

# Dietary pattern, inflammation, and incidence of type 2 diabetes in women<sup>1-3</sup>

Matthias B Schulze, Kurt Hoffmann, JoAnn E Manson, Walter C Willett, James B Meigs, Cornelia Weikert, Christin Heidemann, Graham A Colditz, and Frank B Hu

## ABSTRACT

**Background:** Inflammation is considered a key mechanism leading to type 2 diabetes, but dietary exposures that lead to inflammation and diabetes are largely unknown.

**Objective:** Our objective was to investigate the relation between a dietary pattern associated with biomarkers of inflammation and the incidence of type 2 diabetes.

**Design:** We conducted a nested case-control study of 656 cases of type 2 diabetes and 694 controls among women in the Nurses' Health Study and 2 prospective cohort studies of 35 340 women in the Nurses' Health Study and 89 311 women in the Nurses' Health Study II who were followed for incident diabetes.

**Results:** Through the use of reduced rank regression, we identified a dietary pattern that was strongly related to inflammatory markers in the nested case-control study. This pattern, which was high in sugar-sweetened soft drinks, refined grains, diet soft drinks, and processed meat but low in wine, coffee, cruciferous vegetables, and yellow vegetables, was associated with an increased risk of diabetes (multivariate-adjusted odds ratio comparing extreme quintiles: 3.09; 95% CI: 1.99, 4.79). We identified 1517 incident cases of confirmed type 2 diabetes in the Nurses' Health Study (458 991 person-years) and 724 incident cases in the Nurses' Health Study II (701 155 person-years). After adjustment for body mass index and other potential lifestyle confounders, the relative risks comparing extreme quintiles of the pattern were 2.56 (95% CI: 2.10, 3.12; *P* for trend < 0.001) in the Nurses' Health Study and 2.93 (95% CI: 2.18, 3.92; *P* for trend < 0.001) in the Nurses' Health Study II.

**Conclusion:** The dietary pattern identified may increase chronic inflammation and raise the risk of developing type 2 diabetes. *Am J Clin Nutr* 2005;82:675–84.

**KEY WORDS** Body mass index, diet pattern, incidence, inflammation, non-insulin-dependent diabetes mellitus, type 2 diabetes, prospective studies, questionnaires, reduced rank regression, risk factors

## INTRODUCTION

Accumulating evidence supports the hypothesis that chronic low-grade inflammation and activation of the innate immune system are closely involved in the pathogenesis of type 2 diabetes. Since this hypothesis was first proposed in 1997 (1), many studies have shown that circulating markers of inflammation are strong predictors of the development of type 2 diabetes (2).

Dietary factors may have effects on inflammation and endothelial function independent of smoking, hypercholesterolemia, hyperhomocysteinemia, and hypertension (3). Several dietary factors have also been identified to be associated with diabetes risk, including types of carbohydrate and fat, magnesium and cereal fiber intakes, and consumption of coffee, meats, and sugar-sweetened beverages (4). Dietary patterns, which reflect different combinations of food intake, have also been associated with diabetes risk (5–8) and with markers of inflammation (9, 10). However, previous attempts to derive food patterns were largely determined by variation in food choices among individuals and could not incorporate biomarkers into the analyses on the basis of a priori hypotheses. We therefore derived a dietary pattern that was strongly associated with markers of inflammation and endothelial dysfunction and evaluated the association between this pattern and the risk of developing type 2 diabetes in 2 large cohort studies of women.

## SUBJECTS AND METHODS

### Study populations

The Nurses' Health Study (NHS) cohort was established in 1976 when 121 700 female registered nurses aged 30–55 y and residing in 11 US states responded to mailed questionnaires regarding their medical history and health practices. The respondents reflect the racial composition of women trained as registered nurses at that time; 97% were white (11). Since 1976,

<sup>1</sup> From the Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany (MBS, KH, CW, and CH); the Division of Preventive Medicine (JEM) and the Channing Laboratory (JEM, WCW, GAC, and FBH), Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; the Departments of Epidemiology (JEM, WCW, GAC, and FBH) and Nutrition (WCW and FBH), Harvard School of Public Health, Boston, MA; and the General Medicine Division, Massachusetts General Hospital, Boston, MA (JBM).

<sup>2</sup> Supported by research grants (CA50385, CA87969, HL60712, and DK58845) from the National Institutes of Health. MBS is supported by a grant from the Deutsche Krebshilfe. JBM is supported in part by an American Diabetes Association Career Development Award. FBH is the recipient of an American Heart Association Established Investigator Award.

<sup>3</sup> Reprints not available. Address correspondence to MB Schulze, German Institute of Human Nutrition Potsdam-Rehbruecke, Department of Epidemiology, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany. E-mail: mschulze@mail.dife.de.

Received March 2, 2005.

Accepted for publication May 13, 2005.

questionnaires have been administered biennially to update health information and to identify new cases of disease. During 1989 through 1990, 32 826 women free of diagnosed diabetes, ischemic heart disease, stroke, or cancer provided blood samples. By 2000, 737 of these women had developed definite diabetes. Control women providing baseline blood samples were matched to diabetes cases by year of birth, date of blood draw, race, and fasting status (at least 8 h overnight) at the time of the blood draw. In addition, to improve statistical control for obesity at the upper extreme of the distribution, control subjects were also matched by body mass index (BMI) for case subjects in the top 10% of the BMI distribution, giving a sample of 785 control women. A total of 656 cases and 694 controls completed a dietary questionnaire in 1986, 1990, or both and had complete information on body weight, height, physical activity, smoking, and biomarkers.

Among those NHS participants who did not provide blood, 51 895 women completed a dietary questionnaire in 1984. After the exclusion of women with a history of diabetes, cancer (except nonmelanoma skin cancer), or cardiovascular disease; with implausible energy intakes (eg, <500 or >3500 kcal/d); or without information on physical activity or body weight, 35 340 women remained for cohort analyses separate from the nested case-control analysis.

The Nurses' Health Study II (NHS-II) is a prospective cohort study of 116 671 female US nurses. Participants were aged 24–44 y at the study start in 1989. As with the NHS cohort, the NHS-II cohort is followed up through the use of biennial mailed questionnaires. The follow-up rate exceeds 90% for every 2-y period, and we estimate nearly complete (98%) ascertainment of mortality. For the analyses presented here, women were excluded from the baseline population if they did not complete a dietary questionnaire in 1991; if the reported dietary intake was implausible with regard to total energy intake; if they had a history of diabetes, cancer (except nonmelanoma skin cancer), or cardiovascular disease reported on either the 1989 or 1991 questionnaire; or if they had not provided data on body weight and physical activity in 1991. These exclusions left a total of 89 311 women for the analyses.

Both studies were approved by the Human Research Committees at the Harvard School of Public Health and the Brigham and Women's Hospital. Completion of the self-administered questionnaire was considered to imply informed consent.

### Dietary assessment

Dietary intake information was collected through semiquantitative food-frequency questionnaires (FFQs). Women were asked how often they had consumed a commonly used unit or portion size of each food on average during the previous year, with 9 possible frequency responses ranging from "never" to "more than 6 times a day." For this analysis, we used information collected through the 1984, 1986, 1990, and 1994 FFQs in the NHS and the 1991 and 1995 FFQs in the NHS-II, respectively. Foods were classified into 39 food groups on the basis of nutrient profiles and culinary usage (12). Nutrient intakes were computed by multiplying the frequency response by the nutrient content of the specified portion sizes. Values for nutrients were derived from US Department of Agriculture sources (13) supplemented with information from manufacturers and biochemical analyses

at Harvard School of Public Health. Intakes of cereal fiber, magnesium, and caffeine and the glycemic index were energy adjusted by using the residuals method (14). Intakes of carbohydrates and *trans* fatty acids were expressed as nutrient density (% of total energy intake) (14). The validity and reliability of the FFQ used in the NHS have been described elsewhere (15, 16). Briefly, the mean correlation coefficient between frequencies of intake of 55 foods from 2 FFQs administered 12 mo apart was 0.57, and the mean corrected correlation coefficient between dietary records and a subsequent FFQ was 0.66 in the NHS (15).

