to health outcomes. Thus, the use of factor analysis for the identification of a pattern associated with a disease of interest is of limited value. Reduced rank regression (RRR) has recently been proposed as an approach that can be used to provide the link between overall diet and disease [11, 12]. This method combines the dietary information of the particular study population with prior scientific knowledge of the pathway from diet to disease.

The aim of this study was, first, to identify a dietary pattern cross-sectionally associated with plasma concentrations of diabetes-related biomarkers using RRR and, secondly, to investigate whether the pattern obtained is prospectively associated with type 2 diabetes.

Subjects and methods

Study population A nested case–control study on the risk of developing type 2 diabetes was designed within the European Prospective Investigation into Cancer and Nutrition (EPIC)—Potsdam Study. The EPIC—Potsdam Study is a prospective cohort study of 27,548 individuals aged 35–65 years recruited from the general population. Baseline examinations, including self-administered questionnaires on dietary and lifestyle habits, a computer-guided interview on further lifestyle habits and medical history, anthropometric measurements, and blood sampling were carried out between August 1994 and September 1998. Follow-up questionnaires are mailed every 2–3 years to obtain updated measurements of baseline characteristics and to evaluate newly developed diseases, including type 2 diabetes and current medication. Case subjects were defined as participants without type 2 diabetes at baseline who developed type 2 diabetes during the first follow-up period. We identified 399 potential cases of incident type 2 diabetes from self-reports on incident diseases, current medication and current dietary treatment for diabetes. These cases were verified by the patients' medical records. Up to November 2001, 201 cases of incident diabetes were medically verified. The subsequent biochemical analysis of the diabetes-associated antibodies GAD65 and IA-2 revealed that nine of these individuals had type 1 diabetes, while the remaining 192 individuals had type 2 diabetes. Each of the case subjects was matched with two control subjects in terms of sex and age (± 1 year). Subjects with missing values for exposure variables or one of the confounding variables used in the subsequent statistical analyses were excluded. Consequently, the final study population consisted of 192 diabetic subjects and 382 control subjects.

The EPIC—Potsdam Study was approved by the Ethics Committee of the Landesärztekammer Brandenburg, Germany. Each participant gave written consent before enrolment into the study.

Dietary assessment All participants were asked to complete a semi-quantitative food-frequency questionnaire. This questionnaire assessed the average frequency of in-

take and the portion size of 148 foods consumed during the 12 months prior to examination. Frequency of intake was measured using ten categories, ranging from 'never' to 'five times per day or more'. Portion sizes were estimated using photographs of standard portion sizes. Information on frequency of intake and portion size was used to calculate the amount of each food item in grams consumed on average per day. The validity and reproducibility of the food-frequency questionnaire have been described previously [13, 14]. The 148 single food items of the questionnaire were aggregated into 48 food groups. The grouping criteria were based on the 44 food groups previously described in detail [15], and on additional categories for cooked vegetables (legumes and cooked vegetables), cheese (low- and high-fat), other dairy products (low- and high-fat) and soft drinks (low- and high-energy).

Covariate assessment Body height and weight, and waist and hip circumferences were measured by trained and quality-monitored personnel, with participants wearing light underwear [16]. BMI was calculated as body weight in kilograms divided by body height in metres squared, and WHR as waist circumference divided by hip circumference.

Information on leisure-time activity, smoking status, educational level, and intake of lipid-lowering and anti-inflammatory drugs was obtained from a self-administered questionnaire and a personal computer-guided interview. Leisure-time activity was calculated as the average of hours of physical activity per week during the summer and winter. Smoking status was defined as current smoker, ex-smoker for <5 years, and ex-smoker for ≥ 5 years or non-smoker. Educational level was defined as less than high school education, high school education, or university degree. Users of lipid-lowering or anti-inflammatory drugs were defined as subjects reporting a regular intake of lipid-lowering drugs, or use of antiphlogistics, analgesics, antirheumatics or corticoids, respectively.

Selection of response variables Type 2 diabetes is a complex condition with various antecedents. For our analysis we chose the four plasma parameters HbA_{1c}, HDL cholesterol, C-reactive protein (CRP), and adiponectin as response variables as they are affected by diet and are related to the pathophysiology of diabetes.

HbA_{1c} reflects elevated blood glucose concentrations [17] and is commonly used as an indicator of long-term glycaemic control [18]. Since impairment of glucose homeostasis increases the risk of type 2 diabetes development [19–21], we used HbA_{1c} as one of the four response variables in the analysis. HDL cholesterol was selected because studies report that insulin resistance and type 2 diabetes are commonly accompanied by decreased HDL cholesterol concentrations, with an increased flux of NEFA to the liver considered as an underlying metabolic effect [23]. Studies indicate that diabetes is associated with a chronic subclinical inflammatory process [24] that is characterised by changes in concentrations of acute-phase proteins [25]. Of these, an elevated CRP concentration is currently the best marker of the future risk of type 2

diabetes [26–28]. The adipocyte-derived peptide adiponectin has been reported to have insulin-sensitising properties and anti-inflammatory effects [29–31]. Although the mechanisms linking adiponectin to diabetes risk remain speculative, the plasma adiponectin level seems to be an independent predictor of type 2 diabetes [32].

