# Generalizability of a Diabetes-Associated Country-Specific Exploratory Dietary Pattern Is Feasible Across European Populations

Franziska Jannasch,[1,2,3] Janine Kröger,[1,2] Claudia Agnoli,[4] Aurelio Barricarte,[5,6,7] Heiner Boeing,[8] Valerie Cayssials,[9] Sandra Colorado-Yohar,[5,10,11] Christina C Dahm,[12] Courtney Dow,[13] Guy Fagherazzi,[13] Paul W Franks,[14,15] Heinz Freisling,[16] Marc J Gunter,[16] Nicola D Kerrison,[17] Timothy J Key,[18] Kay-Tee Khaw,[19] Tilman Kühn,[20] Cecilie Kyro,[21] Francesca Romana Mancini,[13] Olatz Mokoroa,[5,22] Peter Nilsson,[14] Kim Overvad,[12,23] Domenico Palli,[24] Salvatore Panico,[25] Jose Ramón Quirós García,[26] Olov Rolandsson,[15] Carlotta Sacerdote,[27,28] Mariá-José Sánchez,[5,29] Mohammad Sediq Sahrai,[16] Ruth Schübel,[20] Ivonne Sluijs,[30] Annemieke MW Spijkerman,[31] Anne Tjonneland,[21] Tammy YN Tong,[18] Rosario Tumino,[32,33] Elio Riboli,[34] Claudia Langenberg,[17] Stephen J Sharp,[17] Nita G Forouhi,[17] Matthias B Schulze,[1,2,35] and Nicholas J Wareham[17]

[1]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Germany; [2]German Center for Diabetes Research (DZD), München-Neuherberg, Germany; [3]NutriAct–Competence Cluster Nutrition Research Berlin-Potsdam, Nuthetal, Germany; [4]Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; [5]CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain; [6]Navarra Public Health Institute, Pamplona, Spain; [7]IdiSNA, Navarra Institute for Health Research, Pamplona, Spain; [8]Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; [9]Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain; [10]Department of Epidemiology, Murcia Health Council, IMIB-Arrixaca, Spain; [11]Research Group on Demography and Health, National Faculty of Public Health, University of Antioquia, Medellín, Colombia; [12]Section for Epidemiology, Department of Public Health, Aarhus University, Denmark; [13]Inserm, Centre for Research in Epidemiology and Population Health [CESP], Villejuif, France; Université Paris-Sud, Villejuif, France; [14]Department of Clinical Sciences, Lund University, Malmö, Sweden; [15]Department of Public Health and Clinical Medicine, Umeå University, Sweden; [16]Section of Nutrition and Metabolism, International Agency for Research on Cancer (IARC-WHO), Lyon, France; [17]MRC Epidemiology Unit, University of Cambridge, United Kingdom; [18]Cancer Epidemiology Unit, University of Oxford, Oxford, United Kingdom; [19]Department of Public Health and Primary Care, University of Cambridge, Addenbrooke's Hospital, Cambridge, United Kingdom; [20]German Cancer Research Center [DKFZ], Heidelberg, Germany; [21]Danish Cancer Society Research Center, Copenhagen, Denmark; [22]Public Health Division of Gipuzkoa, Biodonostia Research Institute, San Sebastian, Spain; [23]Department of Cardiology, Center for Cardiovascular Research, Aalborg Hospital, Aarhus University Hospital, Aalborg, Denmark; [24]Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network—ISPRO, Florence, Italy; [25]Dipartimento di Medicina Clinica e Chirurgia, Federico II University, Naples, Italy; [26]Public Health Directorate, Asturias, Spain; [27]Unit of Cancer Epidemiology, Citta'della Salute e della Scienza Hospital-University of Turin and Center for Cancer Prevention (CPO), Torino, Italy; [28]Human Genetics Foundation (HuGeF), Torino, Italy; [29]Andalusian School of Public Health, Granada, Spain; [30]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands; [31]National Institute for Public Health and the Environment [RIVM], Bilthoven, Netherlands; [32]Cancer Registry and Histopathology Unit, 'Civile – M.P. Arezzo' Hospital, Ragusa, Italy; [33]Associazone Iblea per la Ricerca Epidemiologica—Onlus, Ragusa, Italy; [34]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, United Kingdom; and [35]Institute of Nutritional Sciences, University of Potsdam, Nuthetal, Germany

---

## ABSTRACT

**Background:** Population-specificity of exploratory dietary patterns limits their generalizability in investigations with type 2 diabetes incidence.

**Objective:** The aim of this study was to derive country-specific exploratory dietary patterns, investigate their association with type 2 diabetes incidence, and replicate diabetes-associated dietary patterns in other countries.

**Methods:** Dietary intake data were used, assessed by country-specific questionnaires at baseline of 11,183 incident diabetes cases and 14,694 subcohort members (mean age 52.9 y) from 8 countries, nested within the European Prospective Investigation into Cancer and Nutrition study (mean follow-up time 6.9 y). Exploratory dietary patterns were derived by principal component analysis. HRs for incident type 2 diabetes were calculated by Prentice-weighted Cox proportional hazard regression models. Diabetes-associated dietary patterns were simplified or replicated to be applicable in other countries. A meta-analysis across all countries evaluated the generalizability of the diabetes-association.

**Results:** Two dietary patterns per country/UK-center, of which overall 3 dietary patterns were diabetes-associated, were identified. A risk-lowering French dietary pattern was not confirmed across other countries: pooled HR$_{France}$ per 1 SD: 1.00; 95% CI: 0.90, 1.10. Risk-increasing dietary patterns, derived in Spain and UK-Norfolk, were confirmed, but only the latter statistically significantly: HR$_{Spain}$: 1.09; 95% CI: 0.97, 1.22 and HR$_{UK-Norfolk}$: 1.12; 95% CI: 1.04, 1.20. Respectively,

this dietary pattern was characterized by relatively high intakes of potatoes, processed meat, vegetable oils, sugar, cake and cookies, and tea.

**Conclusions:** Only few country/center-specific dietary patterns (3 of 18) were statistically significantly associated with diabetes incidence in this multicountry European study population. One pattern, whose association with diabetes was confirmed across other countries, showed overlaps in the food groups potatoes and processed meat with identified diabetes-associated dietary patterns from other studies. The study demonstrates that replication of associations of exploratory patterns with health outcomes is feasible and a necessary step to overcome population-specificity in associations from such analyses. *J Nutr* 2019;149:1047–1055.

