

Published in final edited form as:

Eur J Clin Nutr. 2012 October ; 66(10): 1082–1092. doi:10.1038/ejcn.2012.85.

Fruit and vegetable intake and type 2 diabetes: EPIC-InterAct prospective study and meta-analysis

Andrew J Cooper, MPhil¹, Nita G Forouhi, PhD, MRCP, FFFPH¹, Zheng Ye, PhD¹, Brian Buijsse, PhD², Larraitz Arriola, MD, MSc^{3,4,5}, Beverley Balkau, PhD^{6,7}, Aurelio Barricarte, MD, PhD^{8,5}, Joline WJ Beulens, PhD⁹, Heiner Boeing, PhD², Frederike L Büchner, MSc¹⁰, Christina C Dahm, PhD^{11,12}, Blandine de Lauzon-Guillain, PhD^{13,7}, Guy Fagherazzi, PhD¹³, Paul W Franks, PhD^{14,15}, Carlos Gonzalez, MD, PhD¹⁶, Sara Grioni, BSc¹⁷, Rudolf Kaaks, PhD¹⁸, Timothy J Key, DPhil¹⁹, Giovanna Masala, MD, MPH²⁰, Carmen Navarro, MD, PhD, MSc^{21,5,22}, Peter Nilsson, MD, PhD²³, Kim Overvad, PhD^{11,12}, Salvatore Panico, MD²⁴, Jose Ramón Quirós, MD²⁵, Olov Rolandsson, MD, PhD¹⁵, Nina Roswall, PhD²⁶, Carlotta Sacerdote, PhD^{27,28}, María-José Sánchez, MD, PhD^{29,5}, Nadia Slimani, PhD³⁰, Ivonne Sluijs, PhD⁹, Annemieke MW Spijkerman, PhD¹⁰, Birgit Teucher, PhD¹⁸, Anne Tjønneland, Dr. Med. Sci³¹, Rosario Tumino, MD, MSc, DLSHTM^{32,33}, Stephen J Sharp, MSc¹, Claudia Langenberg, MD, PhD¹, Edith JM Feskens, PhD³⁴, Elio Riboli, MD, MPH, ScM³⁵, and Nicholas J Wareham, MD, PhD¹ The InterAct Consortium

⁽¹⁾MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Box 285, Cambridge, CB2 0QQ, United Kingdom ⁽²⁾Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany ⁽³⁾Public Health Division of Gipuzkoa, Basque Government, Av. Navarra 4, 20013 San Sebastian, Spain ⁽⁴⁾Instituto BIO-Donostia, Basque Government, Spain ⁽⁵⁾Consortium for Biomedical Research in Epidemiology and Public Health (CIBER Epidemiología y Salud Pública (CIBERESP)), Spain ⁽⁶⁾Inserm, CESP Centre for research in Epidemiology and Population Health, U1018: Epidemiology of diabetes, obesity and chronic kidney disease over the lifecourse, 16 Av Paul Vaillant Couturier, 94807 Villejuif cedex, France ⁽⁷⁾University Paris Sud 11, UMRS 1018, 94807 Villejuif cedex, France ⁽⁸⁾Navarre Public Health Institute, Leyre 15, 31003 Pamplona, Navarra, Spain ⁽⁹⁾Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Stratum 6.131, PO Box 85.500, 3508 GA Utrecht, The Netherlands ⁽¹⁰⁾National Institute for Public Health and the Environment (RIVM), PO Box 1, 3720 BA Bilthoven, The Netherlands ⁽¹¹⁾Department of Epidemiology, School of Public Health, Bartholins Allé 2, DK-8000 Aarhus C, Denmark ⁽¹²⁾Department of Cardiology, Aalborg Hospital, Aarhus University Hospital, Sdr. Skovvej 15, DK-9000 Aalborg, Denmark ⁽¹³⁾Inserm, CESP Centre for Research in Epidemiology and Population Health, U1018: Nutrition, Hormones and Women's Health, IGR, 114 rue Edouard Vaillant, 94805 Villejuif cedex, France ⁽¹⁴⁾Department of Clinical Sciences, Clinical Research Center, Malmö General Hospital (UMAS) Lund University, 20502 Malmö, Sweden ⁽¹⁵⁾Department of Public Health and Clinical Medicine, Umea University, 90187 Umea, Sweden ⁽¹⁶⁾Unit Nutrition, Environment and Cancer, Department of Epidemiology, Gran Via s/n km 2,7, L'Hospitalet de Llobregat, 08907 Barcelona, Spain ⁽¹⁷⁾Fondazione IRCCS Istituto Nazionale

Corresponding author and author for requests for reprints: Dr. Nita G. Forouhi, MRC Epidemiology Unit, Institute of Metabolic Science, Box 285, Addenbrooke's Hospital, Hills road, Cambridge, CB2 0QQ, United Kingdom. Tel: +44 (0)1223 769145. Fax: +44 (0)1223 330316. nita.forouhi@mrc-epid.cam.ac.uk.

Conflict of interest The authors declare no conflict of interest.

Author contributions Author contributions were as follows: AJC had access to all data for this study. AJC and NGF take responsibility for the manuscript contents. AJC and ZY analysed the data. AJC drafted the manuscript and NGF made extensive revisions to subsequent drafts. All authors have contributed to conception and design, and interpretation of data, revising the article critically for important intellectual content and final approval of the version to be published.