### Ascertainment of type 2 diabetes

Women reporting a new diagnosis of diabetes on any of the biennial questionnaires were sent supplementary questionnaires asking about diagnostic tests and treatment of their diabetes and history of ketoacidosis or ketosis to distinguish between type 1 and type 2 diabetes. In accordance with the criteria of the National Diabetes Data Group (17), confirmation of diabetes required at least one of the following: 1) an elevated plasma glucose concentration [fasting plasma glucose  $\geq$  7.8 mmol/L (140 mg/dL), random plasma glucose  $\geq$  11.1 mmol/L (200 mg/dL), or plasma glucose  $\geq$  11.1 mmol/L (200 mg/dL) after  $\geq$  2 h during an oral-glucose-tolerance test] plus at least one classic symptom (excessive thirst, polyuria, weight loss, or hunger); 2) no symptoms, but at least 2 elevated plasma glucose concentrations (by the above criteria) on different occasions; or 3) treatment with hypoglycemic medication (insulin or oral hypoglycemic agent). The diagnostic criteria for type 2 diabetes were changed in 1997 (18) such that lower fasting glucose levels ( $\geq$  7 mmol/L, or 126 mg/dL) would now be considered diagnostic. Thus, we used the American Diabetes Association criteria for diagnosis of diabetes after 1998. In a substudy of the NHS, 98% of the self-reported diabetes cases documented by the same supplementary questionnaire were confirmed by medical record review (19, 20).

### Assessment of nondietary exposures

Information on age, weight, smoking status, contraceptive use, and postmenopausal hormone therapy was collected through biennial questionnaires. We calculated BMI as the ratio of weight (in kg) to squared height (in  $m^2$ ), the latter being assessed at baseline only. Self-reports of body weight have been highly correlated with technician-measured weights ( $r = 0.96$ ) in the NHS (21). Waist and hip circumferences were self-reported in 1986 and 1996 in the NHS and in 1993 in the NHS-II. Physical activity, assessed in 1986, 1988, 1992, 1994, and 1996 in the NHS and in 1991 and 1997 in the NHS-II, was computed as metabolic equivalents (MET) per week by using the duration per week of various forms of exercise, with each activity weighted by its intensity level. Physical activity reported on the questionnaire was highly correlated with activity recorded in diaries or by 24-h recall in the NHS-II (0.79 and 0.62) (22).

### Laboratory procedures

Women in the NHS willing to provide blood specimens were sent instructions and a phlebotomy kit (including sodium heparin blood tubes, needles, and a tourniquet). Blood specimens were returned by overnight mail in a frozen water bottle. On arrival, the samples were centrifuged to separate plasma from the buffy coat and red cells and were frozen in liquid nitrogen until analyzed.

Ninety-seven percent of samples arrived within 26 h of phlebotomy.

Frozen plasma samples from case and control subjects were selected for simultaneous analysis. C-reactive protein (CRP) concentrations were measured by use of a high-sensitivity latex-enhanced immunonephelometric assay on a BN II analyzer (Dade Behring, Newark, DE). Interleukin 6 (IL-6) was measured by a quantitative sandwich enzyme immunoassay technique (Quantikine HS Immunoassay kit; R&D Systems, Minneapolis), and plasma concentrations of soluble fractions of tumor necrosis factor  $\alpha$  receptor 2 (sTNFR2) were measured by use of an enzyme-linked immunosorbent assay (ELISA) kit using immobilized monoclonal antibody to human sTNFR2 (Genzyme, Cambridge, MA). Concentrations of E-selectin, soluble intracellular cell adhesion molecule 1 (sICAM-1), and soluble vascular cell adhesion molecule 1 (sVCAM-1) were measured by commercial ELISA (R & D Systems). The interassay CVs for each analyte were as follows: 3.4–3.8% for CRP, 5.8–8.2% for IL-6, 6.2% for sTNFR2, 6.4–6.6% for E-selectin, 6.1–10.1% for sICAM-1, and 8.5–10.2% for sVCAM-1.

### Statistical analysis

In the nested case-control sample of the NHS, we calculated the mean intake from the 1986 and 1990 FFQs for the 39 food groups to reduce within-subject variation and best represent long-term diet (23). We subsequently applied reduced rank regression (RRR) to derive a dietary pattern predictive of diabetes risk. RRR identifies linear functions of predictors (eg, food groups) that explain as much response (eg, biomarker) variation as possible. RRR can be interpreted as a principal component analysis applied to responses and a subsequent linear regression of principal components on predictors, although it is somewhat more sophisticated and efficient than this 2-step procedure. Thus, we calculated linear functions of food group intake (dietary patterns) that explain as much variation in inflammatory biomarkers as possible. The first factor obtained by RRR was retained for subsequent analyses because it explains the largest amount of variation among the biomarkers. A more detailed description of the method, including the SAS code and its application in nutritional epidemiology, can be found elsewhere (24). Results of the RRR analysis based on food intakes expressed as energy densities were similar and are not reported here.

Pearson correlation coefficients were used to evaluate associations between the derived pattern and inflammatory biomarkers in the nested case-control sample of the NHS. We divided the distributions of the derived dietary pattern into quintiles based on the control subjects and calculated the geometric mean biomarker concentrations across pattern quintiles with adjustment for anthropometric and lifestyle characteristics. We used unconditional logistic regression models to estimate the odds ratio (OR) of diabetes in each quintile by using the lowest quintile as the reference category and to estimate the significance of trend in ORs across increasing biomarker quintiles. Multivariate models were adjusted for age (5-y groups), BMI (<21.0, 21.0–22.9, 23.0–24.9, 25.0–26.9, 27.0–28.9, 29.0–30.9, 31.0–32.9, 33.0–34.9, and  $\geq 35.0$ ), physical activity (MET score in quintiles), family history of diabetes in a first-degree relative (yes or no), smoking (missing, never, past, or current), postmenopausal hormone use (missing, never, or ever), total energy intake (quintiles), and fasting status. The results of conditional logistic regression with conditioning on matching factors, such as age,

fasting status, and race, were similar to the unconditional regression models and are not reported here.

To cross-validate the results from the nested case-control study, we first identified the important component foods of the RRR pattern by stepwise linear regression with a significance level of  $P < 0.05$  for entry into and staying in the model with the biomarker response score as the dependent variable and food groups as independent variables. A diet pattern score was then calculated among women in the NHS who did not provide blood as well as in the NHS-II as the linear combination of these standardized component food items (25). We estimated the relative risk (RR) for each quintile of pattern score compared with the lowest quintile by using Cox proportional hazards analysis stratified by 5-y age categories and 2-y intervals. Duration of follow-up was calculated as the interval between the return of the 1984 questionnaire (NHS) and the 1991 questionnaire (NHS-II), respectively, and the diagnosis of diabetes (type 1 or type 2), death, or June 1, 1998 (NHS) or 1999 (NHS-II), respectively. We cumulatively updated food intake information with subsequent questionnaires, except for those individuals who reported a diagnosis of cancer (except nonmelanoma skin cancer) or cardiovascular disease during follow-up, because changes in diet after development of these conditions may confound the relation between dietary intake and diabetes (23).

We used information on covariates obtained from the baseline or subsequent questionnaires in multivariate analyses, including BMI (9 categories), physical activity (MET score in quintiles), family history of diabetes in a first-degree relative (yes or no), smoking (missing, never, past, or current), postmenopausal hormone use (missing, never, or ever), and total energy intake (quintiles). We further adjusted for nutrient intake, in particular, magnesium intake (quintiles), caffeine intake (quintiles), cereal fiber intake (quintiles), *trans* fatty acid intake (quintiles), alcohol intake (0, 0.1–4.9, 5.0–9.9, and  $\geq 10$  g/d), the glycemic index (quintiles), and the ratio of polyunsaturated to saturated fat (quintiles) to evaluate whether these nutrients may mediate the dietary pattern–disease association. Nondietary covariates were updated during follow-up by using the most recent data for each 2-y follow-up interval. The significance of linear trends across quintiles of the pattern score was tested by assigning each participant the median value for the quintile and modeling this value as a continuous variable. We also tested for effect modification by BMI, physical activity, and family history of diabetes by performing analyses stratified by these variables and by evaluating interaction terms. All statistical analyses were performed by using SAS statistical software (version 8.0; SAS Institute Inc, Cary, NC).