**Keywords:** dietary patterns, principal component analysis, diet-disease association, type 2 diabetes mellitus, replication, meta-analysis

## Introduction

Numerous studies have determined the association between diet and incident type 2 diabetes and have summarized evidence of the properties of specific single food groups that promote or reduce the risk. For instance, processed meat intake and sugar-sweetened beverages have been associated with greater diabetes risk and wholegrain intake with lower diabetes risk (1–4). However, single food group investigations do not consider the complexity of human diet and the potential interplay of nutrients (5). To address these limitations, dietary patterns can serve as an alternative approach to examine which combinations of foods potentially contribute to development of noncommunicable diseases such as type 2 diabetes. If no previous knowledge about overall dietary patterns in a particular population is available, exploratory methods such as principal component analysis (PCA) and cluster analysis, or mixed approaches such as reduced rank regression, give a first insight. However, these approaches result by definition in population-specific dietary patterns with likely limited comparability (6). Indeed, when we recently summarized prospective studies on exploratory patterns and diabetes risk (7), heterogeneity in the pattern structure limited attempts to meta-analyze evidence. Given that single-study findings are unlikely to inform dietary recommendations, the lack of replication of dietary pattern-disease associations remains a major limitation. This might be particularly challenging for European populations with a high degree of heterogeneity of diets (8, 9).

Although methodological approaches to replicate diet-disease associations for exploratory patterns exist and have been used for patterns derived by reduced rank regression (7, 10), we are not aware of studies that have replicated population-specific patterns from factor analysis or PCA. Because methodological differences in dietary assessments might impact the ability to replicate dietary pattern-disease associations (11), multicenter studies across different populations with standardized dietary assessment and harmonized data to minimize heterogeneity would be very useful in this context. The EPIC (European Prospective Investigation into Cancer and Nutrition)-InterAct study provides a promising setting, as for the collection of dietary information, special attempts were made to harmonize the data across the countries (12). Thus, the study sets a perfect starting point to derive exploratory dietary patterns and to the potential for replication. Hence, the aim of the current investigation was the identification of country-specific dietary patterns with PCA and their association with type 2 diabetes risk. Subsequently, the potential for

replication of identified diabetes-associated dietary patterns across all EPIC-InterAct countries was evaluated to overcome the population-specificity of exploratory dietary patterns.

## Methods

### Study population and design

The EPIC-InterAct study is a case-cohort study, nested within the prospective EPIC study (12). Between 1992 and 2000, more than a half million ($n = 519,978$) participants with an age range of 35–70 y were enrolled for the EPIC study. Participants were recruited from the general population despite some exceptions (13–15). Written informed consent was provided by each participant and the study was approved by local ethics committees and the Internal Review Board of the International Agency for Research on Cancer (14). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

The EPIC-InterAct case-cohort was established based on the incident cases of type 2 diabetes in the full (EPIC) cohort, which occurred between 1991 and 2007 in 8 of 10 participating countries (except for Norway and Greece) and a randomly drawn subcohort (12). From 455,680 eligible participants, those without stored blood samples ($n = 109,625$) and reported diabetes status ($n = 5821$) were excluded, resulting in 340,234 participants. A subcohort of 16,835 participants was randomly selected, stratified by center. After exclusion of post-censoring diabetes cases and individuals with unknown status ($n = 681$), the subcohort comprised 16,154 participants (**Supplemental Figure 1**).

Across the EPIC centers, different sources of evidence were used for ascertainment of incident cases of type 2 diabetes, including self-report, linkage to primary-care or secondary-care registers, medication use (drug registers), hospital admissions, or mortality data. Any of these sources of information was acceptable, where the information was obtained after the date of baseline examination. In Denmark and Sweden, diabetes cases were identified via local and national diabetes and pharmaceutical registers and therefore were considered to be verified. In other centers, if information on diabetes status was ascertained from only 1 source of evidence, further evidence to specify the definition for these cases was sought by including reviews of medical records. Censoring of the follow-up was either done at the date of diagnosis, 31 December 2007 or the date of death, depending on the first occurrence (12).

For the group of ascertained type 2 diabetes cases ($n = 17,928$), exclusion criteria similar to the subcohort were applied and further self-reports of diabetes in Denmark and nondiabetic participants were excluded ($n = 5525$), resulting in 12,403 verified type 2 diabetes cases (12).

In the Swedish study center Umeå, information on certain food groups (vegetables, dairy, other fruits, meat, offal, eggs, and vegetable oils) was not available, which was needed for the principal component analysis. Therefore, data from this study center were excluded ($n = 1845$). Further exclusions on missing food groups, anthropometry, and lifestyle factors such as smoking or physical activity were applied ($n = 776$), resulting in a subcohort of 14,694 participants and 11,183 verified diabetes cases. Because of the random selection of the subcohort members, 719 type 2 diabetes cases overlapped with the subcohort (Supplemental Figure 1). Excluded participants ($n = 2621$) were slightly younger, more likely to be men, with a higher BMI and waist circumference, less physically active, with a lower proportion of highly educated participants, a higher proportion of family history of diabetes and higher glycated hemoglobin (HbA1c), but a lower total energy intake (**Supplemental Table 1**).

### Dietary assessment

Country-specific dietary questionnaires, developed and locally validated in a series of studies (16, 17), were used to assess the usual food intake over the previous 12 mo. Although the reproducibility was generally good, the validity ranged from modest to good (17).

Food items were aggregated in a common food classification system (**Supplemental Table 2**), which was based on the EPIC-Soft classification system (18), but provided a higher level of detail (more subgroups) to be used for dietary pattern analyses. For the purpose of an EPIC-InterAct wide application of principal component analysis, the food groups "sauces," "soups," and "miscellaneous" were not considered as they were not available in all EPIC centers. Total energy intake and intake of specific nutrients were derived from the "EPIC Nutrient Database" (19).

### Assessment of covariates

Standardized self-report questionnaires on sociodemographic and lifestyle information such as age, education, smoking history, physical activity, and history of previous illnesses were used in all EPIC countries (14).

Standardized protocols were used to measure anthropometric data including height, weight, waist and hip circumference. In France and Oxford, self-reports of all 4 anthropometric variables were obtained and adjusted by measured values from a subset of participants (14).