Tumori Milan, Via Venezian, 1, 20133 Milan, Italy ⁽¹⁸⁾German Cancer Research Centre (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany ⁽¹⁹⁾Cancer Epidemiology Unit, University of Oxford, Richard Doll Building, Nuffield Department of Clinical Medicine, Roosevelt Drive, Oxford OX3 7LF, United Kingdom ⁽²⁰⁾Molecular and Nutritional Epidemiology Unit, Cancer Research and Prevention Institute (ISPO), Via delle Oblate 4, 50141, Florence Italy ⁽²¹⁾Department of Epidemiology, Murcia Regional Health Authority, Ronda de Levante, 11. E30008 Murcia, España ⁽²²⁾Unit of Preventive Medicine and Public Health, School of Medicine, University of Murcia, Spain ⁽²³⁾University of Lund, Department of Clinical Sciences Medicine, University Hospital Scania, 205 02 Malmö, Sweden ⁽²⁴⁾Dipartimento di Medicina Clinica e Sperimentale, Università Federico II, Department of Clinical and Experimental Medicine, Federico II University, via Pansini 5 - 80131 Napoli, Italy ⁽²⁵⁾Consejería de Salud y Servicios Sanitarios, C/Ciriaco Miguel Vigil 9, 33006-Oviedo-Asturias ⁽²⁶⁾Department of Diet, Genes and Environment, Danish Cancer Society Research Center, Danish Cancer Society, Strandboulevarden 49, DK-2100 Copenhagen Ø, Denmark ⁽²⁷⁾Center for Cancer Prevention (CPO-Piemonte), Via Santena 7, 10126, Torino, Italy. ⁽²⁸⁾Human Genetic Foundation (HuGeF), Torino, Italy ⁽²⁹⁾Andalusian School of Public Health, Cuesta del Observatorio 4, Ap. Correos 2070, 18080 Granada, Spain ⁽³⁰⁾International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France ⁽³¹⁾Danish Cancer Society Research Center, Strandboulevarden 49, 2100 Copenhagen, Denmark ⁽³²⁾Cancer Registry and Histopathology Unit, "Civile - M.P. Arezzo" Hospital, Azienda Sanitaria Provinciale No 7, Via Dante Nr. 109, 97100 Ragusa, Italy ⁽³³⁾Associazione Iblea per la Ricerca Epidemiologica - Onlus, Piazza Ancione No 2, 97100, Ragusa (Italy) ⁽³⁴⁾Division of Human Nutrition - Section Nutrition and Epidemiology, University of Wageningen, PO Box 8129, 6700 EV Wageningen, The Netherlands ⁽³⁵⁾Department of Epidemiology & Biostatistics, School of Public Health, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG

Abstract

Background/Objective—Fruit and vegetable intake (FVI) may reduce the risk of type 2 diabetes (T2D), but the epidemiological evidence is inconclusive. The aim of this study is to examine the prospective association of FVI with T2D and conduct an updated meta-analysis.

Subjects/Methods—In the EPIC-InterAct (European Prospective Investigation into Cancer-InterAct) prospective case-cohort study nested within eight European countries, a representative sample of 16 154 participants and 12 403 incident cases of T2D were identified from 340 234 individuals with 3.99 million person-years of follow-up. For the meta-analysis we identified prospective studies on FVI and T2D risk by systematic searches of MEDLINE and EMBASE until April 2011.

Results—In EPIC-InterAct, estimated FVI by dietary questionnaires varied more than two-fold between countries. In adjusted analyses the hazard ratio (95% confidence interval) comparing the highest with lowest quartile of reported intake was 0.90 (0.80-1.01) for FVI; 0.89 (0.76-1.04) for fruit, and 0.94 (0.84-1.05) for vegetables. Among FV sub-types, only root vegetables were inversely associated with diabetes 0.87 (0.77-0.99). In meta-analysis using pooled data from five studies including EPIC-InterAct, comparing the highest with lowest category for FVI was associated with a lower relative risk of diabetes (0.93 (0.87-1.00)). Fruit or vegetables separately were not associated with diabetes. Among FV sub-types, only green leafy vegetable intake (RR: 0.84 (0.74-0.94)) was inversely associated with diabetes.

Conclusions—Sub-types of vegetables, such as root vegetables or green leafy vegetables may be beneficial for the prevention of diabetes, while total FVI may exert a weaker overall effect.

Keywords

Fruit; vegetables; type 2 diabetes mellitus; epidemiology; meta-analysis; review

INTRODUCTION

Diabetes is a major chronic disease which is expected to affect in excess of 439 million adults worldwide by 2030,¹ with serious consequences for health and longevity. The primary prevention of diabetes is thus clearly an important public health priority. Dietary modification within the setting of lifestyle intervention trials can delay or prevent the development of type 2 diabetes (T2D).² Although the individual contribution of different foods remains unknown, fruit and vegetable intake (FVI) may explain some of this beneficial effect.

Several plausible mechanisms have been suggested to explain an apparent beneficial effect of FVI on T2D.³⁻⁶ However, findings from prospective studies on the association of FVI with T2D have been inconsistent.⁶⁻¹¹ A recent meta-analysis reported no significant association between FVI, or fruits and vegetables separately, with T2D,¹² confirming findings from an earlier meta-analysis.⁶ Nevertheless, an inverse association between green leafy vegetable (GLV) intake and T2D was found.¹² It is plausible that homogeneity in intake, in addition to measurement error, may have obscured a small but biologically important association of FVI with T2D.^{13, 14} We had the opportunity to further investigate the association between FVI and T2D in EPIC-InterAct,¹⁵ a prospective case-cohort study which includes different European populations with large variation in FVI.¹⁶

Our study therefore had two objectives: first to investigate the association between total FVI and intake of fruit, vegetables, and their sub-types and the risk of T2D in the EPIC-InterAct study; and second, to include these results in an updated meta-analysis of published studies.

SUBJECTS AND METHODS

EPIC-InterAct study

EPIC-InterAct is a large prospective case-cohort study nested within the EPIC study,¹⁷ as described previously.¹⁵ In brief, the recruitment frame (n=340 234) was sampled from 8 of 10 EPIC countries (n=455 680), excluding those without stored blood (n=109 625) or reported diabetes status (n=5 821). Among n=340,234 (with 3.99 million years of follow-up), a subcohort of 16 835 individuals was randomly selected from those with available stored blood and buffy coat, stratified by centre. After exclusion of 681 individuals with prevalent diabetes or without information on diabetes status, 16 154 subcohort individuals were included. Because of random selection, this subcohort also included a random set of 778 individuals who had developed incident T2D during follow-up. Ascertainment of incident T2D involved a review of the existing EPIC datasets at each centre using multiple sources of evidence.¹⁵ Follow-up was censored at the date of diagnosis, the 31st of December 2007, or the date of death, whichever came first. In total, 12 403 incident cases of T2D were verified (including 778 cases from the subcohort). From a total of 27 779 participants, we excluded those with incomplete dietary information or with a ratio of energy intake versus energy expenditure in the top or bottom 1% of the original EPIC study sample (n=736), or with missing information on potential confounding variables (n=955). Participants with prevalent myocardial infarction or stroke were also excluded (n=1 149), leaving 10,821 incident T2D cases and a subcohort of 14 800 individuals (including 682 incident T2D cases) for this analysis.