### RESULTS

Women in the NHS nested case-control population who subsequently developed type 2 diabetes were significantly less physically active ( $\bar{x} \pm \text{SD}$ :  $12.0 \pm 14.8$  METs) and had greater BMIs ( $\bar{x} \pm \text{SD}$ :  $30.2 \pm 5.6$ ) at baseline than did women who remained diabetes free ( $15.8 \pm 28.1$  METs; BMI:  $26.1 \pm 5.1$ ). The age of the participants did not differ significantly between the cases and the controls ( $\bar{x} \pm \text{SD}$ :  $56.2 \pm 6.9$  y; range: 43–69 y). In addition, although the proportion of ex-smokers (41.1% and 42.4%) and current smokers (13.4% and 13.2%) did not differ significantly between the cases and the controls, women who developed diabetes were more likely to have had a family relative with a history of diabetes (47.6% compared with 20.8%).

**TABLE 1**Pearson correlations between diet pattern score and associated food groups and biomarkers of inflammation, Nurses' Health Study<sup>1</sup>

| Pattern and food groups       | Pearson correlation with diet pattern score | Pearson correlation with biomarker <sup>2</sup> |                    |                    |                    |                    |                    |
|-------------------------------|---------------------------------------------|-------------------------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                               |                                             | sTNFR2                                          | IL-6               | CRP                | E-selectin         | sICAM-1            | sVCAM-1            |
| Diet pattern score            | 1.00                                        | 0.12 <sup>3</sup>                               | 0.21 <sup>3</sup>  | 0.23 <sup>3</sup>  | 0.26 <sup>3</sup>  | 0.18 <sup>3</sup>  | 0.14 <sup>3</sup>  |
| Food groups <sup>4</sup>      |                                             |                                                 |                    |                    |                    |                    |                    |
| Positive associations         |                                             |                                                 |                    |                    |                    |                    |                    |
| Sugar-sweetened soft drinks   | 0.47 <sup>3</sup>                           | 0.06 <sup>5</sup>                               | 0.10 <sup>3</sup>  | 0.12 <sup>3</sup>  | 0.12 <sup>3</sup>  | 0.06 <sup>5</sup>  | 0.08 <sup>6</sup>  |
| Refined grains                | 0.46 <sup>3</sup>                           | 0.05                                            | 0.07 <sup>6</sup>  | 0.10 <sup>3</sup>  | 0.14 <sup>3</sup>  | 0.09 <sup>3</sup>  | 0.06 <sup>5</sup>  |
| Processed meat                | 0.39 <sup>3</sup>                           | -0.02                                           | 0.09 <sup>3</sup>  | 0.13 <sup>3</sup>  | 0.12 <sup>3</sup>  | 0.05               | 0.03               |
| Diet soft drinks              | 0.26 <sup>3</sup>                           | 0.03                                            | 0.05               | 0.11 <sup>3</sup>  | 0.07 <sup>5</sup>  | 0.04               | -0.02              |
| Other vegetables <sup>7</sup> | 0.11 <sup>3</sup>                           | <0.01                                           | 0.02               | 0.02               | 0.02               | 0.05               | 0.03               |
| Negative associations         |                                             |                                                 |                    |                    |                    |                    |                    |
| Wine                          | -0.43 <sup>3</sup>                          | -0.04                                           | -0.08 <sup>6</sup> | -0.09 <sup>3</sup> | -0.12 <sup>3</sup> | -0.07 <sup>5</sup> | -0.09 <sup>6</sup> |
| Coffee                        | -0.29 <sup>3</sup>                          | -0.08 <sup>6</sup>                              | -0.03              | -0.09 <sup>6</sup> | -0.07 <sup>6</sup> | -0.02              | -0.06 <sup>5</sup> |
| Cruciferous vegetables        | -0.21 <sup>3</sup>                          | -0.06 <sup>5</sup>                              | -0.03              | <0.01              | -0.08 <sup>6</sup> | -0.05              | -0.05 <sup>5</sup> |
| Yellow vegetables             | -0.21 <sup>3</sup>                          | -0.05 <sup>5</sup>                              | -0.06 <sup>5</sup> | -0.02              | -0.06 <sup>5</sup> | -0.03              | -0.01              |

<sup>1</sup> sTNFR2, soluble tumor necrosis factor  $\alpha$  receptor 2; IL-6, interleukin 6; CRP, C-reactive protein; sICAM-1, soluble intracellular cell adhesion molecule 1; sVCAM-1, soluble vascular cell adhesion molecule 1.

<sup>2</sup> Biomarkers were log transformed.

<sup>3</sup>  $P < 0.001$ .

<sup>4</sup> Food groups were identified through the use of stepwise regression with  $P < 0.05$  for inclusion and exclusion and by modeling the biomarker response score as the dependent variable.

<sup>5</sup>  $P < 0.05$ .

<sup>6</sup>  $P < 0.01$ .

<sup>7</sup> Vegetables other than green leafy vegetables, yellow vegetables, cruciferous vegetables, tomatoes, and legumes.

We identified a dietary pattern with the RRR method that was positively correlated with all inflammatory biomarkers (**Table 1**). Correlation coefficients ranged from 0.12 for sTNFR2 to 0.26 for E-selectin. The pattern represented a diet relatively high in sugar-sweetened soft drinks, refined grains, diet soft drinks, processed meat, and "other vegetables" (other than yellow, cruciferous, and green-leafy vegetables, tomatoes, and legumes) but low in wine, coffee, cruciferous vegetables, and yellow vegetables. These food groups, with the exception of other vegetables, were significantly correlated with at least one inflammatory marker.

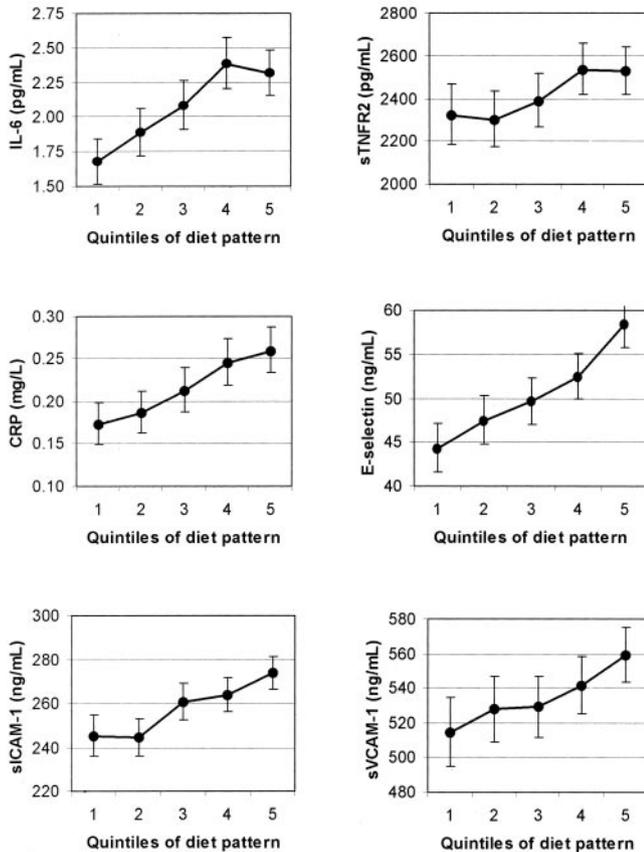
Because anthropometric and lifestyle characteristics may explain associations between the pattern and inflammatory biomarkers, we subsequently calculated the geometric mean biomarker concentrations across quintiles of the pattern with adjustment for BMI and other characteristics (**Figure 1**). The pattern remained strongly associated with all inflammatory markers. Median biomarker concentrations increased from quintile 1 to quintile 5 by 206 pg/mL (8.9%) for sTNFR2, 0.64 pg/mL (38.4%) for IL-6, 0.09 mg/L (50.5%) for CRP, 14.1 ng/mL (32.0%) for E-selectin, 28.7 ng/mL (11.7%) for sICAM-1, and 44.5 ng/mL (8.7%) for sVCAM-1. These results remained virtually unchanged after additional adjustment for a history of hypertension (data not shown).