Evidence suggests that the four diabetes-related biomarkers we chose are affected by dietary components, especially by alcohol consumption [33–36], fat composition [37, 38], and certain types of carbohydrates [39, 40].

Laboratory procedures At the baseline examination of the EPIC—Potsdam Study, peripheral venous citrate blood samples were taken and centrifuged at $1,000\times g$ for 10 min at 4°C. Plasma was removed and stored in aliquots at –80°C until assays of biochemical markers were performed. The diabetes-associated antibodies GAD65 and IA-2 were measured by RIA (Medipan Diagnostics, Selchow, Germany). HbA_{1c} was analysed by enzyme immunoassay (DAKO Diagnostics, Hamburg, Germany; intra-assay CV 2.0–3.0%, interassay CV 1.9%). HDL cholesterol was measured by enzyme calorimetric assay (ABX Diagnostics, Göppingen, Germany; interassay CV 2.1%). CRP was determined by high-sensitivity ELISA (Immun Diagnostik, Bensheim, Germany; intra-assay CV 4.9–6.2%, interassay CV 13.2%). The limit of detection for CRP was 0.124 µg/ml. For the subsequent statistical analysis, CRP values below the limit of quantification were set at 0.7 times the respective limit of quantification [41]. Adiponectin was measured by RIA (Linco Research, St Charles, MO, USA; intra-assay CV 0.1–6.2%, interassay CV 5.0%). All assay procedures were performed as described by the manufacturer.

Statistical analysis Analyses were performed by using SAS software, Version 8.02 (SAS Institute, Cary, NC, USA). To test for significant differences in baseline characteristics between type 2 subjects and control subjects the non-parametric Kruskal–Wallis test was used for continuous variables and the Mantel–Haenszel test was used for categorical variables.

The statistical method RRR, also known as the maximum redundancy analysis, was employed using the PLS procedure in SAS to derive dietary pattern scores. The use of this method in dietary pattern analyses has recently been described in detail and has been compared with other methods elsewhere [11]. Briefly, the method allows the calculation of scores similar to those extracted by factor analysis. Whereas factor analysis determines scores, i.e. linear combinations of predictor variables, by maximising the explained variation of all predictor variables (e.g. food groups), RRR determines linear combinations of predictor variables by accounting for as much of the variation in response variables (e.g. disease-related biomarkers) as possible. In the special case of only one response variable, RRR is identical to multiple linear regression.

We used intake data on 48 food groups as predictors, and logarithmically transformed data for HbA_{1c}, HDL cholesterol, CRP, and adiponectin levels as responses to derive a dietary pattern score. To reduce confounding by

medication, HDL cholesterol values were adjusted for intake of lipid-lowering drugs, whereas values of CRP and adiponectin were adjusted for intake of anti-inflammatory drugs. As the number of extracted scores cannot be higher than the number of selected responses, four scores were obtained. Because the first dietary pattern score explained much more variation than the subsequent ones, this was the only one considered in the subsequent analysis. Factor loadings represent the correlations of each food group with the dietary pattern score. To focus on food groups that significantly contributed to the dietary pattern score we only considered food groups with absolute factor loadings >0.2.

Crude and adjusted (for age, sex, leisure-time activity, smoking status, educational level and total energy intake, with or without adjustment for BMI and WHR) odds ratios (ORs) and 95% CIs for the risk of developing type 2 diabetes according to quintiles of the dietary pattern score were determined by conditional logistic regression analysis.

To examine the robustness of results, several sensitivity analyses were performed using different sets of response variables. First, we selected the four possible combinations of three response variables. Second, we added IL-6 and TNF-α separately to the four originally chosen responses. Finally, we chose all six responses together. The first dietary pattern score was derived in each case. Multivariate-adjusted ORs and 95% CIs for the incidence of diabetes across quintiles of the respective scores were determined again by logistic regression analysis. A *p* value of less than 0.05 was considered statistically significant for all analyses.

Results

The baseline characteristics of the study population are summarised in Table 1. Compared with control subjects, the diabetic subjects had higher mean values for BMI and WHR and were less likely to have a university degree. Leisure-time activity, total energy intake and percentage of current smokers did not differ significantly between the two groups. Intake of lipid-lowering drugs was more frequent among diabetic study participants than control subjects, whereas use of anti-inflammatory drugs was similar. Plasma concentrations of biomarkers differed significantly between diabetic subjects and control subjects, with the former having higher values for HbA_{1c} and CRP, and lower values for HDL cholesterol and adiponectin.

Dietary pattern scores were derived using intake data on 48 food groups as predictors and logarithmically transformed medication-adjusted concentrations of the four biomarkers as responses in RRR. Four scores were obtained. The first score explained 10.2% of the variation in adiponectin (compared with 15.5% maximally explainable by all 48 food groups), 9.4% of HDL cholesterol variation (14.2%), 9.1% of HbA_{1c} variation (14.8%), 1.1% of CRP variation (13.4%), and 7.4% (14.5%) of the total variation in the four selected biomarkers. The second, third and fourth scores accounted for only 3.6%, 1.9% and 1.5% of the total variation in biomarker levels, respectively. We therefore only included the first score in further analyses.