All participants gave written informed consent, and the study was approved by the local ethics committee in the participating countries and the Internal Review Board of the International Agency for Research on Cancer.

Dietary and non-dietary data in EPIC-InterAct

Habitual diet was estimated at baseline using country-specific questionnaires, developed and validated in the source populations.^{18, 19} The major groups and sub-groups of fruits and vegetables are shown in Table 1. To improve the comparability of dietary data across European countries a common standardised food database was developed.²⁰ In addition, a standardised 24-hour recall was collected in a stratified subsample of ~8% (n=36 900) of EPIC study participants,²¹ of whom 2 152 participants were in the EPIC-InterAct eligible population.

Standardised health and lifestyle questionnaires at baseline collected information on lifestyle exposures including history of cigarette smoking (never, former, current), occupational and leisure-time physical activity (inactive, moderately inactive, moderately active, active),²² highest achieved education level (none, primary, technical/professional, secondary, university), and history of previous illness. Height, weight and waist circumference were measured by trained staff using standardised protocols, except in Oxford (UK) and France where self-reported measurements were obtained, and Umea (Sweden) where waist circumference was not recorded.

Statistical Analysis in EPIC-InterAct

Baseline characteristics were summarised by quartiles of total FVI among subcohort participants, using means with SDs, medians with interquartile ranges (IQR), or frequencies. To account for the case-cohort design of EPIC-InterAct, multivariable Prentice-weighted Cox regression models²³ were used to estimate the association between FV intake and T2D. Total FVI and intake of fruit, vegetables, and fruit and vegetable sub-types were analysed comparing quartiles (with the lowest quartile as the reference category) based on intake data from the subcohort participants. Intake was also analysed continuously. To check the proportional hazards assumption of the models, interactions between fruit, vegetables and FVI combined, with current age (i.e. the underlying timescale) were tested. The proportional hazards assumption was not violated for fruit, vegetables or FVI combined (all p-values < 0.26). Hazard ratios (HRs) for the association of FVI with diabetes were investigated using the following modelling strategy. Age was used as the underlying timescale and all models were adjusted for study centre. Model A was adjusted for sex. Model B was additionally adjusted for education level, BMI, physical activity level, smoking status, total energy intake, and alcohol intake. For the analysis of fruit intake we also adjusted for vegetable intake and vice versa. When analysing specific fruit and vegetable sub-types, other sub-types were included as covariates. Hazard ratios (HR) and 95% confidence intervals (95% CI) for associations with diabetes were estimated within each country and displayed in forest plots. Overall combined HRs (95% CI) across countries were calculated using random effects meta-analyses. Between country heterogeneity was assessed using the I^2 statistic.

In sensitivity analyses we also included other potentially confounding variables: waist circumference, and intake of cereal fibre, red and processed meat, coffee, and sugar sweetened beverages. We conducted additional analyses excluding diabetes cases diagnosed within the first two years of recruitment. Effect modification (on the multiplicative scale) was tested by using the interaction term of quartiles of exposure combined with sex, BMI (<25 kg/m² vs. ≥25 kg/m²), and smoking status (never, former, current). The estimated interaction parameter within each country was combined across countries using the same

random effects meta-analysis method used in the main analysis to obtain a p-value for interaction.

A regression calibration model, adapted for a meta-analysis framework, was applied as described by the Fibrinogen Studies Collaboration.²⁴ Single 24-h dietary recalls (24-hDR) were used as the reference method for calibrating dietary questionnaires.^{21, 25} The 24-hDR data were regressed on the dietary questionnaire data for total FVI, fruit intake and vegetable intake separately, and for total energy and alcohol intake. Analyses were adjusted for Model B covariates. Country specific regression dilution ratios (RDRs) were then estimated and used to correct the log HRs estimated from the Prentice-weighted Cox regression models on a continuous scale. Confidence intervals around the corrected log HRs were estimated using the method described by Wood et al.²⁴

All statistical analyses were performed using Stata/SE 11.1 (Stata-Corp, College Station, Texas, USA). All p-values were based on two-sided tests, and statistical significance was set at $p < 0.05$.

Systematic review

Cohort studies published before April 30th 2011 that reported on the association between FVI and T2D were sought by MEDLINE and EMBASE searches, as well as scanning of relevant reference lists and review articles. To ensure a broad search strategy, the words: fruit, vegetable, diabetes, glucose, metabolic syndrome, and hyperglycaemia were searched for using medical subject headings and text word, title word, abstract, and subject headings. No limits on publication date or language were applied. Studies were eligible for inclusion if they had a prospective study design, reported relative risks (RR) or HRs and their corresponding 95% CIs, provided the frequency of FVI using validated dietary assessment questionnaires, and reported on incident diabetes. Two authors (AJC and ZY) reviewed all identified titles ($n=3\ 335$), and subsequently abstracts and full articles (Figure 1). If multiple published studies from the same study cohort were identified, the most recent study or the study with the most detailed information for both FVI and for diabetes risk was included. Any disagreements were resolved by discussion. For each contributing study, information was extracted by two authors (AJC and ZY). Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were followed.