### Statistical methods

Baseline characteristics of the participants in the subcohort were stratified by country and presented as mean ± SD, if normally distributed, median and interquartile range, if not normally distributed, and for categorical variables as relative percentages. The UK cohorts from Norfolk (population-based) and Oxford (high proportion of vegans, vegetarians, and health-conscious individuals) were considered separately. Country-specific dietary patterns in the subcohort were derived with PCA, considering the eigenvalue >1 criterion, scree plot, and interpretability to decide for the final number of principal components (20). Given the assumption that each food group has a variance of 1 in the PCA, retaining only those principal components with an eigenvalue (variance explained by 1 principal component) >1 reduces the dimensionality. In a scree plot these eigenvalues were plotted against all principal components (number equals the number of food groups) to visualize a possible "scree" in the top-to-ground slope. Principal components above the scree could be identified as explaining the majority of variance. Subsequently, the interpretability criterion (defined as ≥3 food groups with absolute factor loadings ≥0.4) was applied on these retained principal components to identify only those that reflect the complexity of a dietary pattern. In PCA, factor loadings can be interpreted as correlations between the food groups and resulting principal components (20). Then, dietary pattern scores were calculated for subcohort participants in each country as the sum of country-specific weights ($\beta_i$, i = 1, ..., m) equivalent to factor loadings, multiplied by food groups ($X_i$, i = 1, ..., m) (standardized to respective country-specific subcohort distributions) (Equation 1) (Step 1 in **Figure 1**).

$$Dietary\ pattern\ score\ =\ \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_m X_m \qquad (1)$$

Subsequently, dietary pattern scores were calculated for the cases external to the subcohort in the respective country with use of standardized food groups according to Equation 1. Cox proportional hazard regression models, weighted according to Prentice (21) and stratified by center and integers of age (y), were performed to investigate the association between the dietary patterns and the risk of type 2 diabetes within the whole case-cohort in each country (Step 2 in Figure 1). Two levels of adjustment were applied: Model 1 was adjusted for sex, physical activity [classified into "inactive," "moderately inactive," "moderately active," and "active" according to the validated Cambridge Physical Activity Index (22)], educational level (none, primary, technical/professional, secondary, university), smoking status (never, former, current), and total energy intake (continuous). Model 2 was also adjusted for BMI (continuous) and waist circumference (continuous).

To facilitate the applicability of the identified diabetes-associated country-specific dietary patterns to other study populations, a simplification approach (6) was applied. For this purpose, simplified sum scores were constructed by summing only those unweighted food groups ($X_i$,

| | Aim | Method | Study population | Result |
|---|---|---|---|---|
| **Step 1** | Identify number and composition of country-specific dietary patterns | Principal component analysis (eigenvalue>1, scree plot, interpretability) | Subcohort in each country | Country-specific dietary patterns |
| **Step 2** | Identify significantly associated dietary patterns | Calculation of standardized food intake for external cases according to country-specific subcohort | Cases external to the subcohort in each country | Country-specific T2D-associated dietary patterns |
| | | Prentice-weighted Cox proportional hazard regression models | Whole case-cohort in each country | |
| **Step 3** | Test for potential simplification of country-specific T2D-associated dietary patterns in each country subcohort | Calculation of: $simplified\ sum\ score^* = X_1 + X_2 + \cdots X_m, \mathrm{m} < \mathrm{n}$ | Subcohort in each country | Decision for/against simplification  if $\mathrm{r} \geq 0.90$, then simplified sum score **(A)**    if $\mathrm{r} < 0.90$, then: $replicative\ score^{**} = \beta_1 X_1 + \beta_2 X_1 + \cdots + \beta_n X_n$ **(B)** |
| | | Correlation with original dietary pattern score | Subcohort in each country | |
| **Step 4** | Score calculation | Calculation of standardized food intake according to the distribution in the full subcohort | Whole EPIC-InterAct case-cohort | Scores for all country-specific T2D-associated dietary patterns in the whole EPIC-InterAct case-cohort |
| | | Calculation of the scores according to **(A)** or **(B)** | Whole EPIC-InterAct case-cohort | |
| **Step 5** | Confirmation of the T2D-association | Prentice-weighted Cox proportional hazard regression models | Whole EPIC-InterAct case-cohort | Pooled hazard ratios for T2D-associated dietary patterns across all countries |
| | | Meta-analysis | Whole EPIC-InterAct case-cohort | |

**FIGURE 1** Scheme of the statistical steps to investigate the association between dietary patterns and type 2 diabetes risk. Subcohort in each country = randomly drawn subcohort in each country including noncases and cases of type 2 diabetes; whole case-cohort in each country = randomly drawn subcohort in each country and type 2 diabetes cases external to the subcohort; whole EPIC-InterAct case-cohort = sum of all randomly drawn subcohorts ($n = 14,694$) and external type 2 diabetes cases ($n = 11,183$), with an overlap of $n = 719$ verified incident type 2 diabetes cases in the subcohort, across all included EPIC-InterAct countries. *Sum of a reduced number ($m < n$) of standardized, unweighted food groups characterized by high factor loadings in the original dietary pattern score. **Sum of all 36 standardized food groups ($X_n$) multiplied by standardized scoring coefficients ($\beta_n$). EPIC, European Prospective Investigation into Cancer and Nutrition.

$i = 1, ..., n$), which had high factor loadings in the original country-specific dietary patterns (6) (Equation 2). The food groups were again standardized to the country-specific subcohort distribution.

$$Simplified\ sum\ score = X_1 + X_2 + X_3 + \ldots + X_n\ (n < m) \qquad (2)$$

Because no common cutoff has been defined as "high" factor loading, we applied 4 different cutoffs for absolute factor loadings $\geq 0.4$, $\geq 0.3$, $\geq 0.2$, and $\geq 0.1$ to calculate simplified sum scores. Depending on the chosen cutoff for the factor loading, the number of food groups (n) to sum up differed, but was always smaller than all 36 food groups (m). The resulting simplified sum scores were standardized according to the distribution of the country-specific subcohorts. The original dietary pattern score (Equation 1) was sufficiently reflected by the simplified sum score (Equation 2), if Spearman correlation coefficients of $r \geq 0.90$ were achieved. However, if $r < 0.90$ within the respective country, we did not consider the use of simplified sum scores, but calculated replicative scores in the other countries instead: sum of the product of all 36 standardized food groups and PCA-derived standardized scoring coefficients according to Equation 1 (Step 3 in Figure 1). Both for the calculation of simplified sum scores and replicative scores, food groups were standardized according to the distribution of the full subcohort (Step 4 in Figure 1).

To investigate the association between the simplified sum scores or the replicative scores with incident type 2 diabetes, Prentice-weighted Cox proportional hazard regression models (Model 2 in Step 2) were calculated. The HRs from each country were subsequently meta-analyzed to evaluate whether the association with type 2 diabetes could be confirmed in other countries (Step 5 in Figure 1).