Table 1 Baseline characteristics of the study population

Characteristic	Diabetic subjects (<i>n</i> =192)	Control subjects (<i>n</i> =382)	<i>p</i> value ^a
Sex (male)	58.9	58.9	Matching variable
Age (years)	55.5±6.8	55.5±6.8	Matching variable
BMI (kg/m ²)	30.8±4.8	26.7±3.6	<0.0001
WHR	0.95±0.09	0.89±0.09	<0.0001
Leisure-time activity (h per week)	10.8±7.7	11.2±7.1	0.24
Total energy intake (kJ per day)	9,250±3,040	9,110±2,730	0.85
Current smoker	19.8	20.7	0.80
University degree	28.1	39.5	0.007
Receiving anti-inflammatory drugs	19.8	15.2	0.16
Receiving lipid-lowering drugs	9.90	4.97	0.03
HbA _{1c} (%)	6.39±2.15	4.73±0.74	<0.0001
HDL cholesterol (mmol/l)	0.90±0.23	1.06±0.28	<0.0001
CRP (µg/ml)	4.12±5.14	2.45±4.37	<0.0001
Adiponectin (µg/ml)	5.96±3.50	7.95±4.58	<0.0001

Data are means±SD or percentages

^aValues are based on the Mantel–Haenszel test or the Kruskal–Wallis test

The association between each food group and the dietary pattern score is indicated in Table 2. Food groups were sorted according to decreasing absolute values of factor loadings. The factor loadings represent the correlations of the single food groups with the score. The mean intakes of the 48 food groups across quintiles of the score are also provided. Red meat, beer, legumes, poultry, high-caloric soft drinks, processed meat, and bread (excluding wholegrain bread) were inversely correlated with the score (factor loadings <−0.2) and were characterised by a steady decrease in intake across quintiles of the dietary pattern score ($p<0.0001$ for trend). In contrast, fresh fruit was directly associated with the score (factor loading >0.2) and showed an increased intake according to quintiles of the score ($p<0.0001$ for trend). All other food groups made a smaller contribution to the dietary pattern score (absolute factor loadings <0.2).

The mean plasma biomarker concentrations according to quintiles of dietary pattern score are presented in Table 3. HbA_{1c} values consistently decreased across quintiles of the score ($p<0.0001$ for trend), whereas concentrations of HDL cholesterol and adiponectin increased ($p<0.0001$ for trend).

No significant trend was evident for CRP concentration across quintiles of the score ($p=0.18$ for trend).

Table 4 shows the risk of developing type 2 diabetes for each quintile of the dietary pattern score. In the crude model (model 1), the risk was consistently and significantly reduced across quintiles of the score. Subjects in the highest quintile had an 80% lower diabetes risk than those in the lowest quintile (OR=0.20, 95% CI 0.10–0.40, $p<0.0001$ for trend). After adjusting for known risk factors (except for BMI and WHR) a further slight decrease in diabetes risk was observed (OR=0.18, 95% CI 0.09–0.37, $p<0.0001$ for trend). Additional adjustment for BMI and WHR, which are strong predictors of type 2 diabetes, reduced the magnitude of the decrease in risk estimates across quintiles of the score. However, subjects in the highest quintile still had a >70% lower risk of developing diabetes than subjects in the lowest quintile (OR=0.27, 95% CI 0.13–0.64, $p=0.0006$ for trend).

To examine the dependence of the results on the choice of biomarkers we derived the first dietary pattern scores for varying sets of biomarkers (Table 5). The risk of type 2 diabetes significantly decreased with increasing quintiles for each of the dietary pattern scores.

Discussion

We have successfully identified a dietary pattern that is associated with plasma concentrations of four diabetes-related biomarkers and with the incidence of type 2 diabetes. A high pattern score reflected a diet that included a high intake of fresh fruit and low intakes of high-caloric soft drinks, beer, red meat, processed meat, poultry, legumes and bread (excluding wholegrain bread). Together, these eight food groups explained about 70% of the variation in dietary pattern score (data not shown). Furthermore, subjects with a higher dietary pattern score were characterised by higher plasma concentrations of HDL cholesterol and adiponectin, and by lower plasma concentrations of HbA_{1c} and CRP.

The results of our study show parallels with previous studies. In a population-based cross-sectional study, a so-called “healthy balanced” dietary pattern was identified. This pattern was mainly characterised by a frequent fruit intake, and was positively associated with HDL cholesterol concentration and inversely correlated with type 2 diabetes [7]. In a large prospective study of men, a ‘prudent pattern’ with a high intake of fruit was identified that was associated with a moderate reduction in the risk of subsequent diabetes [9]. Within the same study population the authors found that red meat and processed meat were the two main components of a diet named “western pattern”. This pattern contributed to an increased incidence of type 2 diabetes [9]. Alone, a high intake of processed meat, adjusted for known risk factors or other dietary components, was associated with an increased risk of diabetes in men and women, respectively [42, 43]. Other studies have reported that a dietary pattern including a high intake of refined bread is

Table 2 Factor loadings and intake of food groups in the study population ($n=574$) according to quintiles of the first identified dietary pattern score