Meta analysis: Standardisation of FVI

As studies reported FVI using different measurement units (e.g. g/day, servings/day), intake was standardised into portions per day (portions/day) using a portion size of 80 grams.²⁶ FVI was subsequently classified into three intake categories: high (H), medium (M), and a reference category of low (R). Assigning categories ensured that the extracted data were used appropriately since all but one study⁸ assessed FVI using FFQs, which are not suitable for determining absolute quantity of intake,²⁷ but rather classifying individuals in terms of their relative intake.²⁸ We assigned common lower and upper category values for FVI based on the pooled average of the median intake of the respective values specified by each individual study. When the median consumption value per category was not given, the midpoint of the upper and lower boundaries in each category was assigned. If the upper boundary of the highest category was not given, the same scale of FVI as the preceding category was assigned. Similarly, if the lower boundary of the reference category was not provided, the same scale of FVI as the subsequent category was assigned. If the midpoint of the reference category was below zero then a value of zero was assigned. A medium category of intake was assumed to be mid-way between the upper and reference categories. A similar protocol for fruit intake and for vegetable intake was used. GLV and root vegetable intake categories were based on weekly, not daily, consumption values because of

a low quantity of daily intake. H, M, and R categories were assigned for total FVI (H=8, M=5, and R=2 portions/day), fruit (H=4, M=2.5, and R=1 portions/day), vegetables (H=5, M=3, and R=1 portions/day), GLV (H=9, M=5, and R=1 portions/week), and root vegetables (H=4, M=2, and R=0 portions/week).

Meta analysis: statistical approach

We assumed that RRs and HRs included in published studies were equivalent in order to combine results across studies. We used RR estimates from multivariable models with the most complete adjustment for potential confounders, except in one study⁷ where we used a lesser adjusted model that we considered most similar to the other included studies in terms of covariate adjustment. We assumed a log-linear association between intake and T2D risk. Linear interpolation was used to compare the highest and middle intake categories with the reference intake category, thereby ensuring comparison across studies was with the same reference, middle, and upper categories of intake. Pooled RRs and 95% CIs of diabetes comparing the highest and middle categories of intake with the reference category were obtained using random effects meta-analyses. Heterogeneity was assessed using the I^2 statistic. To explore heterogeneity, we conducted subgroup analyses by sex (men & women vs. women only), duration of follow-up (<10 years vs. ≥10 years), location (USA vs. Europe vs. China), and dietary assessment method (FFQ vs. others). Publication bias was assessed visually and by using Begg's test to test funnel plot asymmetry.

RESULTS

EPIC-InterAct Study

The median (IQR) duration of follow-up was 11.0 (7.4-12.7) years. Among subcohort participants there was a greater than twofold variation between countries in estimated FVI (Table 2), which is similar to the marked variation in FVI previously described when estimated by the 24-hour recall.¹⁶ The highest (median [IQR]) estimated daily FVI was in Spain (531 [358-730] g/d) and the lowest in Germany (237 [180-320] g/d). Women reported greater FVI (399 [262-562] g/d) than did men (322 [197-501] g/d). Baseline characteristics of subcohort participants are shown by sex-specific quartiles of total FVI in Table 3. Men and women in the highest quartile of FVI had a higher BMI and waist circumference, were less likely to smoke, and were more likely to have lower education. The percentage of participants reporting being physically active increased across increasing quartiles of FVI for men but not for women. Energy intake and percentage energy from protein were higher across increasing quartiles of FVI, whereas percentage of energy from total fat was lower.

Total FVI was inversely associated with T2D comparing the highest with lowest quartile of intake (Model A, HR: 0.84; 95% CI: 0.73, 0.96; p for trend =0.05) (Table 4). After adjustment for potential confounders, including BMI, the inverse association was attenuated and no longer significant (Model B, HR: 0.90; 95% CI: 0.80, 1.01; p for trend =0.42) (Figure 2). We also found no evidence that a 100 g/d increment in FVI was associated with diabetes (Model B HR: 0.99; 95% CI: 0.96, 1.01), and the result from the calibration analysis was similar. The association between FVI and diabetes did not differ by sex, BMI or smoking status (p for interaction= 0.41, 0.72, and 0.43 respectively). Inclusion of waist circumference, and intake of cereal fibre, red and processed meat, coffee, and sugar sweetened beverages did not change our results, neither did excluding participants with a diagnosis date of diabetes within the first two years of follow-up (data not shown).

The association of fruit intake with T2D was moderate in magnitude but non-significant after adjustment for potential confounders (HR: 0.89; 95% CI: 0.76, 1.04; p for trend =0.30) (Table 4). Similarly, a 100 g/d increment in fruit intake was not associated with diabetes in

uncalibrated (Model B HR: 1.00; 95% CI: 0.97, 1.02) or calibrated analyses. Comparing the highest with lowest quartile of fruit intake, the percentage of total variability due to between country heterogeneity was $I^2=38\%$, with no country except the UK showing a significant inverse association with diabetes. Citrus- and non-citrus fruits were not associated with diabetes incidence. Total vegetable intake was not associated with diabetes (Table 4), and this was consistent across countries ($I^2=0.0\%$). Among vegetable sub-types however, being in the highest versus lowest quartile of root vegetable intake was associated with a 13% reduction in diabetes incidence after adjustment for potential confounders (HR: 0.87; 95% CI: 0.77, 0.99; p for trend =0.001), with little heterogeneity between countries ($I^2=12\%$). GLV intake was inversely associated with diabetes (p for trend=0.03), although the results comparing the highest with lowest quartile of intake were not statistically significant (Table 4).