The influence of potential energy-misreporting was investigated by excluding participants in the top and bottom 1% of the energy intake/energy requirement ratio. To account for a change in diet from a chronic disease, participants reported to have a cardiovascular disease (angina, stroke, or myocardial infarction) at baseline or participants who developed incident type 2 diabetes within the first 2 y of follow-up or had HbA1c values $\geq 6.5\%$ [measured in the erythrocyte faction of the blood serum samples stored at $-196°$C with the Tosoh-G8 HPLC analyzer (23)] were excluded. Potential confounding by a history of diabetes in first-degree relatives was addressed by further adjusting for it (information not available in all centers).

Statistical analyses were performed with the software packages Statistical Analysis System (SAS) Enterprise Guide 6.1 with SAS version 9.4 (SAS Institute Inc.). Meta-analyses were undertaken with Cochrane Software package Review Manager 5.3.

**TABLE 1** Country-specific baseline characteristics of the EPIC-InterAct subcohort ($n = 14,694$)[1]

| Characteristics | France ($n = 562$) | Italy ($n = 1927$) | Spain ($n = 3509$) | UK-Norfolk ($n = 900$) | UK-Oxford ($n = 316$) | Netherlands ($n = 1398$) | Germany ($n = 2044$) | Sweden ($n = 1919$) | Denmark ($n = 2119$) |
|---|---|---|---|---|---|---|---|---|---|
| Age, y | 56.3 ± 6.5 | 50.2 ± 7.9 | 49.1 ± 7.8 | 59.5 ± 9.4 | 50.2 ± 11.6 | 52.6 ± 10.8 | 50.3 ± 8.6 | 57.8 ± 7.7 | 56.6 ± 4.4 |
| Men, % | 0 | 32.3 | 38.4 | 42.7 | 27.9 | 16.7 | 41.5 | 40.3 | 53.7 |
| BMI, kg/m$^2$ | 23.1 ± 3.7 | 25.8 ± 4.1 | 28.2 ± 4.2 | 25.7 ± 3.7 | 24.3 ± 3.9 | 25.3 ± 3.9 | 25.8 ± 4.1 | 25.3 ± 3.9 | 26.0 ± 3.9 |
| WC, cm | | | | | | | | | |
| Men | — | 91.9 ± 9.7 | 99.2 ± 8.7 | 92.7 ± 9.6 | 89.0 ± 11.4 | 91.3 ± 11.2 | 94.4 ± 9.7 | 93.4 ± 10.3 | 95.9 ± 9.8 |
| Women | 76.0 ± 9.0 | 79.8 ± 10.6 | 87.1 ± 11.0 | 79.3 ± 9.9 | 74.3 ± 8.7 | 81.1 ± 10.3 | 79.8 ± 11.1 | 77.8 ± 10.6 | 81.6 ± 11.1 |
| Physically active, % | 9.6 | 14.8 | 11.4 | 14.8 | 15.2 | 40.6 | 20.3 | 16.4 | 35.4 |
| Never smoking, % | 64.6 | 45.0 | 54.9 | 46.9 | 58.9 | 40.2 | 47.1 | 39.1 | 33.7 |
| Post-secondary education, % | 39.9 | 14.5 | 11.3 | 12.3 | 42.4 | 21.7 | 34.9 | 22.4 | 20.4 |

[1] Data are shown as means ± SDs or as relative percentages. EPIC, European Prospective Investigation into Cancer and Nutrition; WC, waist circumference.

## Results

### Characterization of the study population

Country-specific baseline characteristics of the EPIC-InterAct subcohort ($n = 14,694$) are shown in **Table 1**. The participants were middle-aged, with wide ranges for BMI and waist circumference, and with a rather low percentage of participants being physically active and a low percentage of participants having post-secondary education in most EPIC-InterAct countries.

Relative contributions of macronutrients to the total energy intake were comparable across the countries. Except for few food groups, consistently consumed in small amounts in all countries, for example, "other fruits" and nuts, food intake largely varied between countries. Some food groups appeared to represent a country-specific intake, for example, a 3-fold higher intake of legumes in Spain compared to other countries or "other cereals" (including flour, breakfast cereals, and salty biscuits, among others) in the 2 UK centers. By contrast, some countries were characterized by a very low intake of specific food groups, for example, intake of leafy vegetables in Denmark (**Supplemental Table 3**).

### Country-specific dietary patterns and their association with type 2 diabetes risk

In each EPIC-InterAct country, 2 dietary patterns were identified with PCA according to the defined criteria (**Supplemental Table 4**). The structure of dietary patterns, thus the factor loadings of the 36 included food groups, showed substantial heterogeneity between countries. We subsequently evaluated these dietary patterns within the countries in which they were derived. Of the 18 identified dietary patterns, 3 were significantly associated with diabetes risk in the most adjusted models (**Table 2**): in Norfolk, the dietary pattern 2 explained 7.5% and was characterized by high factor loadings of potatoes, processed meat, vegetable oils, sugar, cakes and cookies, and tea (**Supplemental Figure 2**). This pattern was statistically significantly associated with an increased diabetes risk (HR: 1.41; 95% CI: 1.19, 1.67), although the risk increase was attenuated in Model 2 (HR: 1.24; 95% CI: 1.02, 1.51) (**Table 2**). The dietary pattern 2 in France explained 8.3% of variance in the food groups and was characterized by high factor loadings of nuts, other fruits, processed meat, fish, eggs, cake and cookies, coffee, and other alcoholic beverages (**Supplemental Figure 3**). It