Food group	Factor loading	Food group intake (g/day)					<i>p</i> value for trend
		Q1	Q2	Q3	Q4	Q5	
Red meat	-0.36	65.4±3.6	46.7±2.2	42.9±2.3	31.8±1.7	25.7±1.7	<0.0001
Beer	-0.35	636±61	243±28	209±26	131±19	47.2±9.2	<0.0001
Poultry	-0.27	22.6±2.0	15.8±1.2	12.2±1.0	10.1±0.8	8.37±0.7	<0.0001
Legumes	-0.27	45.8±3.1	29.6±2.0	26.0±1.7	22.0±1.5	20.2±1.6	<0.0001
High-caloric soft drinks	-0.27	139±27	56.4±10.3	27.5±8.7	21.6±5.6	18.1±6.3	<0.0001
Processed meat	-0.26	91.2±7.5	70.7±3.5	63.8±3.5	57.4±3.2	42.0±2.9	<0.0001
Bread (except wholegrain bread)	-0.24	170±7	168±8	128±7	128±7	102±6	<0.0001
Fresh fruit	0.24	105±7	122±8	120±6	146±9	204±12	<0.0001
Cooked potatoes	-0.19	106±5	104±5	94.0±3.8	82.5±4.5	69.8±3.9	<0.0001
Eggs	-0.18	21.7±1.1	21.3±1.8	16.4±1.1	16.3±1.2	11.6±0.9	<0.0001
Sauce	-0.17	24.5±2.0	16.7±1.0	17.4±1.3	14.8±1.3	11.2±1.1	<0.0001
Sweet bread spreads	0.16	8.77±0.94	10.3±1.0	11.8±1.1	13.0±1.1	16.7±1.7	<0.0001
Fried potatoes	-0.15	21.3±1.7	20.1±1.6	13.7±1.0	13.8±1.2	12.8±1.4	<0.0001
Wholegrain bread	0.13	32.2±4.4	26.5±3.3	42.2±4.7	41.7±4.9	54.0±5.3	<0.0001
Tea	0.13	185±27	246±24	283±33	280±46	396±61	0.0004
Cooked vegetables	-0.12	67.3±3.6	60.0±2.8	61.6±3.1	50.3±2.3	51.0±2.7	<0.0001
Butter	-0.12	13.6±1.8	10.2±1.2	8.63±1.16	8.12±1.03	7.50±0.94	0.0003
Spirits	-0.12	6.22±1.35	3.29±0.67	4.48±1.24	2.78±0.71	1.41±0.43	0.0008
Coffee	0.10	395±32	322±21	370±30	410±23	515±35	0.0003
Vegetarian dishes	0.10	0.32±0.13	0.37±0.11	0.88±0.33	0.62±0.18	1.41±0.34	0.002
Garlic	-0.10	0.18±0.05	0.14±0.04	0.09±0.02	0.06±0.01	0.07±0.02	0.002
Cornflakes	0.10	0.15±0.05	1.31±0.45	0.90±0.24	0.84±0.24	2.13±0.66	0.005
Fruit juice	-0.09	200±24	216±21	208±18	161±16	125±12	0.001
High-fat cheese	-0.08	32.8±2.7	28.5±2.2	27.1±2.7	23.1±1.9	23.4±2.2	0.001
High-fat dairy products	-0.08	131±27	114±13	74.7±8.8	96.3±12.8	89.3±10.0	0.04
Animal fat (except butter)	-0.08	0.52±0.12	0.30±0.07	0.22±0.05	0.31±0.06	0.24±0.06	0.02
Canned fruit	-0.08	24.7±3.1	22.3±2.5	23.4±2.6	21.9±2.7	15.0±1.8	0.02
Soup	-0.08	55.2±4.0	47.9±4.9	46.8±4.0	49.7±3.9	37.7±3.7	0.01
Wine	0.08	46.8±9.7	62.4±11.4	51.6±6.5	50.5±5.9	78.5±10.6	0.08
Low-caloric soft drinks	-0.07	28.6±12.0	7.19±3.66	12.8±7.0	3.24±1.64	1.22±0.55	0.004
Cake, cookies	0.07	56.0±5.0	63.2±5.5	50.1±4.2	72.4±6.3	76.8±8.6	0.01
Vegetable oils	-0.06	3.65±0.34	3.49±0.30	3.57±0.17	2.64±0.20	3.02±0.26	0.02
Fish	-0.06	28.7±2.1	25.6±1.8	31.4±2.8	25.9±3.2	21.7±2.2	0.08
Mushrooms	-0.06	2.30±0.23	2.10±0.22	1.99±0.20	1.83±0.18	1.91±0.19	0.10
Confectionery, ice cream	-0.05	23.5±2.3	23.0±1.9	19.7±2.1	20.7±1.9	17.5±1.9	0.03
Decaffeinated coffee	-0.05	58.9±14.8	63.5±16.6	72.2±17.6	43.4±13.0	33.7±12.4	0.14
Pizza	-0.04	6.47±0.97	6.32±0.83	5.60±0.71	5.61±0.69	4.33±0.53	0.04
Nuts	-0.04	3.42±0.47	4.18±1.27	2.57±0.61	2.99±0.72	2.01±0.34	0.09
Dessert	-0.04	17.1±2.0	14.8±1.3	11.6±1.0	15.9±1.9	12.6±1.5	0.11
Pasta, rice	-0.03	16.3±1.4	16.2±1.4	14.8±1.1	12.6±0.9	14.5±1.5	0.08
Raw vegetables	0.03	48.2±4.0	57.2±4.6	48.1±3.1	47.8±2.8	54.6±3.5	0.77
Water	0.02	424±50	322±36	394±35	429±39	495±42	0.05
Other alcoholic beverages	-0.02	1.90±0.66	1.44±0.29	1.43±0.30	1.23±0.21	1.81±0.56	0.77