Meta-analysis

Seven eligible prospective studies met the inclusion criteria, including EPIC-InterAct (Figure 1). Characteristics of the included studies are shown in Table 5. Of the seven studies (including EPIC-InterAct), four were based in the USA,^{7, 8, 10, 11} two in Europe,²⁹ and one in China.³⁰ Four studies included women only^{7, 10, 11, 30} and three included both men and women.^{8, 29} FVI was assessed using an FFQ,^{7, 10, 11, 30} 24-hour recall⁸ or diet history interview.²⁹ For total FVI, there were five contributing studies including EPIC-InterAct,^{7, 8, 10, 11} yielding 179 956 participants and 19 123 T2D cases. Six studies examined the association between diabetes and intake of fruit and vegetables separately,^{7, 10, 11, 29, 30} five with GLV intake,^{7, 10, 29, 30} and three with root vegetable intake.^{10, 30} The age of participants ranged from 25 to 79 years. Study length ranged from 4.6 to 23 years. Participant exclusion criteria differed by study, with most,^{7, 11, 29, 30} but not all,^{8, 10} excluding participants with extreme values for total energy intake. All studies, except two,^{11, 29} excluded participants with prevalent T2D at baseline. Assessment of T2D was by self-report with confirmation in all but one study which used data from a nationwide drug reimbursement register.²⁹ All studies used multivariable analyses to adjust for potential confounding factors, including age, sex (in studies including men and women) and BMI. Two studies did not additionally adjust for physical activity level and energy intake.^{8, 29}

Figure 3 shows the RR of diabetes for the individual studies reporting data on the association between FVI combined and T2D. Compared with individuals in the lowest category of FVI, the pooled RR of T2D was 0.97 (95% CI: 0.93, 1.00; $p=0.09$) for those in the middle category, and 0.93 (95% CI: 0.87, 1.00; $p=0.05$) for those in the highest category of intake. There was little heterogeneity across studies ($I^2=10\%$). Visual inspection of the funnel plots of precision against the log of the RR of T2D did not suggest publication bias for the intake of FV combined, fruit, vegetables, GLV, or root vegetables, which was confirmed formally with Begg's tests for funnel plot asymmetry: FV combined ($p=1.00$), fruit ($p=0.85$), vegetables ($p=0.19$), GLV ($p=0.14$), or root vegetables ($p=0.12$).

Among FVI sub-types, only GLV intake was consistently inversely associated with diabetes (RR comparing highest with lowest intake category: 0.84, 95% CI: 0.74, 0.94; $p=0.004$), and this association was irrespective of sex, duration of follow-up, location, or dietary assessment method (Table 6). Although a reduced RR of T2D was evident for individuals in the middle versus lowest category of root vegetable intake, this association was not consistent comparing the highest with lowest category of intake (Table 7).

Because I^2 values indicated heterogeneity across studies for the intake of fruit, vegetables, GLV and root vegetables (Table 7), we conducted subgroup analyses by sex (men and women vs. women only), duration of follow-up (<10 years vs. 10 years), location (USA vs. Europe vs. China), and dietary assessment method (FFQ vs. others) (Table 6). Examination

of the RRs of diabetes by study characteristics showed that much of the variability across studies was due to duration of follow-up and the dietary assessment method used to estimate FVI. Studies with ≥ 10 years of follow-up generally found stronger inverse associations between FVI and risk of T2D than studies with <10 years of follow-up. Studies using an FFQ to estimate FV intake appeared to consistently report weaker associations between intake and risk of T2D when compared with studies that used other dietary assessment methods (e.g. 24-hour recall).

DISCUSSION

We provide a comprehensive assessment of the association between FVI and diabetes in our meta-analysis of $>179\,000$ individuals (with $>19\,000$ incident T2D cases), including new data from the large EPIC-InterAct study in Europe. For total FVI there was a weak inverse association with incident diabetes, which was non-significant in EPIC-InterAct, and of modest magnitude in the meta-analysis, with a 7% lower RR of diabetes in those with the highest versus lowest FVI. However, the association between diabetes and FVI was most pronounced for specific sub-types of vegetables, including root vegetables and green leafy vegetables (GLV), suggesting that persons at risk of diabetes may benefit from consuming higher quantities of these specific vegetable sub-types.

The suggestion from our meta-analysis of a modest inverse association of FVI with T2D is consistent with findings from biomarker,^{9, 31} and dietary pattern studies,³²⁻³⁶ but not with the findings from some prospective cohort studies. When exploring potential sources of heterogeneity across studies we observed that associations between FVI and T2D tended to be weaker when intake was assessed using a FFQ compared with other dietary assessment methods (i.e. 24-hour recall). This finding is consistent with previous studies indicating that detection of diet-disease associations may be sensitive to the dietary assessment method used to estimate intake,^{9, 37} suggesting that the inconsistency between studies examining the association between FV intake and risk of diabetes could be due to the extent of measurement error associated with the FFQ.³⁷ There is some suggestion that fruit may be differentially associated with diabetes compared with vegetables.^{7, 9, 30} We observed no significant association of diabetes with fruit intake, but specific sub-types of vegetables were inversely associated with diabetes. In our meta-analysis, we observed that those in the highest compared with lowest category of GLV intake (9 vs. 1 portion per week) had a 16% reduced risk of diabetes, which is consistent with the findings reported in a previous meta-analysis.¹² The inverse association of root vegetable intake with diabetes that we observed in EPIC-InterAct (HR: 0.87, 95% CI: 0.77-0.99) was not replicated in our meta-analysis when combined with the findings from two other studies.^{10, 30}

Although several biological mechanisms have been proposed to explain an inverse association of FVI with diabetes, no clear mechanism(s) exists as yet. Antioxidants in fruits and vegetables have been hypothesised to protect against diabetes. Yet several supplementation trials, including supplementation with β -carotene and vitamin C, have reported null associations on adverse metabolic traits, including diabetes.³⁸⁻⁴⁰ It is plausible that an inverse association with diabetes only occurs in the presence of a complex mixture of antioxidants, as found in whole fruits and vegetables.⁴¹ An inverse association of FVI with diabetes may also be mediated through body weight,^{42, 43} in which case adjustment for BMI would constitute over-adjustment. However, exclusion of BMI from the EPIC-InterAct analysis (Model B) made little difference to the observed estimates (data not shown). Regarding specific mechanisms related to GLV, it has been shown that supplementation with magnesium, of which GLV are an important source, may improve glucose metabolism⁴⁴ and reduce the risk of diabetes.⁴⁵ An inverse association of GLV with diabetes risk may also work through the nitric oxide-1-arginine pathway.^{46, 47} Further

mechanistic studies are needed to help explain the apparent inverse association of FVI, and of GLV and root vegetables in particular, with T2D.