**TABLE 2** HR for the 2 derived dietary patterns per each country[1]

| | | Country/center-specific dietary pattern 1 | | Country/center-specific dietary pattern 2 | |
|---|---|---|---|---|---|
| | | HR (95% CI) | P | HR (95% CI) | P |
| France ($n = 828$) | Model 1 | 1.14 (0.99, 1.31) | 0.08 | 1.15 (0.92, 1.42) | 0.22 |
| | Model 2 | 1.06 (0.90, 1.26) | 0.49 | 0.64 (0.49, 0.85) | 0.002 |
| Italy ($n = 3193$) | Model 1 | 1.22 (1.11, 1.35) | <0.0001 | 1.13 (1.01, 1.26) | 0.04 |
| | Model 2 | 1.10 (0.98, 1.23) | 0.10 | 1.01 (0.89, 1.14) | 0.93 |
| Spain ($n = 5766$) | Model 1 | 1.32 (1.19, 1.46) | <0.0001 | 0.99 (0.93, 1.06) | 0.86 |
| | Model 2 | 1.14 (1.03, 1.27) | 0.02 | 1.02 (0.95, 1.09) | 0.67 |
| UK-Norfolk ($n = 1605$) | Model 1 | 0.89 (0.78, 1.02) | 0.09 | 1.41 (1.19, 1.67) | <0.0001 |
| | Model 2 | 0.89 (0.77, 1.03) | 0.11 | 1.24 (1.02, 1.51) | 0.03 |
| UK-Oxford ($n = 536$) | Model 1 | 0.92 (0.74, 1.14) | 0.44 | 1.61 (1.25, 2.07) | 0.0002 |
| | Model 2 | 1.11 (0.88, 1.39) | 0.38 | 1.22 (0.94, 1.60) | 0.14 |
| Netherlands ($n = 2119$) | Model 1 | 1.25 (1.09, 1.45) | 0.0022 | 1.02 (0.90, 1.15) | 0.74 |
| | Model 2 | 1.10 (0.93, 1.29) | 0.27 | 0.92 (0.79, 1.06) | 0.24 |
| Germany ($n = 3553$) | Model 1 | 1.00 (0.92, 1.09) | 0.99 | 1.20 (1.08, 1.34) | 0.0008 |
| | Model 2 | 0.97 (0.88, 1.07) | 0.55 | 1.08 (0.96, 1.21) | 0.19 |
| Sweden ($n = 3539$) | Model 1 | 1.08 (0.97, 1.21) | 0.18 | 1.03 (0.96, 1.11) | 0.39 |
| | Model 2 | 1.08 (0.95, 1.23) | 0.25 | 1.00 (0.91, 1.09) | 0.91 |
| Denmark ($n = 4019$) | Model 1 | 0.92 (0.85, 0.99) | 0.02 | 1.30 (1.15, 1.46) | <0.0001 |
| | Model 2 | 0.98 (0.90, 1.06) | 0.60 | 1.08 (0.94, 1.24) | 0.26 |

[1] Model 1 was adjusted for sex, physical activity, smoking, education, and total energy intake; Model 2 was further adjusted for BMI and waist circumference.
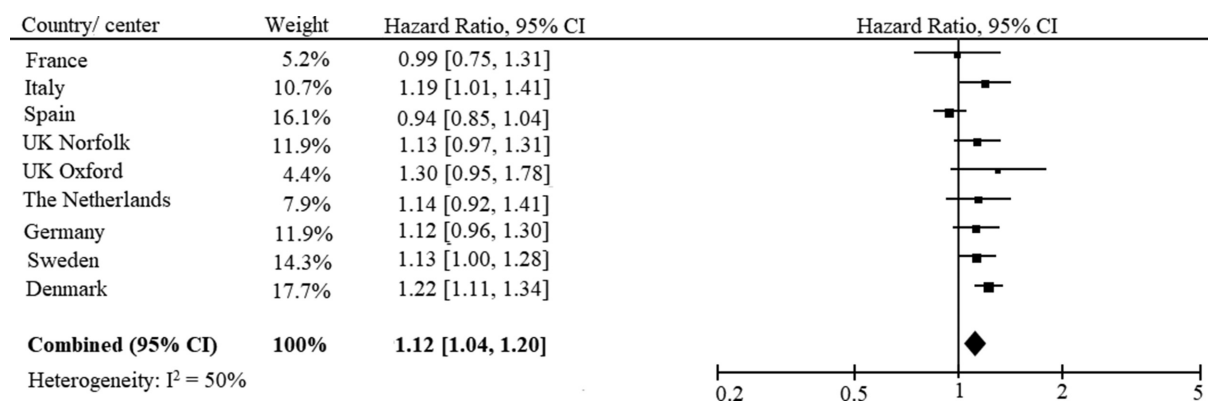
**FIGURE 2** Meta-analysis of HR and 95% CI for the risk of type 2 diabetes per 1 SD of the "Replicative Norfolk" score across all EPIC-InterAct countries. EPIC, European Prospective Investigation into Cancer and Nutrition.

was not associated with risk of diabetes in Model 1 (HR: 1.15; 95% CI: 0.92, 1.42); however, adjustment for anthropometric markers (Model 2) indicated an inverse association (HR: 0.64; 95% CI: 0.49, 0.85 ) (Table 2). The dietary pattern 1 in Spain was characterized by high contributions of potatoes, legumes, bread, red meat, processed meat, eggs, vegetable oils, wine and spirits, and explained 9.7% of the total variance in the food groups (**Supplemental Figure 4**). It was associated with an increased diabetes risk in Model 1 (HR: 1.32; 95% CI: 1.19, 1.46), whereas the HR was markedly attenuated, when also adjusted for anthropometric markers (HR: 1.14; 95% CI: 1.03, 1.27) (Table 2).

**Replication of selected country-specific dietary patterns across other countries**

Those dietary patterns significantly associated with type 2 diabetes in Model 2 (Table 2), were subsequently replicated in all other countries. As a first step, the Spearman correlation coefficients of the simplified sum scores (by step-wise lowering the cutoff for the factor loadings) with the original dietary patterns were moderate in France ($r$ = 0.59–0.69) and Norfolk ( $r$ = 0.66–0.72) (**Supplemental Table 5**) and judged not to sufficiently reflect the original dietary patterns. Hence, replicative scores were calculated instead. In contrast, the simplified sum score in the Spanish subcohort showed good correlations with the original scores ranging from $r$ = 0.90–0.93. As lowering the cutoff of absolute factor loadings from ≥0.3 to ≥0.2 did not result in a markedly improved correlation, but in a highly increased number of food groups, the simplified sum score consisting of those foods with loadings ≥0.3 (potatoes, legumes, bread, red meat, processed meat, eggs, vegetable oils, wine, spirits, and pasta and rice) was considered.

Subsequently, the simplified and replicative scores were calculated in all EPIC-InterAct countries and evaluated with regard to diabetes risk. The "Replicative Norfolk" score was associated with an increased diabetes risk in the meta-analysis (**Figure 2**) (pooled HR for 1 SD increment, model 2 adjustments: 1.12; 95% CI: 1.04, 1.20). Inconsistency between the countries was existent ($I^2$ = 50%), although lower than for the other 2 replicated scores. With the exception of France and Spain, the positive direction of an association with diabetes risk in all other countries could be replicated, although not always significantly. To investigate whether the pooled estimate in the meta-analysis was driven by few countries, we excluded each country in a sensitivity analysis. The exclusion of Spain resulted in a strengthened association (HR: 1.16; 95% CI: 1.10,

1.23), whereas the exclusion of Denmark attenuated the pooled estimate (HR: 1.09; 95% CI: 1.02, 1.18). Hence, the overall clear positive association remained.

For the "Replicative France" score, the meta-analysis resulted in a pooled HR: 1.00; 95% CI: 0.90, 1.10 (**Supplemental Figure 5**), thus not confirming the inverse association of the Component 2 observed initially in France. The inconsistency measure indicated high heterogeneity between countries ($I^2$ = 67%). A clear inverse association with type 2 diabetes risk was exclusively observed in France, whereas a clear risk increase of 44% was observed in UK-Norfolk.