Table 2 (continued)

Food group	Factor loading	Food group intake (g/day)					<i>p</i> value for trend
		Q1	Q2	Q3	Q4	Q5	
Muesli	0.01	4.32±1.60	3.83±1.02	5.53±1.13	4.73±1.03	4.95±1.31	0.58
Margarine	-0.01	15.0±1.4	16.2±1.4	15.7±1.2	15.6±1.3	14.6±1.1	0.74
Chips, salt sticks	0.01	2.54±0.43	1.42±0.18	2.90±0.76	2.03±0.31	2.31±0.55	0.91
Low-fat cheese	0.01	6.51±1.25	8.30±1.89	8.17±1.44	8.79±1.37	6.38±0.96	0.96
Low-fat dairy products	-0.0003	108±30	112±20	95.1±12.8	119±14	85.5±11.7	0.53

Data are means±SD

Q Quintile of dietary pattern score

correlated with a detrimental biomarker profile [44] and an increased risk of diabetes [8, 9].

Studies investigating the relationship between alcohol consumption or type of alcoholic beverage and diabetes risk have shown inconsistent results. In our study, subjects with a lower dietary pattern score consumed larger amounts of beer and showed an increased risk for type 2 diabetes. Furthermore, these subjects were characterised by a higher consumption of high-caloric soft drinks. Both types of beverage may have several diabetogenic effects, including contributing to an excess energy intake, to overweight and disturbances in carbohydrate metabolism [45–48].

In one prospective study of men, poultry and legumes were included in the “prudent pattern” that was inversely associated with diabetes risk [9]. Furthermore, it was suggested that legumes are good sources of dietary fibre, which has a positive effect on concentrations of cholesterol, glucose and insulin in the blood [49], and contributes to improved glycaemic control [50]. Nevertheless, we found no direct evidence in the literature to suggest that either poultry or legumes are associated with the risk of type 2 diabetes. In our study we observed that the intake of poultry was directly correlated with a dietary pattern associated with an increased risk of diabetes, as was the intake of processed meat and red meat. Moreover, we found that a high consumption of legumes was linked to a dietary pattern score associated with an increased risk of type 2 diabetes. This finding may be explained by the fact that most of the intake of legumes was attributable to a stew of lentils, peas or beans, which is preferably accompanied by bacon, sausages, beef or pork in this region.

Only a few epidemiological studies have identified dietary patterns associated with the risk of type 2 diabetes. Compared with the single dietary component approach, the

patterning approach takes account of the complexity of the dietary intake and the intercorrelations and cumulative effects of various nutrients or foods. Thus, patterns are appropriate measures of overall diet. Furthermore, we are not aware of any other studies that have derived patterns associated with diabetes risk using biomarker information. Analysis by RRR allowed us to use data on the diabetes-related biomarkers CRP, adiponectin, HDL cholesterol and HbA_{1c} to derive a pattern score associated with diabetes risk. Compared with other data reduction techniques (e.g. principal component or factor analysis), RRR has the advantage that it may extract dietary pattern scores by maximising the explained variation in the concentration of biomarkers or other intermediate variables in the pathway that links diet to the disease of interest. Thus, RRR may be a useful tool for the identification of diet–disease associations. However, the method is limited by the fact that it requires prior knowledge of the disease-related variables chosen as responses.

A further advantage associated with our study is the prospective study design. The identification of the study population and assessment of diet before diabetes was diagnosed reduces the possibility of selection and recall bias. Our results should not be severely biased by misclassification of diabetes, since all potential cases were verified by medical records. The remaining misclassification based on non-identified incident or prevalent cases would serve to attenuate the observed risk estimates rather than overestimate the risks. Moreover, the high follow-up rate of 96% minimised the bias due to loss to follow-up.

Nevertheless, our study has several limitations. First, we assessed diet using a food-frequency questionnaire. Measurement errors may be introduced by the under- or over-reporting of the amounts of food groups usually eaten per

Table 3 Plasma concentration of response variables in the study population (*n*=574) according to quintiles of first dietary pattern score

Response variable	Q1	Q2	Q3	Q4	Q5	<i>p</i> value for trend
HbA _{1c} (%)	6.02±0.22	5.34±0.13	5.22±0.13	5.06±0.15	4.78±0.07	<0.0001
HDL cholesterol (mmol/l)	0.90±0.02	0.94±0.02	1.04±0.03	1.02±0.02	1.14±0.02	<0.0001
CRP (µg/ml)	3.07±0.45	3.00±0.35	3.78±0.64	3.07±0.38	2.11±0.28	0.18
Adiponectin (µg/ml)	5.40±0.27	6.14±0.28	7.54±0.46	7.82±0.43	9.55±0.45	<0.0001

Data are means±SD

Q Quintile of dietary pattern score

Table 4 Risk of type 2 diabetes in the study population ($n=574$) according to quintiles of first dietary pattern score