The pattern was strongly associated with diabetes risk in the nested case-control study (**Table 2**). The age-adjusted OR comparing extreme quintiles was 4.70 (95% CI: 3.20, 6.90). Adjustment for BMI attenuated this association (OR: 3.09), but further adjustment for physical activity, smoking, family history of diabetes, fasting status, and energy intake had no effect (multivariate-adjusted OR: 3.09; 95% CI: 1.99, 4.79;  $P$  for trend  $< 0.001$ ). In addition, further adjustment for a history of hypertension did not alter these results (data not shown).

Because body fat may confound associations between food intake and inflammatory markers, in a separate analysis we adjusted all biomarker values for BMI before their use as responses in the RRR. The dietary pattern remained similar with sugar-sweetened beverages and refined grains being positively associated with the pattern and wine, coffee, and cruciferous vegetables being negatively associated. However, diet soft drinks were no longer an important component of the pattern. Associations between the pattern and diabetes risk were weaker than in the analysis with unadjusted biomarkers, but the pattern remained significantly associated with an increased risk (multivariate-adjusted RR for extreme quintiles: 1.76; 95% CI: 1.18, 2.64).

To validate the results in the 2 independent cohorts, we first identified food groups significantly explaining variation in biomarker levels in the nested case-control study by stepwise linear regression (positive association: sugar-sweetened soft drinks, refined grains, diet soft drinks, processed meat, and other vegetables; negative association: wine, coffee, cruciferous vegetables, and yellow vegetables). We next calculated pattern scores among participants in the NHS who did not provide blood and in the NHS-II cohort on the basis of these food groups. The simplified diet pattern score had a correlation coefficient of 0.75 with the original RRR pattern score in the nested case-control study in the NHS and thus reflected a dietary pattern similar to the one identified with the RRR method.

In both the NHS and the NHS-II, the pattern score was associated with increasing BMI and lower age and physical activity at baseline in 1984 and 1991, respectively (**Table 3**). Also, women with a higher pattern score were more likely to have a family history of diabetes. The dietary pattern represented a diet relatively high in *trans* fat and with a high glycemic index but low in alcohol, magnesium, caffeine, and cereal fiber and with a low ratio of polyunsaturated to saturated fat.



**FIGURE 1.** Geometric mean concentrations and 95% CIs of interleukin 6 (IL-6), soluble tumor necrosis factor  $\alpha$  receptor 2 (sTNFR2), C-reactive protein (CRP), E-selectin, soluble intracellular cell adhesion molecule 1 (sICAM-1), and soluble vascular cell adhesion molecule 1 (sVCAM-1) by quintiles of diet pattern score adjusted for age, BMI (9 categories), physical activity (quintiles), family history of diabetes, smoking (never, past, current, or missing), postmenopausal hormone use (never, ever, or missing), energy intake (quintiles), and fasting status. The comparison between quintile 5 and quintile 1 was significant for all biomarkers,  $P < 0.05$ . Quintile cutoffs were based on distributions in controls.

During 458 991 person-years of follow-up, we documented 1517 new cases of type 2 diabetes among women in the NHS who did not provide blood. The pattern score was associated with an increased risk of diabetes, but this association was substantially attenuated after adjustment for BMI (Table 4). The multivariate RR for quintiles of the pattern score were 1.00, 1.50, 1.61, 1.96,

and 2.56 (95% CI: 2.10, 3.12;  $P < 0.001$ ). To evaluate whether the association was mediated in part by specific nutrients associated with the pattern, we adjusted in a further analysis for *trans* fatty acids, magnesium, caffeine, cereal fiber, glycemic index, alcohol, and the ratio of polyunsaturated to saturated fat. The association between the pattern and diabetes risk was attenuated but remained significant (RR: 1.99; 95% CI: 1.60, 2.49). The association appeared to be consistent across strata of BMI ( $< 24$  versus  $\geq 24$ ), history of diabetes in a first-degree relative, and physical activity (Table 4). In addition, the association remained significant after additional adjustment for the waist-to-hip ratio in an analysis among those 21 394 women (733 cases) who reported waist and hip circumferences in 1986 (RR for extreme quintiles: 2.22; 95% CI: 1.68, 2.94).

During 701 155 person-years of follow-up, we documented 724 new cases of type 2 diabetes in the NHS-II. A higher pattern score was strongly associated with an increased risk of type 2 diabetes (age-adjusted RR: 6.89; Table 4). This association was substantially attenuated after adjustment for BMI, but it remained strong (RR: 3.05; 95% CI: 2.29, 4.06;  $P < 0.001$ ). Additional adjustment for other lifestyle covariates had little effect on the observed association; the multivariate RRs for quintiles of the pattern score were 1.00, 1.92, 1.64, 2.10, and 2.93 (95% CI: 2.18, 3.92;  $P < 0.001$ ). Adjustment for *trans* fatty acids, magnesium, caffeine, cereal fiber, glycemic index, alcohol, and the ratio of polyunsaturated to saturated fat attenuated the association (RR: 2.15; CI: 1.55, 2.97). The association between the pattern and diabetes risk was relatively consistent across strata of physical activity and history of diabetes in a first-degree relative (Table 4). However, the association appeared to be stronger among women in the NHS-II with high BMI ( $\geq 27$ ; RR for extreme quintiles: 3.15) than among women with low BMI (RR: 1.93;  $P$  for interaction: 0.02). We additionally adjusted for the waist-to-hip ratio in a separate analysis among 42 967 women (273 cases) who reported their waist and hip circumferences in 1993, but the pattern remained significantly associated with an increased diabetes risk (RR for extreme quintiles: 2.37; 95% CI: 1.48, 3.79).

## DISCUSSION

Using a novel statistical method, we derived a diet pattern score that was strongly associated with markers of inflammation and endothelial dysfunction. This pattern strongly predicted risk of type 2 diabetes in a nested case-control analysis, independent

**TABLE 2**

Risk of type 2 diabetes by quintile of diet pattern score: unconditional logistic regression, 656 cases and 694 controls, Nurses' Health Study<sup>1</sup>

|                                       | Quintile of diet pattern score <sup>2</sup> |                   |                   |                   |                   | <i>P</i> for trend |
|---------------------------------------|---------------------------------------------|-------------------|-------------------|-------------------|-------------------|--------------------|
|                                       | 1                                           | 2                 | 3                 | 4                 | 5                 |                    |
| No. of cases                          | 54                                          | 80                | 117               | 164               | 241               |                    |
| Age-adjusted OR                       | 1.00                                        | 1.49 (0.98, 2.26) | 2.18 (1.46, 3.26) | 3.10 (2.10, 4.57) | 4.70 (3.20, 6.90) | <0.001             |
| Age and BMI-adjusted OR               | 1.00                                        | 1.23 (0.77, 1.95) | 1.80 (1.16, 2.80) | 2.20 (1.43, 3.37) | 3.09 (2.03, 4.69) | <0.001             |
| Multivariate-adjusted OR <sup>3</sup> | 1.00                                        | 1.28 (0.79, 2.07) | 1.89 (1.19, 2.98) | 2.25 (1.45, 3.51) | 3.09 (1.99, 4.79) | <0.001             |

<sup>1</sup> OR, odds ratio.

<sup>2</sup> 95% CI in parentheses. A high diet pattern score reflects a diet high in sugar-sweetened soft drinks, refined grains, diet soft drinks, processed meat, and other vegetables and low in wine, coffee, cruciferous vegetables, and yellow vegetables. Quintiles are based on the control population.

<sup>3</sup> Adjusted for age, BMI (9 categories), physical activity (quintiles), family history of diabetes, smoking (never, past, current, or missing), postmenopausal hormone use (never, ever, or missing), and energy intake (quintiles).