For the "Simple Spain" score the pooled HR: 1.09; 95% CI: 0.97, 1.22 and inconsistency between the countries was high ($I^2$ = 74%) (**Supplemental Figure 6**). A higher diabetes risk was observed in UK-Norfolk, besides Spain. Although in the other countries the simplified score was not significantly associated with incident type 2 diabetes, a clear inverse association was observed in Italy. A sensitivity analysis was conducted to investigate whether the exclusion of wine would result in a stronger positive association with diabetes risk. However, the pooled estimate was not substantially changed (HR: 1.08; 95% CI: 0.98, 1.20).

We subsequently applied several sensitivity analyses to evaluate the robustness of our findings. Excluding participants with a history of cardiovascular diseases, for example, myocardial infarction, stroke, or angina ($n$ = 1195), who developed type 2 diabetes within the first 2 y of follow-up and with baseline HbA1c ≥6.5% ($n$ = 2492), did not change the overall results [pooled HR (Norfolk): 1.11; 95% CI: 1.02, 1.20, $I^2$ = 42%; pooled HR (France): 0.98; 95% CI: 0.88, 1.09, $I^2$ = 75%; pooled HR (Spain): 1.10; 95% CI: 0.98, 1.23, $I^2$ = 75%]. Exclusion of participants in the top and bottom 1% of the energy intake/energy requirement ratio ($n$ = 551) did not alter the results [pooled HR (Norfolk): 1.10; 95% CI: 1.05, 1.15, $I^2$ = 14%; pooled HR (France): 1.00; 95% CI: 0.91, 1.09, $I^2$ = 69%; pooled HR (Spain): 1.08; 95% CI: 0.97, 1.21, $I^2$ = 74%]. Adjusting for family history of type 2 diabetes ($n$ = 3102) in countries where information was ascertained, gave the same diabetes-associated dietary patterns as in the main analyses.

## Discussion

In this European case-cohort study, 2 dietary patterns per country were identified by PCA with highly specific structures

of contributing food groups. Only a small proportion (3 of 18 dietary patterns) was significantly associated with diabetes risk (1 in France, 1 in Spain, 1 in UK-Norfolk). Of these dietary patterns, the positive association of the UK-Norfolk pattern could potentially be replicated across other countries in the EPIC-InterAct study. In contrast, no meaningful association could be established for the pattern from France, given the heterogeneous findings in terms of direction of association and overall null association from the meta-analysis. For the Spanish pattern, the positive association with diabetes risk could not be replicated, although in 5 of 8 countries the HRs pointed towards a positive association, even if the significance level was not reached.

The "Replicative Norfolk" score was quite consistently associated with an increased diabetes risk across most countries, as confirmed by the sensitivity analyses, where the overall association was not mainly driven by a single country contribution. Food groups with high positive factor loadings were potatoes, processed meat, vegetable oils, sugar, cake and cookies, and tea. Among these, evidence for risk-increasing properties exists for processed meat (24) and partly for potatoes (including French fries) (25). In previous EPIC-InterAct analyses, cakes and cookies, as well as vegetable oils, were not associated with diabetes risk when adjusted for BMI (26). Olive oil intake, in particular, contributed highly to lower diabetes risk (27). However, vegetable oil intake in non-Mediterranean countries is potentially driven by contributions of oils other than olive oil (26). Overall, these data suggest that vegetable oils are unlikely to explain the positive association of the "Replicative Norfolk" score with diabetes risk. For tea consumption, a systematic review summarized an inverse association in Asian studies, but no association in European and American studies (28). However, a previous publication in EPIC-InterAct showed an inverse association between the consumption of ≥4 cups of tea/d and diabetes risk (29). Our food group "sugar" included confectionary, honey, jams, chocolate, syrups, and ice cream. This heterogeneous mix of food items is characterized by a high content of monosaccharides and disaccharides. In terms of diabetes prevention, evidence for sugar from solid sources on an association remains inconsistent (30). Hence, evidence for diabetes risk increasing properties of single food groups is limited to processed meat and potatoes. Still, the replicated positive association could possibly be caused in part by other contributing food groups, specifically the group "sugar." It is also possible that food groups such as cakes and cookies, vegetable oils, and tea have rather diluted detrimental effects of more harmful food groups.

The replication of the "Simple Spain" score resulted in no significant association with diabetes risk. Exclusions of single countries in a sensitivity analysis resulted in no major changes of the pooled estimate. The pattern was characterized by potatoes, legumes, bread, red meat, processed meat, eggs, vegetable oils, wine, spirits, pasta and rice. For the food groups potatoes, red and processed meat (as discussed earlier), eggs (31), bread (mainly white bread), and pasta and rice (32, 33), recent studies underline greater risk. Thus, evidence for the majority of contributing food groups to this dietary pattern supports diabetes risk-increasing properties. Still, such positive association could be counterbalanced by moderate alcohol consumption, which is inversely associated with diabetes risk (34). However, the exclusion of wine from the "Simple Spain" score in a sensitivity analysis did not result in a stronger positive association.

The diabetes-association of the "Replicative France" score, characterized by contributions of nuts, "other fruits," processed meat, fish, eggs, cake and cookies, coffee, and "other alcoholic beverages," could not be replicated. For some components, evidence for an inverse association with diabetes risk is plausible, for example, nuts (34) and coffee (36), whereas processed meat and eggs were discussed before as diabetes risk factors. Evidence remains inconclusive in terms of intake of fish (37), whereas for "other fruits" evidence for no influence on diabetes risk is probable (38). "Other alcoholic beverages," including fortified wine, could potentially explain the inverse association in France, as adjusting for this food group resulted in an attenuated association (data not shown). We observed a markedly higher intake of this food group in the female French study population in comparison to other countries. In a previous EPIC-InterAct publication, moderate consumption of fortified wine, compared to low intake, was associated with a 32% lower risk of diabetes in women (39). As the intake was markedly lower in other countries, exposure differences for this beverage might have been too low to affect diabetes risk.

The "Replicative Norfolk" score, identified as the only pattern to be significantly associated with an increased diabetes risk, showed rather limited comparability in pattern composition with other exploratory patterns previously described in European studies (8, 9). A comparison with patterns from the United States resulted in overlaps for potatoes (including French fries), processed meat, sugar, and cake and cookies, but tea and vegetable oils remain specific for the "Replicative Norfolk" score (40–42). Taking into account the aim of exploratory methods to deduce dietary patterns explaining maximum variance in food groups, but not necessarily being associated with the outcome of interest, limited comparability was expected. This was also confirmed in a recent systematic literature review, where, despite few exploratory dietary patterns with partially overlapping food groups, many population-specific dietary patterns remained incomparable (7). Still, among these, some were associated with diabetes risk as well (8, 43, 44).