	Q1	Q2	Q3	Q4	Q5	<i>p</i> value for trend
Diabetic subjects/ control subjects	54/60	45/70	37/78	31/84	25/90	
Model 1	1.0 (reference)	0.63 (0.34–1.08)	0.44 (0.25–0.79)	0.30 (0.16–0.56)	0.20 (0.10–0.40)	<0.0001
Model 2	1.0 (reference)	0.60 (0.35–1.05)	0.40 (0.22–0.73)	0.27 (0.14–0.53)	0.18 (0.09–0.37)	<0.0001
Model 3	1.0 (reference)	0.59 (0.30–1.15)	0.51 (0.25–1.05)	0.26 (0.12–0.56)	0.27 (0.13–0.64)	0.0006

Data are numbers or ORs (95% CIs). Model 1: adjusted for age (years) and sex (matching variables). Model 2: adjusted for all variables included in model 1, plus leisure-time activity (h per week), smoking status (current smoker, ex-smoker for <5 years, and ex-smoker for ≥ 5 years or non-smoker), educational level (less than high school education, high school education, or university degree) and total energy intake (kJ per day). Model 3: Adjusted for all variables included in model 2 plus BMI (kg/m^2) and WHR

Q Quintile of dietary pattern score

day. Second, a crucial issue in selecting biomarkers as response variables for RRR is the time of blood sampling. Ideally, all biomarkers should be biomarkers of exposure or disease, or they should refer to the same stage of the pathway from exposure to disease [51]. In our study, the four responses CRP, adiponectin, HDL cholesterol and HbA_{1c} probably reflect different stages in the pathogenesis of type 2 diabetes. HbA_{1c} in particular may refer to a later stage of the pathogenesis of diabetes than the other three biomarkers. However, elevated HbA_{1c} values that reflect higher average glucose concentrations and that are below the threshold for diabetes may refer to a preceding stage of type 2 diabetes. Additionally, the sensitivity analysis showed that a dietary pattern score derived using a biomarker set without HbA_{1c} is still significantly associated with a reduced risk of diabetes (Table 5). Third, although the variations in the selected biomarkers elucidated by the dietary pattern score were small, these translated into relatively large variations when comparing the maximally explainable biomarker variation by all food groups. Fourth, a cohort

approach could be more attractive than the nested case–control approach used in this study, because this design produces more realistic results. A cohort study would probably produce smaller risk estimates across quintiles than those observed in this study. However, a study based on a nested case–control design with available biomarker values for all cases and matched controls seemed the most appealing approach for our aims. Fifth, patterns may be difficult to reproduce in other studies, especially in study populations with different dietary habits.

In summary, we have identified a dietary pattern score that is directly associated with the intake of fresh fruit and inversely associated with the intake of high-caloric soft drinks, beer, red meat, poultry, processed meat, legumes and bread (excluding wholegrain bread). A higher pattern score was cross-sectionally related to higher concentrations of HDL cholesterol and adiponectin, and lower concentrations of HbA_{1c} and CRP. It was also prospectively associated with a substantially lower risk of diabetes. The observed association between the pattern and diabetes risk

Table 5 Risk of type 2 diabetes in the study population ($n=574$) according to quintiles of first dietary pattern scores derived from different combinations of response variables

Set of response variables	Q1	Q2	Q3	Q4	Q5	<i>p</i> value for trend
HDL-c, HbA _{1c} , CRP	1.0 (reference)	0.82 (0.42–1.60)	0.44 (0.22–0.89)	0.30 (0.14–0.66)	0.23 (0.10–0.55)	<0.0001
HDL-c, CRP, adiponectin	1.0 (reference)	0.72 (0.36–1.43)	0.64 (0.31–1.31)	0.35 (0.16–0.78)	0.54 (0.25–1.19)	0.047
HbA _{1c} , CRP, adiponectin	1.0 (reference)	0.60 (0.28–1.15)	0.34 (0.16–0.69)	0.45 (0.22–0.94)	0.28 (0.13–0.64)	0.002
HDL-c, HbA _{1c} , adiponectin	1.0 (reference)	0.54 (0.27–1.08)	0.43 (0.20–0.93)	0.29 (0.13–0.62)	0.27 (0.11–0.67)	0.001
HDL-c, HbA _{1c} , CRP, adiponectin, IL-6	1.0 (reference)	0.85 (0.44–1.63)	0.44 (0.22–0.88)	0.33 (0.15–0.74)	0.31 (0.14–0.72)	0.0007
HDL-c, HbA _{1c} , CRP, adiponectin, TNF- α	1.0 (reference)	0.63 (0.33–1.23)	0.48 (0.24–0.98)	0.28 (0.13–0.61)	0.26 (0.11–0.63)	0.0006
HDL-c, HbA _{1c} , CRP, adiponectin, IL-6, TNF- α	1.0 (reference)	0.91 (0.47–1.77)	0.42 (0.21–0.86)	0.36 (0.16–0.80)	0.35 (0.15–0.80)	0.001

Data are ORs (95% CIs) adjusted for age (years) and sex (matching variables), leisure-time activity (h per week), smoking status (current smoker, ex-smoker for <5 years, and ex-smoker for ≥ 5 years or non-smoker), educational level (less than high school education, high school education, or university degree), total energy intake (kJ per day), BMI (kg/m^2) and WHR

HDL-c, HDL cholesterol

Q Quintile of dietary pattern score

was independent of BMI, WHR, smoking status, leisure-time activity, energy intake, educational level, age and sex. The results suggest that the identified dietary pattern, which was linked to a specific biomarker profile, may be a predictor of the subsequent risk of type 2 diabetes.