**TABLE 3**  
Age-adjusted characteristics at baseline by quintile of diet pattern score, Nurses' Health Study (NHS) and Nurses' Health Study II (NHS-II)<sup>1</sup>

| Characteristic                                      | Quintile of diet pattern score in the NHS ( <i>n</i> = 35 367) |             |             |             |             | Quintile of diet pattern score in the NHS-II ( <i>n</i> = 89 311) |             |             |             |             |
|-----------------------------------------------------|----------------------------------------------------------------|-------------|-------------|-------------|-------------|-------------------------------------------------------------------|-------------|-------------|-------------|-------------|
|                                                     | 1                                                              | 2           | 3           | 4           | 5           | 1                                                                 | 2           | 3           | 4           | 5           |
| Age (y)                                             | 52.2 ± 6.8 <sup>2</sup>                                        | 51.6 ± 7.0  | 50.9 ± 7.1  | 49.9 ± 7.2  | 48.5 ± 7.1  | 37.6 ± 4.3                                                        | 36.5 ± 4.5  | 35.8 ± 4.7  | 35.4 ± 4.7  | 35.1 ± 4.7  |
| BMI (kg/m <sup>2</sup> )                            | 24.0 ± 3.8                                                     | 24.4 ± 4.1  | 24.9 ± 4.5  | 25.4 ± 4.9  | 26.2 ± 5.6  | 23.7 ± 4.4                                                        | 24.0 ± 4.7  | 24.3 ± 4.9  | 25.0 ± 4.7  | 26.2 ± 4.7  |
| Physical activity (METs/wk)                         | 17.6 ± 28.2                                                    | 14.7 ± 20.0 | 13.7 ± 20.2 | 12.9 ± 19.8 | 11.4 ± 18.7 | 26.5 ± 31.3                                                       | 21.9 ± 27.5 | 20.0 ± 25.6 | 18.8 ± 25.2 | 17.7 ± 25.0 |
| Currently smoking (%)                               | 29.6                                                           | 27.1        | 24.5        | 24.6        | 27.2        | 14.8                                                              | 11.6        | 10.0        | 10.8        | 13.6        |
| Family history of diabetes (%)                      | 16.1                                                           | 15.9        | 16.8        | 17.5        | 17.7        | 15.2                                                              | 15.8        | 15.8        | 17.2        | 18.0        |
| Currently receiving hormone replacement therapy (%) | 12.9                                                           | 12.9        | 13.0        | 12.5        | 11.3        | 2.3                                                               | 2.4         | 2.3         | 2.6         | 2.8         |
| Nutrient intake                                     |                                                                |             |             |             |             |                                                                   |             |             |             |             |
| Energy (kcal/d)                                     | 1659 ± 499                                                     | 1605 ± 494  | 1646 ± 496  | 1753 ± 501  | 2033 ± 552  | 1699 ± 515                                                        | 1653 ± 504  | 1693 ± 505  | 1817 ± 520  | 2096 ± 570  |
| Alcohol (g/d)                                       | 11.4 ± 14.7                                                    | 6.9 ± 10.1  | 5.9 ± 10.2  | 5.3 ± 9.5   | 5.4 ± 10.4  | 5.4 ± 8.6                                                         | 3.1 ± 5.5   | 2.5 ± 4.8   | 2.3 ± 4.9   | 2.2 ± 5.1   |
| Carbohydrates (% of energy)                         | 45.6 ± 8.4                                                     | 46.2 ± 7.9  | 46.4 ± 7.6  | 46.6 ± 7.5  | 47.7 ± 8.2  | 50.0 ± 8.0                                                        | 49.9 ± 7.4  | 49.7 ± 7.2  | 49.4 ± 7.1  | 50.0 ± 7.9  |
| Polyunsaturated:saturated fat                       | 0.57 ± 0.20                                                    | 0.56 ± 0.18 | 0.55 ± 0.18 | 0.53 ± 0.16 | 0.53 ± 0.16 | 0.56 ± 0.19                                                       | 0.53 ± 0.16 | 0.52 ± 0.15 | 0.51 ± 0.14 | 0.50 ± 0.14 |
| <i>trans</i> Fat (% of energy)                      | 1.7 ± 0.6                                                      | 1.9 ± 0.6   | 1.9 ± 0.6   | 2.0 ± 0.6   | 2.0 ± 0.6   | 1.4 ± 0.6                                                         | 1.6 ± 0.6   | 1.7 ± 0.6   | 1.7 ± 0.6   | 1.8 ± 0.6   |
| Magnesium (mg/d)                                    | 335 ± 77                                                       | 305 ± 71    | 285 ± 67    | 264 ± 63    | 243 ± 59    | 363 ± 76                                                          | 333 ± 71    | 313 ± 67    | 297 ± 65    | 272 ± 62    |
| Caffeine (mg/d)                                     | 427 ± 266                                                      | 362 ± 242   | 298 ± 219   | 270 ± 202   | 254 ± 193   | 372 ± 271                                                         | 266 ± 227   | 208 ± 199   | 187 ± 179   | 185 ± 159   |
| Glycemic index                                      | 51.6 ± 3.8                                                     | 52.7 ± 3.6  | 53.4 ± 3.5  | 54.1 ± 3.4  | 55.0 ± 3.4  | 52.3 ± 3.3                                                        | 53.3 ± 3.2  | 54.0 ± 3.1  | 54.4 ± 3.1  | 55.3 ± 3.2  |
| Cereal fiber (g/d)                                  | 4.2 ± 2.6                                                      | 4.3 ± 2.5   | 4.2 ± 2.4   | 4.0 ± 2.1   | 3.7 ± 1.8   | 6.1 ± 3.7                                                         | 5.9 ± 3.3   | 5.7 ± 2.9   | 5.4 ± 2.6   | 4.9 ± 2.4   |

<sup>1</sup> A high diet pattern score reflects a diet high in sugar-sweetened soft drinks, refined grains, diet soft drinks, processed meat, and other vegetables and low in wine, coffee, cruciferous vegetables, and yellow vegetables. Tests for trend (based on ordinal variables containing median values for each quintile) were all significant ( $P < 0.01$ ), except for carbohydrates in the NHS-II.

<sup>2</sup>  $\bar{x} \pm SD$  (all such values).

**TABLE 4**  
Risk of type 2 diabetes by quintile of diet pattern score, Cox proportional hazards analysis, Nurses' Health Study (NHS) and Nurses' Health Study II (NHS-II)<sup>a</sup>