This study has some important strengths. We included data from EPIC-InterAct which estimated FVI using dietary questionnaires and standardised 24-hour recalls across eight different European countries. Notably, EPIC-InterAct includes countries with a high degree of heterogeneity in FVI and with different confounding structures. By combining results from EPIC-InterAct with previously published studies we more than doubled the total number of T2D cases included in previous meta-analyses.^{6, 12} We were able to explore possible sources of heterogeneity across studies, and were able to demonstrate that one likely source of heterogeneity that is likely, in part, to explain the inconsistent findings between studies is the dietary assessment method used to estimate FVI. Although the reference and upper intake categories varied among individual studies, by using linear interpolation we were able to ensure the same categories of intake were used for all studies, which was not the case previously.^{6, 12} We acknowledge that the upper intake category of 8 portions/day for total FVI is higher than the public health message to consume at least 5 portions/day.⁴⁸ However, since all but one study⁸ included in our meta-analysis used FFQs to assess FVI, which substantially over-estimate intake,²⁷ the assigned upper category of 8 portions/day is not inconsistent with current WHO recommendations.⁴⁸ It is therefore possible that previous studies failed to find an association between FVI and T2D risk because of setting an upper intake category too low. Limitations of our study and others included in the meta-analysis also merit consideration. All but one study included in the meta-analysis³⁰ used a single baseline assessment of FVI to determine long-term exposure. This, in addition to the measurement error associated with self-report instruments will likely have biased our risk estimates towards the null. Also, any potential misclassification of individuals with undiagnosed T2D as non-diabetic in any of the included studies may have attenuated our overall findings. A possible weakness inherent in any systematic review or meta-analysis is the possibility that the wrong conclusion is made as a result of publication bias.⁴⁹ However, we found no evidence to suggest publication bias as a likely alternative explanation for our findings. The studies included in the meta-analysis differed in adjustment for covariates, particularly in relation to lifestyle behaviours which tend to cluster with FV intake.^{50, 51} Therefore, we are unable to exclude residual confounding by unmeasured or imperfectly measured lifestyle factors as a plausible explanation for our findings.

In conclusion, our findings from the meta-analysis of cohort studies, including new data from EPIC-InterAct with wide variation in FVI across Europe, provide evidence that specific groups of vegetables, principally GLV and root vegetables, may be beneficial in preventing diabetes, while higher total FVI is weakly inversely associated with diabetes.

Acknowledgments

We thank all EPIC participants and staff for their contribution to the study. We thank Nicola Kerrison (MRC Epidemiology Unit, Cambridge) for managing the data for the EPIC-InterAct Project.

Funding: Funding for the InterAct project was provided by the EU FP6 programme (grant number LSHM_CT_2006_037197). In addition, InterAct investigators acknowledge funding from the following agencies: LA: We thank to the participants of the Spanish EPIC cohort for their contribution to the study as well as to the team of trained nurses who participated in the recruitment; JWJB: Verification of diabetes cases in EPIC-NL was additionally funded by NL Agency grant IGE05012 and an Incentive Grant from the Board of the UMC Utrecht; PWF: Swedish Research Council, Novo Nordisk, Swedish Diabetes Association, Swedish Heart-Lung Foundation; RK: German Cancer Aid; TJK: Cancer Research UK; CN: Health Research Fund (FIS) of the Spanish Ministry of Health; Murcia Regional Government (N° 6236); PN: Swedish Research Council; KO: Danish Cancer Society; SP: Compagnia di San Paolo; JRQ: Asturias Regional Government; OR: The Västerboten County Council; IS: Verification of diabetes cases was additionally funded by NL Agency grant IGE05012 and an Incentive Grant from the Board of the UMC Utrecht; AMWS: Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands

Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands; BT: German Cancer Aid; AT: Danish Cancer Society; RT: AIRE-ONLUS Ragusa, AVIS-Ragusa, Sicilian Regional Government; ER: ER was supported in this work by the Imperial College Biomedical Research Centre.

References

1. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract.* 2010; 87:4–14. [PubMed: 19896746]
2. Gillies CL, Abrams KR, Lambert PC, et al. Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis. *Bmj.* 2007; 334:299. [PubMed: 17237299]
3. Barclay AW, Petocz P, McMillan-Price J, et al. Glycemic index, glycemic load, and chronic disease risk--a meta-analysis of observational studies. *Am J Clin Nutr.* 2008; 87:627–37. [PubMed: 18326601]
4. Wang J, Luben R, Khaw KT, Bingham S, Wareham NJ, Forouhi NG. Dietary energy density predicts the risk of incident type 2 diabetes: the European Prospective Investigation of Cancer (EPIC)-Norfolk Study. *Diabetes Care.* 2008; 31:2120–5. [PubMed: 18689693]
5. Larsson SC, Wolk A. Magnesium intake and risk of type 2 diabetes: a meta-analysis. *J Intern Med.* 2007; 262:208–14. [PubMed: 17645588]
6. Hamer M, Chida Y. Intake of fruit, vegetables, and antioxidants and risk of type 2 diabetes: systematic review and meta-analysis. *J Hypertens.* 2007; 25:2361–9. [PubMed: 17984654]
7. Bazzano LA, Li TY, Joshipura KJ, Hu FB. Intake of fruit, vegetables, and fruit juices and risk of diabetes in women. *Diabetes Care.* 2008; 31:1311–7. [PubMed: 18390796]
8. Ford ES, Mokdad AH. Fruit and vegetable consumption and diabetes mellitus incidence among U.S. adults. *Prev Med.* 2001; 32:33–9. [PubMed: 11162324]
9. Harding AH, Wareham NJ, Bingham SA, et al. Plasma vitamin C level, fruit and vegetable consumption, and the risk of new-onset type 2 diabetes mellitus: the European prospective investigation of cancer--Norfolk prospective study. *Arch Intern Med.* 2008; 168:1493–9. [PubMed: 18663161]
10. Liu S, Serdula M, Janket SJ, et al. A prospective study of fruit and vegetable intake and the risk of type 2 diabetes in women. *Diabetes Care.* 2004; 27:2993–6. [PubMed: 15562224]
11. 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. [PubMed: 10731498]
12. Carter P, Gray LJ, Troughton J, Khunti K, Davies MJ. Fruit and vegetable intake and incidence of type 2 diabetes mellitus: systematic review and meta-analysis. *Bmj.* 2010; 341:c4229. [PubMed: 20724400]
13. Day N, McKeown N, Wong M, Welch A, Bingham S. Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. *Int J Epidemiol.* 2001; 30:309–17. [PubMed: 11369735]
14. Kipnis V, Midthune D, Freedman LS, et al. Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol.* 2001; 153:394–403. [PubMed: 11207158]
15. The InterAct Consortium. 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. [PubMed: 21717116]
16. Agudo A, Slimani N, Ocke MC, et al. Consumption of vegetables, fruit and other plant foods in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts from 10 European countries. *Public Health Nutr.* 2002; 5:1179–96. [PubMed: 12639226]
17. Bingham S, Riboli E. Diet and cancer--the European Prospective Investigation into Cancer and Nutrition. *Nat Rev Cancer.* 2004; 4:206–15. [PubMed: 14993902]
18. Margetts BM, Pietinen P. European Prospective Investigation into Cancer and Nutrition: validity studies on dietary assessment methods. *Int J Epidemiol.* 1997; 26(Suppl 1):S1–5. [PubMed: 9126528]