Acknowledgements We thank all of the study participants. We also thank E. Kohlsdorf and W. Bernigau for data management; K. Sprengel for excellent technical support; the Department of Clinical Biochemistry, Greifswald University (Germany), under the supervision of H.-J. Rose for HbA_{1c} analyses; and T. Pischon for critical comments on the manuscript.

The recruitment phase of the EPIC—Potsdam Study was supported by the Federal Ministry of Science, Germany (01 EA 9401) and the European Union (SOC 95201408 05F02). The EPIC—Potsdam Study is now supported by German Cancer Aid (70-2488-Ha I) and the European Community (SOC 98200769 05F02). J. Spranger is supported by grants from the German Diabetes Association, the Eli-Lilly International Foundation and the Federal Ministry of Education and Research. None of the authors are aware of any duality of interest.

References

- King H, Aubert RE, Herman WH (1998) Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections. *Diabetes Care* 21:1414–1431
- Centers for Disease Control and Prevention (2004) The burden of heart disease, stroke, cancer, and diabetes, United States. In: *The burden of chronic diseases and their risk factors. National and state perspectives 2004*. US Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, p 29
- Hu FB, van Dam RM, Liu S (2001) Diet and risk of Type II diabetes: the role of types of fat and carbohydrate. *Diabetologia* 44:805–817
- Salmeron J, Manson JE, Stampfer MJ, Colditz GA, Wing AL, Willett WC (1997) Dietary fiber, glycemic load, and risk of non-insulin-dependent diabetes mellitus in women. *JAMA* 277:472–477
- Lopez-Ridaura R, Willett WC, Rimm EB et al (2004) Magnesium intake and risk of type 2 diabetes in men and women. *Diabetes Care* 27:134–140
- Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 13:3–9
- Williams DE, Prevost AT, Whiclow MJ, Cox BD, Day NE, Wareham NJ (2000) A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *Br J Nutr* 83:257–266
- Gittelsohn J, Wolever TM, Harris SB, Harris-Giraldo R, Hanley AJ, Zinman B (1998) Specific patterns of food consumption and preparation are associated with diabetes and obesity in a Native Canadian community. *J Nutr* 128:541–547
- van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB (2002) Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. *Ann Intern Med* 136:201–209
- Newby PK, Tucker KL (2004) Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* 62:177–203
- Hoffmann K, Schulze MB, Schienkiewitz A, Nothlings U, Boeing H (2004) Application of a new statistical method to derive dietary patterns in nutritional epidemiology. *Am J Epidemiol* 159:935–944
- Hoffmann K, Zyriax BC, Boeing H, Windler E (2004) A dietary pattern derived to explain biomarker variation is strongly associated with the risk of coronary artery disease. *Am J Clin Nutr* 80:633–640
- Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J (1997) Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. *European Prospective Investigation into Cancer and Nutrition*. *Int J Epidemiol* 26(Suppl 1):S59–S70
- Kroke A, Klipstein-Grobusch K, Voss S et al (1999) Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* 70:439–447
- Schulze MB, Hoffmann K, Kroke A, Boeing H (2003) Risk of hypertension among women in the EPIC—Potsdam study: comparison of relative risk estimates for exploratory and hypothesis-oriented dietary patterns. *Am J Epidemiol* 158:365–373
- Klipstein-Grobusch K, Georg T, Boeing H (1997) Interviewer variability in anthropometric measurements and estimates of body composition. *Int J Epidemiol* 26(Suppl 1):S174–S180
- Koenig RJ, Peterson CM, Jones RL, Saudek C, Lehrman M, Cerami A (1976) Correlation of glucose regulation and hemoglobin A_{1c} in diabetes mellitus. *N Engl J Med* 295:417–420
- Little RR, Rohlfing CL, Wiedmeyer HM, Myers GL, Sacks DB, Goldstein DE (2001) The national glycohemoglobin standardization program: a five-year progress report. *Clin Chem* 47:1985–1992
- Pratley RE, Weyer C (2002) Progression from IGT to type 2 diabetes mellitus: the central role of impaired early insulin secretion. *Curr Diab Rep* 2:242–248
- Edelstein SL, Knowler WC, Bain RP et al (1997) Predictors of progression from impaired glucose tolerance to NIDDM: an analysis of six prospective studies. *Diabetes* 46:701–710
- de Vegt F, Dekker JM, Jager A et al (2001) Relation of impaired fasting and postload glucose with incident type 2 diabetes in a Dutch population: the Hoorn Study. *JAMA* 285:2109–2113
- Ginsberg HN (1996) Diabetic dyslipidemia: basic mechanisms underlying the common hypertriglyceridemia and low HDL cholesterol levels. *Diabetes* 45(Suppl 3):S27–S30
- Spranger J, Kroke A, Mohlig M et al (2003) Inflammatory cytokines and the risk to develop type 2 diabetes: results of the prospective population-based European Prospective Investigation into Cancer and Nutrition (EPIC)—Potsdam study. *Diabetes* 52:812–817
- Pickup JC, Crook MA (1998) Is type II diabetes mellitus a disease of the innate immune system? *Diabetologia* 41:1241–1248
- Hu FB, Meigs JB, Li TY, Rifai N, Manson JE (2004) Inflammatory markers and risk of developing type 2 diabetes in women. *Diabetes* 53:693–700
- Freeman DJ, Norrie J, Caslake MJ et al (2002) C-reactive protein is an independent predictor of risk for the development of diabetes in the West of Scotland Coronary prevention study. *Diabetes* 51:1596–1600
- Barzilay JI, Abraham L, Heckbert SR et al (2001) The relation of markers of inflammation to the development of glucose disorders in the elderly: the Cardiovascular Health study. *Diabetes* 50:2384–2389
- Ouchi N, Kihara S, Arita Y et al (2000) Adiponectin, an adipocyte-derived plasma protein, inhibits endothelial NF- κ B signaling through a cAMP-dependent pathway. *Circulation* 102:1296–1301
- Stefan N, Vozarova B, Funahashi T et al (2002) Plasma adiponectin concentration is associated with skeletal muscle insulin receptor tyrosine phosphorylation, and low plasma concentration precedes a decrease in whole-body insulin sensitivity in humans. *Diabetes* 51:1884–1888
- Berg AH, Combs TP, Du X, Brownlee M, Scherer PE (2001) The adipocyte-secreted protein Acrp30 enhances hepatic insulin action. *Nat Med* 7:947–953
- Spranger J, Kroke A, Mohlig M et al (2003) Adiponectin and protection against type 2 diabetes mellitus. *Lancet* 361:226–228