The "Simple Spain" score, by contrast, overlapped in several food groups (refined grains, eggs, red meat, processed meat, and potatoes including French fries), identified in diabetes-associated patterns from single-country studies (7).

To discuss strengths and limitations of this analysis, the grouping of food items for PCA and the assumption of a set of foods to constitute all relevant food items remains difficult to justify (45). However, attempts were made in EPIC-InterAct to harmonize food groups and facilitate their use in dietary pattern analyses. Nevertheless, we could not consider different preparation techniques or the influence of meal timing. Hence, we still assume country-specificity in the resulting dietary patterns (11). Furthermore, the overall variance explained in the food groups ranged between 7.5% and 9.7%. Hence, these identified diabetes-associated dietary patterns did not reflect a high proportion of variance in the food groups. To minimize subjective decisions in PCA, we applied the eigenvalue >1 criterion which is straightforward. However, the second criterion of an identifiable turning point in the scree plots to separate principal components explaining relatively large proportions of variance in the food groups, may be ambiguous. Consequently, an additional criterion was consistently applied: the interpretability criterion, defined as cutoff for the number of food groups (≥3) and their factor loadings (≥0.4) (20). Pattern scores reflect relative, rather than absolute, quantitative differences in intake in comparison to a population average.

Although the initial derivation of patterns was based on the food intake distribution of the respective country, we used the full subcohort distribution for replication purposes. If intake distributions between country-specific subcohorts and the full subcohort differed greatly, effect estimates in the meta-analysis could have limited comparability to the country-specific estimates, therefore limiting the internal validity. Furthermore, the replication of only those dietary patterns being significantly associated with diabetes risk could be considered as a restrictive approach and it cannot be ruled out that other than the 3 identified country-specific patterns would show meaningful associations, if evaluated across all EPIC-InterAct countries. However, a lack of statistical significance for other patterns was mostly accompanied by weak associations, which supports our procedure. As strength of the replication, different cutoffs for factor loadings were systematically applied to investigate the capability of the simplified sum scores to reflect the original dietary pattern. The capability was quantified as Spearman correlation coefficient of $r \geq 0.90$, an arbitrary cutoff but ensuring sufficient reflection of the dietary pattern. As the Spanish dietary pattern was the only successfully simplified pattern, the replication of the other 2 diabetes-related patterns was done by applying the original pattern structure. However, this specific approach was feasible in the study setting of EPIC-InterAct with harmonized food groups. In other study settings, replication of original pattern structures may be strongly limited by lacking concordant food items and food groups. Although the use of simplified pattern scores appears to be more suitable in such cases, investigators should be aware that simplified sum scores may not sufficiently reflect the original pattern scores.

To conclude from this analysis, of the 18 dietary patterns derived by PCA in EPIC-InterAct, only 3 patterns were significantly associated with risk of type 2 diabetes in country-specific analyses. Of these, the UK-Norfolk pattern, characterized by high intakes of potatoes, processed meat, vegetable oils, sugar, cake and cookies, and tea, was largely consistently associated with increased diabetes risk across 8 countries. A Spanish pattern may also be related to increased risk, although associations were more heterogeneous across countries. Still, the UK-Norfolk and the Spanish dietary patterns show some similarity in food group composition (potatoes, processed meat), which are also characteristic for previously reported diabetes-related exploratory dietary patterns. Our study furthermore demonstrates that replication of associations of exploratory patterns with health outcomes is feasible and a necessary step to focus on generalizable associations from such analyses.

## References

1. Ley SH, Hamdy O, Mohan V, Hu FB. Prevention and management of type 2 diabetes: dietary components and nutritional strategies. Lancet 2014;383:1999–2007.
2. InterAct Consortium, Bendinelli B, Palli D, Masala G, Sharp SJ, Schulze MB, Guevara M, van der A DL, Sera F, Amiano P, et al. Association between dietary meat consumption and incident type 2 diabetes: the EPIC-InterAct study. Diabetologia 2013;56:47–59.
3. InterAct Consortium, Romaguera D, Norat T, Wark PA, Vergnaud AC, Schulze MB, van Woudenbergh GJ, Drogan D, Amiano P, Molina-Montes E, et al. Consumption of sweet beverages and type 2 diabetes incidence in European adults: results from EPIC-InterAct. Diabetologia 2013;56:1520–30.
4. InterAct Consortium. Dietary fibre and incidence of type 2 diabetes in eight European countries: the EPIC-InterAct Study and a meta-analysis of prospective studies. Diabetologia 2015;58:1394–408.
5. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002;13:3–9.
6. 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.
7. Jannasch F, Kroger J, Schulze MB. Dietary patterns and type 2 diabetes: a systematic literature review and meta-analysis of prospective studies. J Nutr 2017;147:1174–82.
8. Bauer F, Beulens JW, van der A DL, Wijmenga C, Grobbee DE, Spijkerman AM, van der Schouw YT, Onland-Moret NC. Dietary patterns and the risk of type 2 diabetes in overweight and obese individuals. Eur J Nutr 2013;52:1127–34.
9. Montonen J, Knekt P, Harkanen T, Jarvinen R, Heliovaara M, Aromaa A, Reunanen A. Dietary patterns and the incidence of type 2 diabetes. Am J Epidemiol 2005;161:219–27.
10. Weikert C, Schulze MB. Evaluating dietary patterns: the role of reduced rank regression. Curr Opin Clin Nutr Metab Care 2016;19(5):341–6.
11. Imamura F, Jacques PF. Invited commentary: dietary pattern analysis. Am J Epidemiol 2011;173:1105–8; discussion 9–10.
12. InterAct Consortium, Langenberg C, Sharp S, Forouhi NG, Franks PW, Schulze MB, Kerrison N, Ekelund U, Barroso I, Panico S, et al. Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. Diabetologia 2011;54:2272–82.
13. Bingham S, Riboli E. Diet and cancer—the European Prospective Investigation into Cancer and Nutrition. Nat Rev Cancer 2004;4:206–15.
14. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5:1113–24.
15. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 1997;26 (Suppl 1):S6–14.
16. Riboli E, Elmstahl S, Saracci R, Gullberg B, Lindgarde F. The Malmo Food Study: validity of two dietary assessment methods for measuring nutrient intake. Int J Epidemiol 1997;26 (Suppl 1):S161–73.
17. Margetts BM, Pietinen P. European Prospective Investigation into Cancer and Nutrition: Validity Studies on Dietary Assessment Methods. Int J Epidemiol 1997;26:S1–189.