19. Riboli E, Hunt KJ, Slimani N, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr.* 2002; 5:1113–24. [PubMed: 12639222]
20. Slimani N, Deharveng G, Unwin I, 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. [PubMed: 17375121]
21. Slimani N, Ferrari P, Ocke M, et al. Standardization of the 24-hour diet recall calibration method used in the European prospective investigation into cancer and nutrition (EPIC): general concepts and preliminary results. *Eur J Clin Nutr.* 2000; 54:900–17. [PubMed: 11114689]
22. Wareham NJ, Jakes RW, Rennie KL, et al. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr.* 2003; 6:407–13. [PubMed: 12795830]
23. Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. *J Clin Epidemiol.* 1999; 52:1165–72. [PubMed: 10580779]
24. Fibrinogen Studies Collaboration. Correcting for multivariate measurement error by regression calibration in meta-analyses of epidemiological studies. *Stat Med.* 2009; 28:1067–92. [PubMed: 19222086]
25. Slimani N, Deharveng G, Charrondiere RU, 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. [PubMed: 10094230]
26. FAO/WHO. Fruit and Vegetables for Health: Report of a Joint FAO/WHO Workshop; Kobe, Japan. 2004; 2004.
27. Michels KB, Welch AA, Luben R, Bingham SA, Day NE. Measurement of fruit and vegetable consumption with diet questionnaires and implications for analyses and interpretation. *Am J Epidemiol.* 2005; 161:987–94. [PubMed: 15870164]
28. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol.* 1985; 122:51–65. [PubMed: 4014201]
29. 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. [PubMed: 15674312]
30. Villegas R, Shu XO, Gao YT, et al. Vegetable but not fruit consumption reduces the risk of type 2 diabetes in Chinese women. *J Nutr.* 2008; 138:574–80. [PubMed: 18287369]
31. Hozawa A, Jacobs DR Jr, Steffes MW, Gross MD, Steffen LM, Lee DH. Associations of serum carotenoid concentrations with the development of diabetes and with insulin concentration: interaction with smoking: the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Epidemiol.* 2006; 163:929–37. [PubMed: 16597706]
32. Esposito K, Kastorini CM, Panagiotakos DB, Giugliano D. Prevention of type 2 diabetes by dietary patterns: a systematic review of prospective studies and meta-analysis. *Metab Syndr Relat Disord.* 2010; 8:471–6. [PubMed: 20958207]
33. Brunner EJ, Mosdol A, Witte DR, et al. Dietary patterns and 15-y risks of major coronary events, diabetes, and mortality. *Am J Clin Nutr.* 2008; 87:1414–21. [PubMed: 18469266]
34. Nettleton JA, Steffen LM, Ni H, Liu K, Jacobs DR Jr. Dietary patterns and risk of incident type 2 diabetes in the Multi-Ethnic Study of Atherosclerosis (MESA). *Diabetes Care.* 2008; 31:1777–82. Epub 2008 Jun 10. [PubMed: 18544792]
35. Odegaard AO, Koh WP, Butler LM, et al. Dietary patterns and incident type 2 diabetes in Chinese men and women: the Singapore Chinese Health Study. *Diabetes.* 34:880–5. Epub 2011 Feb 17.
36. Yu R, Woo J, Chan R, et al. Relationship between dietary intake and the development of type 2 diabetes in a Chinese population: the Hong Kong Dietary Survey. *Public Health Nutr.* 2011; 5:1–9.
37. Bingham SA, Luben R, Welch A, Wareham N, Khaw KT, Day N. Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet.* 2003; 362:212–4. [PubMed: 12885485]
38. The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. The ATBC Cancer Prevention Study Group. *Ann Epidemiol.* 1994; 4:1–10. [PubMed: 8205268]