33. Kroenke CH, Chu NF, Rifai N et al (2003) A cross-sectional study of alcohol consumption patterns and biologic markers of glycemic control among 459 women. *Diabetes Care* 26:1971–1978
34. Rimm EB, Williams P, Fosher K, Criqui M, Stampfer MJ (1999) Moderate alcohol intake and lower risk of coronary heart disease: meta-analysis of effects on lipids and haemostatic factors. *BMJ* 319:1523–1528
35. Imhof A, Froehlich M, Brenner H, Boeing H, Pepys MB, Koenig W (2001) Effect of alcohol consumption on systemic markers of inflammation. *Lancet* 357:763–767
36. Sierksma A, Patel H, Ouchi N et al (2004) Effect of moderate alcohol consumption on adiponectin, tumor necrosis factor- α , and insulin sensitivity. *Diabetes Care* 27:184–189
37. Harding AH, Sargeant LA, Welch A et al (2001) Fat consumption and HbA(1c) levels: the EPIC—Norfolk study. *Diabetes Care* 24:1911–1916
38. Ascherio A, Katan MB, Zock PL, Stampfer MJ, Willett WC (1999) Trans fatty acids and coronary heart disease. *N Engl J Med* 340:1994–1998
39. Ajani UA, Ford ES, Mokdad AH (2004) Dietary fiber and C-reactive protein: findings from national health and nutrition examination survey data. *J Nutr* 134:1181–1185
40. Brand-Miller J, Hayne S, Petocz P, Colagiuri S (2003) Low-glycemic index diets in the management of diabetes: a meta-analysis of randomized controlled trials. *Diabetes Care* 26:2261–2267
41. Hallez S, Derounane A (1982) Nouvelle methode de traitement de séries de données tronquées dans l' étude de la pollutions atmospherique [article in French]. *Sci Total Environ* 22:115–123
42. van Dam RM, Willett WC, Rimm EB, Stampfer MJ, Hu FB (2002) Dietary fat and meat intake in relation to risk of type 2 diabetes in men. *Diabetes Care* 25:417–424
43. Schulze MB, Manson JE, Willett WC, Hu FB (2003) Processed meat intake and incidence of Type 2 diabetes in younger and middle-aged women. *Diabetologia* 46:1465–1473
44. Wirfalt E, Hedblad B, Gullberg B et al (2001) Food patterns and components of the metabolic syndrome in men and women: a cross-sectional study within the malmo diet and cancer cohort. *Am J Epidemiol* 154:1150–1159
45. Martin-Villa MC, Vidal-Valverde C, Rojas-Hidalgo E (1981) Soluble sugars in soft drinks. *Am J Clin Nutr* 34:2151–2153
46. Janssens JP, Shapira N, Debeuf P et al (1999) Effects of soft drink and table beer consumption on insulin response in normal teenagers and carbohydrate drink in youngsters. *Eur J Cancer Prev* 8:289–295
47. Selby JV, Newman B, King MC, Friedman GD (1987) Environmental and behavioral determinants of fasting plasma glucose in women. A matched co-twin analysis. *Am J Epidemiol* 125:979–988
48. Jose Gerard M, Klatsky AL, Siegelau AB, Friedman GD, Feldman R (1977) Serum glucose levels and alcohol-consumption habits in a large population. *Diabetes* 26:780–785
49. Marlett JA, McBurney MI, Slavin JL (2002) Position of the American Dietetic Association: health implications of dietary fiber. *J Am Diet Assoc* 102:993–1000
50. Jenkins DJ, Kendall CW, Marchie A et al (2003) Type 2 diabetes and the vegetarian diet. *Am J Clin Nutr* 78:610S–616S
51. Branca F, Hanley AB, Pool-Zobel B, Verhagen H (2001) Biomarkers in disease and health. *Br J Nutr* 86(Suppl 1):S55–S92