18. Slimani N, Deharveng G, Charrondiere RU, van Kappel AL, Ocke MC, Welch A, Lagiou A, van Liere M, Agudo A, Pala V, et al. Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. Comput Methods Programs Biomed 1999;58:251–66.

19. Slimani N, Deharveng G, Unwin I, Southgate DA, Vignat J, Skeie G, Salvini S, Parpinel M, Moller A, Ireland J, et al. The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. Eur J Clin Nutr 2007;61:1037–56.

20. Hatcher L. A step-by-step approach to using the SAS® system for factor analysis and structural equation modeling. Cary (NC): SAS Institute Inc.: Books by Users; 1998.

21. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 1986;73(1):1–11.

22. InterAct Consortium, Ekelund U, Palla L, Brage S, Franks PW, Peters T, Balkau B, Diaz MJ, Huerta JM, Agnoli C, et al. Physical activity reduces the risk of incident type 2 diabetes in general and in abdominally lean and obese men and women: the EPIC-InterAct Study. Diabetologia 2012;55:1944–52.

23. Forouhi NG, Koulman A, Sharp SJ, Imamura F, Kroger J, Schulze MB, Crowe FL, Huerta JM, Guevara M, Beulens JW, et al. Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct case-cohort study. Lancet Diabetes Endocrinol 2014;2(10):810–8.

24. Feskens EJ, Sluik D, van Woudenbergh GJ. Meat consumption, diabetes, and its complications. Curr Diab Rep 2013;13:298–306.

25. Borch D, Juul-Hindsgaul N, Veller M, Astrup A, Jaskolowski J, Raben A. Potatoes and risk of obesity, type 2 diabetes, and cardiovascular disease in apparently healthy adults: a systematic review of clinical intervention and observational studies. Am J Clin Nutr 2016;104:489–98.

26. Buijsse B, Boeing H, Drogan D, Schulze MB, Feskens EJ, Amiano P, Barricarte A, Clavel-Chapelon F, de Lauzon-Guillain B, Fagherazzi G, et al. Consumption of fatty foods and incident type 2 diabetes in populations from eight European countries. Eur J Clin Nutr 2015;69:455–61.

27. Romaguera D, Guevara M, Norat T, Langenberg C, Forouhi NG, Sharp S, Slimani N, Schulze MB, Buijsse B, Buckland G, et al. Mediterranean diet and type 2 diabetes risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) study: the InterAct project. Diabetes Care 2011;34:1913–8.

28. Yang J, Mao QX, Xu HX, Ma X, Zeng CY. Tea consumption and risk of type 2 diabetes mellitus: a systematic review and meta-analysis update. BMJ Open 2014;4:e005632.

29. InterAct Consortium, van Woudenbergh GJ, Kuijsten A, Drogan D, van der A DL, Romaguera D, Ardanaz E, Amiano P, Barricarte A, Beulens JW, et al. Tea consumption and incidence of type 2 diabetes in Europe: the EPIC-InterAct case-cohort study. PLoS One 2012;7:e36910.

30. Lean ME, Te Morenga L. Sugar and type 2 diabetes. Br Med Bull 2016;120:43–53.

31. Shin JY, Xun P, Nakamura Y, He K. Egg consumption in relation to risk of cardiovascular disease and diabetes: a systematic review and meta-analysis. Am J Clin Nutr 2013;98:146–59.

32. Barclay AW, Petocz P, McMillan-Price J, Flood VM, Prvan T, Mitchell P, Brand-Miller JC. Glycemic index, glycemic load, and chronic disease risk—a meta-analysis of observational studies. Am J Clin Nutr 2008;87:627–37.

33. Aune D, Norat T, Romundstad P, Vatten LJ. Whole grain and refined grain consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. Eur J Epidemiol 2013;28:845–58.

34. Baliunas DO, Taylor BJ, Irving H, Roerecke M, Patra J, Mohapatra S, Rehm J. Alcohol as a risk factor for type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2009;32:2123–32.

35. Afshin A, Micha R, Khatibzadeh S, Mozaffarian D. Consumption of nuts and legumes and risk of incident ischemic heart disease, stroke, and diabetes: a systematic review and meta-analysis. Am J Clin Nutr 2014;100:278–88.

36. Ding M, Bhupathiraju SN, Chen M, van Dam RM, Hu FB. Caffeinated and decaffeinated coffee consumption and risk of type 2 diabetes: a systematic review and a dose-response meta-analysis. Diabetes Care 2014;37:569–86.

37. Yanai H, Hamasaki H, Katsuyama H, Adachi H, Moriyama S, Sako A. Effects of intake of fish or fish oils on the development of diabetes. J Clin Med Res 2015;7:8–12.

38. Boeing H, Bechthold A, Bub A, Ellinger S, Haller D, Kroke A, Leschik-Bonnet E, Muller MJ, Oberritter H, Schulze M, et al. Critical review: vegetables and fruit in the prevention of chronic diseases. Eur J Nutr 2012;51:637–63.

39. Beulens JW, van der Schouw YT, Bergmann MM, Rohrmann S, Schulze MB, Buijsse B, Grobbee DE, Arriola L, Cauchi S, Tormo MJ, et al. Alcohol consumption and risk of type 2 diabetes in European men and women: influence of beverage type and body size The EPIC-InterAct study. J Intern Med 2012;272:358–70.

40. Malik VS, Fung TT, van Dam RM, Rimm EB, Rosner B, Hu FB. Dietary patterns during adolescence and risk of type 2 diabetes in middle-aged women. Diabetes Care 2012;35:12–8.

41. Qi L, Cornelis MC, Zhang C, van Dam RM, Hu FB. Genetic predisposition, Western dietary pattern, and the risk of type 2 diabetes in men. Am J Clin Nutr 2009;89:1453–8.

42. 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.

43. Erber E, Hopping BN, Grandinetti A, Park SY, Kolonel LN, Maskarinec G. Dietary patterns and risk for diabetes: the multiethnic cohort. Diabetes Care 2010;33:532–8.

44. Yu R, Woo J, Chan R, Sham A, Ho S, Tso A, Cheung B, Lam TH, Lam K. Relationship between dietary intake and the development of type 2 diabetes in a Chinese population: the Hong Kong Dietary Survey. Public Health Nutr 2011;14:1133–41.

45. Martinez ME, Marshall JR, Sechrest L. Invited commentary: factor analysis and the search for objectivity. Am J Epidemiol 1998;148:17–9.