|                                                 | Quintile of diet pattern score in the NHS (n = 35 367) |                   |                   |                   |                   | Quintile of diet pattern score in the NHS-II (n = 89 311) |         |                   |                   |                   | P <sup>2</sup>    |        |
|-------------------------------------------------|--------------------------------------------------------|-------------------|-------------------|-------------------|-------------------|-----------------------------------------------------------|---------|-------------------|-------------------|-------------------|-------------------|--------|
|                                                 | 1                                                      | 2                 | 3                 | 4                 | 5                 | P <sup>2</sup>                                            | 1       | 2                 | 3                 | 4                 |                   | 5      |
| No. of cases                                    | 133                                                    | 227               | 271               | 363               | 523               | —                                                         | 57      | 114               | 104               | 156               | 293               | —      |
| Person-years                                    | 91 837                                                 | 91 855            | 91 897            | 91 883            | 91 520            | —                                                         | 139 949 | 140 312           | 140 471           | 140 435           | 139 988           | —      |
| Relative risks (95% CI)                         |                                                        |                   |                   |                   |                   |                                                           |         |                   |                   |                   |                   |        |
| Age-adjusted                                    | 1.00                                                   | 1.74 (1.41, 2.16) | 2.15 (1.75, 2.65) | 2.98 (2.44, 3.64) | 4.66 (3.84, 5.64) | <0.001                                                    | 1.00    | 2.25 (1.64, 3.10) | 2.23 (1.61, 3.08) | 3.32 (2.60, 4.77) | 6.89 (5.18, 9.16) | <0.001 |
| Age- and BMI-adjusted                           | 1.00                                                   | 1.51 (1.22, 1.87) | 1.65 (1.34, 2.03) | 2.01 (1.64, 2.45) | 2.63 (2.17, 3.20) | <0.001                                                    | 1.00    | 1.95 (1.42, 2.67) | 1.66 (1.20, 2.30) | 2.15 (1.59, 2.92) | 3.05 (2.29, 4.06) | <0.001 |
| Multivariate-adjusted <sup>b</sup>              |                                                        |                   |                   |                   |                   |                                                           |         |                   |                   |                   |                   |        |
| All                                             | 1.00                                                   | 1.50 (1.21, 1.86) | 1.61 (1.31, 1.99) | 1.96 (1.61, 2.40) | 2.56 (2.10, 3.12) | <0.001                                                    | 1.00    | 1.92 (1.39, 2.64) | 1.64 (1.18, 2.26) | 2.10 (1.55, 2.86) | 2.93 (2.18, 3.92) | <0.001 |
| Without family history of diabetes <sup>c</sup> | 1.00                                                   | 1.56 (1.21, 2.03) | 1.60 (1.24, 2.07) | 1.98 (1.55, 2.53) | 2.71 (2.13, 3.46) | —                                                         | 1.00    | 1.90 (1.26, 2.85) | 1.58 (1.04, 2.39) | 2.10 (1.42, 3.10) | 2.97 (2.04, 4.31) | —      |
| With family history of diabetes <sup>d</sup>    | 1.00                                                   | 1.38 (0.95, 2.02) | 1.66 (1.16, 2.37) | 1.93 (1.36, 2.73) | 2.29 (1.62, 3.23) | —                                                         | 1.00    | 1.96 (1.18, 3.27) | 1.74 (1.03, 2.92) | 2.13 (1.30, 3.50) | 2.91 (1.81, 4.67) | —      |
| Low physical activity <sup>e</sup>              | 1.00                                                   | 1.29 (0.92, 1.81) | 1.60 (1.17, 2.21) | 1.88 (1.38, 2.56) | 2.55 (1.89, 3.46) | —                                                         | 1.00    | 1.49 (0.94, 2.37) | 1.80 (1.16, 2.81) | 1.77 (1.15, 2.73) | 2.70 (1.80, 4.07) | —      |
| High physical activity <sup>e</sup>             | 1.00                                                   | 1.69 (1.28, 2.23) | 1.62 (1.23, 2.14) | 2.05 (1.57, 2.67) | 2.61 (2.00, 3.40) | —                                                         | 1.00    | 2.45 (1.58, 3.81) | 1.38 (0.84, 2.25) | 2.59 (1.68, 3.99) | 3.19 (2.09, 4.87) | —      |
| Low BMI <sup>f</sup>                            | 1.00                                                   | 1.08 (0.60, 1.94) | 1.88 (1.11, 3.21) | 2.51 (1.49, 4.22) | 2.43 (1.40, 4.23) | —                                                         | 1.00    | 2.44 (1.34, 4.42) | 1.28 (0.64, 2.54) | 1.63 (0.85, 3.14) | 1.93 (1.00, 3.70) | 0.2    |
| High BMI <sup>f</sup>                           | 1.00                                                   | 1.58 (1.26, 2.00) | 1.59 (1.27, 2.00) | 1.94 (1.56, 2.41) | 2.56 (2.07, 3.18) | —                                                         | 1.00    | 1.82 (1.25, 2.65) | 1.80 (1.24, 2.61) | 2.29 (1.61, 3.26) | 3.15 (2.24, 4.41) | <0.001 |

<sup>a</sup> A high diet pattern score reflects a diet high in sugar-sweetened soft drinks, refined grains, diet soft drinks, processed meat, and other vegetables and low in wine, coffee, cruciferous vegetables, and yellow vegetables.

<sup>b</sup> P for linear trends across quintiles of the pattern score on the basis of an ordinal variable containing the median value for each quintile.

<sup>c</sup> Adjusted for age, BMI (9 categories), physical activity (quintiles), family history of diabetes, smoking (never, past, current, or missing), postmenopausal hormone use (never, ever, or missing), and energy intake (quintiles). Stratified analyses were not adjusted for stratifying variable, except BMI strata (adjusted for continuous BMI within each stratum).

<sup>d</sup> P for interaction for family history of diabetes = 0.28 for NHS and 0.20 for NHS-II.

<sup>e</sup> Low baseline activity = lower 2 quintiles of activity level; high activity = upper 3 quintiles of activity level. P for interaction for physical activity = 0.54 for NHS and 0.73 for NHS-II.

<sup>f</sup> Based on baseline population median (NHS: 24; NHS-II: 27). P for interaction for BMI = 0.46 for NHS and 0.017 for NHS-II.

of BMI and other diabetes risk factors. This association was subsequently confirmed in 2 separate cohorts of women. These findings provide evidence that the association between dietary factors and risk of type 2 diabetes may be mediated in part by inflammation and endothelial dysfunction.

The pattern we identified was characterized by high intakes of sugar-sweetened soft drinks, refined grains, diet soft drinks, processed meat, and low intakes of wine, coffee, cruciferous vegetables, and yellow vegetables. Other vegetables (celery, mushrooms, green pepper, corn, mixed vegetables, eggplant, and summer squash) were not correlated with inflammation and were only moderately associated with the pattern; thus, the contribution of this food group to the overall pattern was negligible. Most of these food groups have been identified to be associated with diabetes risk in previous studies and some have also been found to be associated with inflammatory markers. Moderate alcohol consumption (1–3 drinks/d) has been consistently associated with lower incidence of diabetes (26) and lower levels of pro-inflammatory markers (27–32). An inverse association between coffee consumption and risk of type 2 diabetes was observed in several prospective cohort studies (33–36). Sugar-sweetened beverages have been associated with risk of diabetes among women (37); these beverages contribute importantly to glycemic load, which has been associated with inflammatory markers (38). Several epidemiologic studies found that diets rich in whole grains compared with refined grains may protect against type 2 diabetes (39–43). Refined grains may be associated with increased diabetes risk because these foods tend to be low in cereal fiber and have a high glycemic index, which both appear to be associated with increased diabetes risk (4). Other components of whole grains may also have beneficial effects, with isoflavones being potentially associated with decreased inflammation (44). Frequent consumption of meat, in particular processed meat, has been consistently shown to increase the risk of diabetes in prospective studies (6, 45–47). Advanced glycation end products, which are high in processed animal foods high in protein and fat, have been found to promote inflammatory mediators in humans (48). Vegetable consumption was inversely associated with risk of diabetes in the National Health and Nutrition Examination Survey (49) and in a Finnish cohort study (50) but not among older women in the Iowa Women's Health Study (41). The effects of vegetable consumption on inflammatory processes are largely unknown. A cross-sectional study of the elderly observed lower concentrations of CRP with higher fruit and vegetable consumption, although this study did not provide estimates for vegetables alone or for specific subgroups of vegetables (51). In a recent trial over a 2-y period among men and women with the metabolic syndrome, increased consumption of fruit, vegetables, walnuts, whole grains, and olive oil significantly reduced concentrations of CRP, IL-6, IL-7, and IL-18 and improved endothelial function compared with that in a control group that consumed an otherwise healthy diet (<30% fat, <10% saturated fat) (52). These effects were attenuated but not eliminated by additional adjustment for weight change over the course of the study.

Recently, we reported the role of overall dietary patterns derived by using factor analysis in predicting the risk of diabetes in 2 cohort studies (5, 6). A prudent pattern (characterized by a high consumption of vegetables, fruit, fish, poultry, and whole grains) was associated with a modest nonsignificant risk reduction in both studies, whereas a Western pattern (characterized by a high consumption of red meat, processed meat, French fries, high-fat

dairy products, refined grains, and sweets and desserts) was associated with an increased risk of type 2 diabetes. Here, the multivariate RR for extreme quintiles was 1.49 (95% CI: 1.26, 1.76) in the NHS (6) and 1.59 (95% CI: 1.32, 1.93) in the Health Professionals Follow-up Study (5). The prudent pattern was inversely associated with plasma concentrations of CRP and E-selectin, and the Western pattern showed a positive relation with CRP, E-selectin, sICAM-1, and sVCAM-1 after adjustment for age, BMI, physical activity, smoking status, and alcohol consumption in the NHS among control women of the same nested case-control study used in our analysis (10). In addition, the Western pattern was significantly correlated with CRP in the Health Professionals Follow-up Study after adjustment for a variety of risk factors but not after adjustment for BMI (9). In a recent study of 4304 Finnish men and women, a prudent pattern (characterized by higher consumption of fruit and vegetables) was significantly associated with a lower risk of type 2 diabetes, whereas a "conservative" pattern (characterized by consumption of butter, potatoes, and whole milk) was significantly associated with an increased risk (7).

In contrast with our previous analyses and the Finnish study, which derived the dietary patterns by using factor analysis based on the observed covariance among food groups, the present study used the information on inflammatory biomarkers to derive the dietary pattern. The advantage of this approach as opposed to the factor analysis approach is that the derived dietary pattern incorporates information on biological pathways and thus is hypothesis-driven instead of being driven by patterns of eating behavior and could be more predictive of disease risk. Using the same technique, we previously identified a dietary pattern in the prospective EPIC-Potsdam cohort that was characterized by a high intake of fresh fruit and a low intake of sugar-sweetened beverages, beer, meat, poultry, processed meat, legumes, and bread, excluding whole-grain bread (8). Subjects who scored high had high plasma concentrations of HDL cholesterol and adiponectin and low plasma concentrations of glycated hemoglobin. After multivariate adjustment, the RR for type 2 diabetes mellitus for extreme quintiles of the dietary pattern score was 0.27 (95% CI: 0.13, 0.64; *P* for trend < 0.001). However, the pattern was not significantly associated with CRP concentrations, and we were not able to verify these results in independent study samples, as we did in the present study.

The current RRR approach requires response (biomarker) information. This information may not be available in many studies otherwise suitable for evaluating diet-disease associations. Also, the biomarker information available may not reflect the current state of knowledge. In our study, we selected inflammatory markers that previously predicted risk of diabetes in the same cohort (53, 54). However, pathways other than inflammation may also be relevant in the development of diabetes. For example, we observed significant associations between markers of body iron stores (55) and  $\beta$  cell function (56) and risk of diabetes in the NHS, but did not consider these markers in the present analysis for 2 reasons. First, measures of body iron stores were only weakly correlated with inflammatory markers, which limits the usefulness of their additional inclusion in the RRR analysis. Second,  $\beta$  cell function was not measured in a large proportion of the current sample. Although this narrows potential effects of diet on diabetes risk to a single pathway, previous RRR analyses suggest that one single dietary pattern is unlikely to explain several different and independent pathways (8, 57). It might be of

interest to evaluate whether other biomarkers play a role in mediating effects of the dietary pattern on inflammatory markers, for example, lipoproteins. Unfortunately, we did not have lipoprotein markers available for our nested case-control study population.

Obesity induces a state of chronic low-grade inflammation (58), and excess body fat may therefore also explain associations between those food groups identified to be components of the dietary pattern and inflammatory markers. For example, diet soft drinks were directly associated with BMI cross-sectionally in the NHS-II (37), although this association most likely represents a reverse causation because the use of diet soft drinks instead of regular soft drinks is the most frequently reported diet intervention to lose weight in US adults (59). Although diet soft drinks were identified as a component of the RRR pattern, associations between diet soft drinks and inflammatory markers may be confounded by BMI. With use of the BMI-adjusted biomarker levels as responses in RRR, the pattern remained similar with sugar-sweetened beverages and refined grains being positively associated with the pattern and wine, coffee, and cruciferous vegetables being negatively associated. However, diet soft drinks were no longer an important component of the pattern. Adjustment for BMI also partly attenuated the age-adjusted associations between the dietary pattern and risk of diabetes. However, it is also possible that weight gain is one potential pathway by which the dietary pattern is associated with inflammation and diabetes risk. Adjustment for BMI may therefore lead to an underestimation of the true effect of the diet. To determine whether BMI is a confounder or mediator is not possible with our study design, because associations between the dietary pattern, BMI, and inflammatory markers were analyzed cross-sectionally. Nevertheless, component foods such as sugar-sweetened soft drinks, refined grains, wine, coffee, and vegetables appear to relate to inflammatory markers independent of BMI, and the association between the dietary pattern and diabetes risk remained strong after adjustment for BMI and the waist-to-hip ratio. Our data also suggest that the diet pattern score may be more strongly associated with diabetes risk among obese women than among lean women.

The repeated dietary measurements used in this study were advantageous because they dampened measurement errors and took into account changes in eating behaviors over time (23). Compared with the traditional approach of examining the effects of individual nutrients or foods, the dietary pattern approach has the advantage of representing the cumulative effects of overall diet.

The NHS and NHS-II cohorts are study populations of US female nurses and therefore are not representative of the general US female population. Thus, our results should be replicated in other populations. Another potential limitation of our study is the reliance on self-reported confounder information. For example, smoking has been related to inflammation (60) and diabetes risk (4), and smoking cessation has effects on body weight (61). Because smoking and poor diet may also be part of an unhealthy lifestyle, residual confounding due to measurement error in assessing smoking history or due to insufficient control in statistical models might have biased our observations, but it is unlikely that such a bias would explain the strong relation between the dietary pattern and diabetes risk. Similarly, a history of hypertension may influence dietary behavior and is associated with inflammation (62). Although adjustment for hypertension did

not alter our observations, residual confounding might still be present.

In conclusion, our data suggest that a diet high in sugar-sweetened soft drinks, refined grains, diet soft drinks, and processed meat and low in wine, coffee, cruciferous vegetables, and yellow vegetables may increase the risk of developing type 2 diabetes, possibly by exacerbating inflammatory processes. 

MBS contributed to the development of the analysis plan, conducted the statistical analyses, collaborated on the interpretation of the results, and wrote the manuscript. KH, JEM, WCW, CH, and FBH provided significant consultation on the statistical analysis plan, interpretation of results, and writing of the manuscript. CW, GAC, and JBM provided significant consultation on the interpretation of results and writing of the manuscript. None of the authors had any financial or personal interest in any company or organization sponsoring this research, including advisory board affiliations.

## REFERENCES

- Pickup JC, Mattock MB, Chusney GD, Burt D. NIDDM as a disease of the innate immune system: association of acute-phase reactants and interleukin-6 with metabolic syndrome X. *Diabetologia* 1997;40:1286–92.
- Pickup JC. Inflammation and activated innate immunity in the pathogenesis of type 2 diabetes. *Diabetes Care* 2004;27:813–23.
- Brown AA, Hu FB. Dietary modulation of endothelial function: implications for cardiovascular disease. *Am J Clin Nutr* 2001;73:673–86.
- Schulze MB, Hu FB. Primary prevention of diabetes: what can be done and how much can be prevented? *Annu Rev Public Health* 2005;26:445–67.
- van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. *Ann Intern Med* 2002;136:201–9.
- Fung TT, Schulze M, Manson JE, Willett WC, Hu FB. Dietary patterns, meat intake, and the risk of type 2 diabetes in women. *Arch Intern Med* 2004;164:2235–40.
- Montonen J, Knekt P, Harkanen T, et al. Dietary patterns and the incidence of type 2 diabetes. *Am J Epidemiol* 2005;161:219–27.
- Heidemann C, Hoffmann K, Spranger J, et al. A dietary pattern protective for type 2 diabetes mellitus in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study cohort. *Diabetologia* 2005;48:1126–34.
- Fung TT, Rimm EB, Spiegelman D, et al. Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr* 2001;73:61–7.
- Lopez-Garcia E, Schulze MB, Fung TT, et al. Major dietary patterns are related to plasma concentrations of markers of inflammation and endothelial dysfunction. *Am J Clin Nutr* 2004;80:1029–35.
- Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* 2005;5:388–96.
- Hu FB, Rimm E, Smith-Warner SA, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* 1999;69:243–9.
- US Department of Agriculture. Composition of foods: raw, processed, prepared, 1963–1991. Washington, DC: US Government Printing Office, 1992.
- Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *Am J Epidemiol* 1986;124:17–27.
- Salvini S, Hunter DJ, Sampson L, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int J Epidemiol* 1989;18:858–67.
- Willett WC, Lenart E. Reproducibility and validity of food-frequency questionnaires. In: Willett WC, ed. *Nutritional epidemiology*. New York: Oxford University Press, 1998:101–56.
- National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 1979;28:1039–57.
- Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 1997;20:1183–97.
- Manson JE, Rimm EB, Stampfer MJ, et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *Lancet* 1991;338:774–8.

20. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med* 2001;161:1542–8.
21. Willett W, Stampfer MJ, Bain C, et al. Cigarette smoking, relative weight, and menopause. *Am J Epidemiol* 1983;117:651–8.
22. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol* 1994;23:991–9.
23. Hu FB, Stampfer MJ, Rimm E, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. *Am J Epidemiol* 1999;149:531–40.
24. Hoffmann K, Schulze MB, Schienkiewitz A, Nothlings U, Boeing H. Application of a new statistical method to derive dietary patterns in nutritional epidemiology. *Am J Epidemiol* 2004;159:935–44.
25. Schulze MB, Hoffmann K, Kroke A, Boeing H. An approach to construct simplified measures of dietary patterns from exploratory factor analysis. *Br J Nutr* 2003;89:409–19.
26. Howard AA, Arnsten JH, Gourevitch MN. Effect of alcohol consumption on diabetes mellitus: a systematic review. *Ann Intern Med* 2004;140:211–9.
27. Volpato S, Pahor M, Ferrucci L, et al. Relationship of alcohol intake with inflammatory markers and plasminogen activator inhibitor-1 in well-functioning older adults: the Health, Aging, and Body Composition study. *Circulation* 2004;109:607–12.
28. Mukamal KJ, Cushman M, Mittleman MA, Tracy RP, Siscovick DS. Alcohol consumption and inflammatory markers in older adults: the Cardiovascular Health Study. *Atherosclerosis* 2004;173:79–87.
29. Imhof A, Froehlich M, Brenner H, Boeing H, Pepys MB, Koenig W. Effect of alcohol consumption on systemic markers of inflammation. *Lancet* 2001;357:763–7.
30. Imhof A, Woodward M, Doering A, et al. Overall alcohol intake, beer, wine, and systemic markers of inflammation in western Europe: results from three MONICA samples (Augsburg, Glasgow, Lille). *Eur Heart J* 2004;25:2092–100.
31. Shai I, Rimm EB, Schulze MB, Rifai N, Stampfer MJ, Hu FB. Moderate alcohol intake and markers of inflammation and endothelial dysfunction among diabetic men. *Diabetologia* 2004;47:1760–7.
32. Estruch R, Sacanella E, Badia E, et al. Different effects of red wine and gin consumption on inflammatory biomarkers of atherosclerosis: a prospective randomized crossover trial. Effects of wine on inflammatory markers. *Atherosclerosis* 2004;175:117–23.
33. Rosengren A, Dotevall A, Wilhelmsen L, Thelle D, Johansson S. Coffee and incidence of diabetes in Swedish women: a prospective 18-year follow-up study. *J Intern Med* 2004;255:89–95.
34. Salazar-Martinez E, Willett WC, Ascherio A, et al. Coffee consumption and risk for type 2 diabetes mellitus. *Ann Intern Med* 2004;140:1–8.
35. Tuomilehto J, Hu G, Bidel S, Lindstrom J, Jousilahti P. Coffee consumption and risk of type 2 diabetes mellitus among middle-aged Finnish men and women. *JAMA* 2004;291:1213–9.
36. van Dam RM, Feskens EJ. Coffee consumption and risk of type 2 diabetes mellitus. *Lancet* 2002;360:1477–8.
37. Schulze MB, Manson JE, Ludwig DS, et al. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *JAMA* 2004;292:927–34.
38. Liu S, Manson JE, Buring JE, Stampfer MJ, Willett WC, Ridker PM. Relation between a diet with a high glycemic load and plasma concentrations of high-sensitivity C-reactive protein in middle-aged women. *Am J Clin Nutr* 2002;75:492–8.
39. Fung TT, Hu FB, Pereira MA, et al. Whole-grain intake and the risk of type 2 diabetes: a prospective study in men. *Am J Clin Nutr* 2002;76:535–40.
40. Liu S, Manson JE, Stampfer MJ, et al. A prospective study of whole-grain intake and risk of type 2 diabetes mellitus in US women. *Am J Public Health* 2000;90:1409–15.
41. Meyer KA, Kushi LH, Jacobs DR Jr, Slavin J, Sellers TA, Folsom AR. Carbohydrates, dietary fiber, and incident type 2 diabetes in older women. *Am J Clin Nutr* 2000;71:921–30.
42. Montonen J, Knekt P, Jarvinen R, Aromaa A, Reunanen A. Whole-grain and fiber intake and the incidence of type 2 diabetes. *Am J Clin Nutr* 2003;77:622–9.
43. McKeown NM, Meigs JB, Liu S, Wilson PW, Jacques PF. Whole-grain intake is favorably associated with metabolic risk factors for type 2 diabetes and cardiovascular disease in the Framingham Offspring Study. *Am J Clin Nutr* 2002;76:390–8.
44. Anderson JW. Whole grains protect against atherosclerotic cardiovascular disease. *Proc Nutr Soc* 2003;62:135–42.
45. Schulze MB, Manson JE, Willett WC, Hu FB. Processed meat intake and incidence of Type 2 diabetes in younger and middle-aged women. *Diabetologia* 2003;46:1465–73.
46. van Dam RM, Willett WC, Rimm EB, Stampfer MJ, Hu FB. Dietary fat and meat intake in relation to risk of type 2 diabetes in men. *Diabetes Care* 2002;25:417–24.
47. Song Y, Manson JE, Buring JE, Liu S. A prospective study of red meat consumption and type 2 diabetes in middle-aged and elderly women: The Women's Health Study. *Diabetes Care* 2004;27:2108–15.
48. Vlassara H, Cai W, Crandall J, et al. Inflammatory mediators are induced by dietary glycotoxins, a major risk factor for diabetic angiopathy. *Proc Natl Acad Sci U S A* 2002;99:15596–601.
49. Ford ES, Mokdad AH. Fruit and vegetable consumption and diabetes mellitus incidence among U.S. adults. *Prev Med* 2001;32:33–9.
50. Montonen J, Jarvinen R, Heliovaara M, Reunanen A, Aromaa A, Knekt P. Food consumption and the incidence of type II diabetes mellitus. *Eur J Clin Nutr* 2005;59:441–8.
51. Gao X, Bermudez OI, Tucker KL. Plasma C-reactive protein and homocysteine concentrations are related to frequent fruit and vegetable intake in Hispanic and non-Hispanic white elders. *J Nutr* 2004;134:913–8.
52. Esposito K, Marfella R, Ciotola M, et al. Effect of a Mediterranean-style diet on endothelial dysfunction and markers of vascular inflammation in the metabolic syndrome: a randomized trial. *JAMA* 2004;292:1440–6.
53. Meigs JB, Hu FB, Rifai N, Manson JE. Biomarkers of endothelial dysfunction and risk of type 2 diabetes mellitus. *JAMA* 2004;291:1978–86.
54. Hu FB, Meigs JB, Li TY, Rifai N, Manson JE. Inflammatory markers and risk of developing type 2 diabetes in women. *Diabetes* 2004;53:693–700.
55. Jiang R, Manson JE, Meigs JB, Ma J, Rifai N, Hu FB. Body iron stores in relation to risk of type 2 diabetes in apparently healthy women. *JAMA* 2004;291:711–7.
56. Schulze MB, Solomon CG, Rifai N, et al. Hyperproinsulinemia and risk of type 2 diabetes in women. *Diabet Med* 2005; Mar 2 [Epub ahead of print; DOI: 10.1111/j.1464-5491.2005.01585.x].
57. Hoffmann K, Zyriax BC, Boeing H, Windler E. A dietary pattern derived to explain biomarker variation is strongly associated with the risk of coronary artery disease. *Am J Clin Nutr* 2004;80:633–40.
58. Trayhurn P, Wood IS. Adipokines: inflammation and the pleiotropic role of white adipose tissue. *Br J Nutr* 2004;92:347–55.
59. Levy AS, Heaton AW. Weight control practices of U.S. adults trying to lose weight. *Ann Intern Med* 1993;119:661–6.
60. Wannamethee SG, Lowe GD, Shaper AG, Rumley A, Lennon L, Whincup PH. Associations between cigarette smoking, pipe/cigar smoking, and smoking cessation, and haemostatic and inflammatory markers for cardiovascular disease. *Eur Heart J* 2005;Apr 7 [Epub ahead of print; DOI:10.1093/eurheartj/ehi183].
61. Filozof C, Fernandez Pinilla MC, Fernandez-Cruz A. Smoking cessation and weight gain. *Obes Rev* 2004;5:95–103.
62. Virdis A, Schiffrin EL. Vascular inflammation: a role in vascular disease in hypertension? *Curr Opin Nephrol Hypertens* 2003;12:181–7.