39. Liu S, Ajani U, Chae C, Hennekens C, Buring JE, Manson JE. Long-term beta-carotene supplementation and risk of type 2 diabetes mellitus: a randomized controlled trial. *Jama*. 1999; 282:1073–5. [PubMed: 10493207]
40. Song Y, Cook NR, Albert CM, Van Denburgh M, Manson JE. Effects of vitamins C and E and beta-carotene on the risk of type 2 diabetes in women at high risk of cardiovascular disease: a randomized controlled trial. *Am J Clin Nutr*. 2009; 90:429–37. [PubMed: 19491386]
41. Liu RH. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. *Am J Clin Nutr*. 2003; 78:517S–520S. [PubMed: 12936943]
42. Buijsse B, Feskens EJ, Schulze MB, et al. Fruit and vegetable intakes and subsequent changes in body weight in European populations: results from the project on Diet, Obesity, and Genes (DiOGenes). *Am J Clin Nutr*. 2009; 90:202–9. [PubMed: 19458016]
43. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in diet and lifestyle and long-term weight gain in women and men. *New England Journal of Medicine*. 2011; 364:2392–404. [PubMed: 21696306]
44. Song Y, He K, Levitan EB, Manson JE, Liu S. Effects of oral magnesium supplementation on glycaemic control in Type 2 diabetes: a meta-analysis of randomized double-blind controlled trials. *Diabet Med*. 2006; 23:1050–6. [PubMed: 16978367]
45. Schulze MB, Schulz M, Heidemann C, Schienkiewitz A, Hoffmann K, Boeing H. Fiber and magnesium intake and incidence of type 2 diabetes: a prospective study and meta-analysis. *Arch Intern Med*. 2007; 167:956–65. [PubMed: 17502538]
46. Gilchrist M, Benjamin N. Vegetables and diabetes. Is nitrate the answer? *BMJ*. 2010; 341:c5306. 10.1136/bmj.c5306. [PubMed: 20880918]
47. Scherrer U, Sartori C. Defective nitric oxide synthesis: a link between metabolic insulin resistance, sympathetic overactivity and cardiovascular morbidity. *Eur J Endocrinol*. 2000; 142:315–23. [PubMed: 10754469]
48. WHO/FAO. Diet, nutrition and the prevention of chronic diseases: report of a joint FAO/WHO Expert Consultation. World Health Organization; Geneva: 2003. Technical Report Series, NO. 916
49. Egger M, Smith GD. Misleading meta-analysis. *BMJ*. 1995; 310:752–4. [PubMed: 7711568]
50. Agudo A, Pera G. Vegetable and fruit consumption associated with anthropometric, dietary and lifestyle factors in Spain. EPIC Group of Spain. European Prospective Investigation into Cancer. *Public Health Nutr*. 1999; 2:263–71. [PubMed: 10512560]
51. Serdula MK, Byers T, Mokdad AH, Simoes E, Mendlein JM, Coates RJ. The association between fruit and vegetable intake and chronic disease risk factors. *Epidemiology*. 1996; 7:161–5. [PubMed: 8834556]

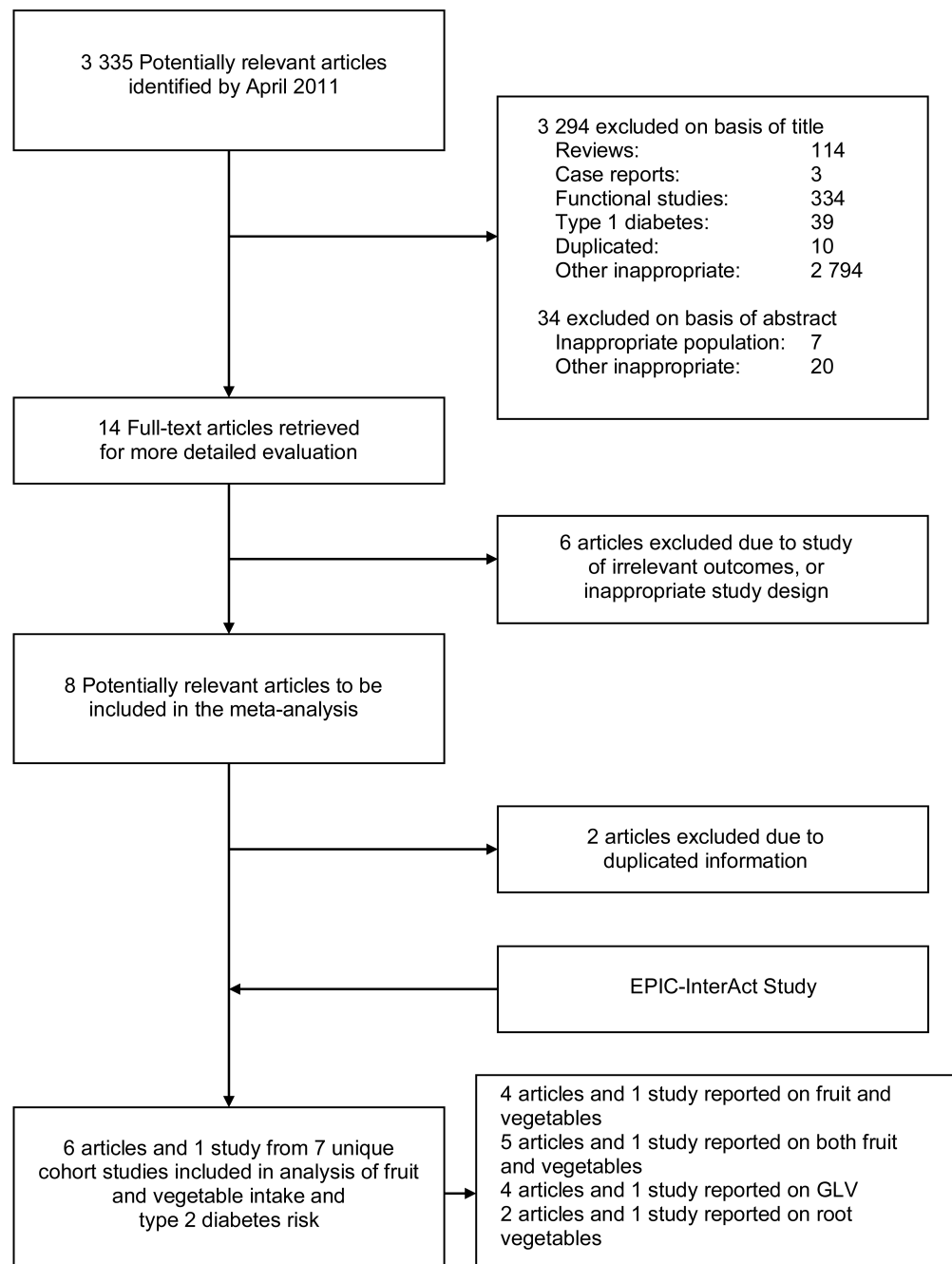


Figure 1. Flow diagram of the selection of prospective studies of fruit and vegetable intake and risk of type 2 diabetes through systematic review of the literature

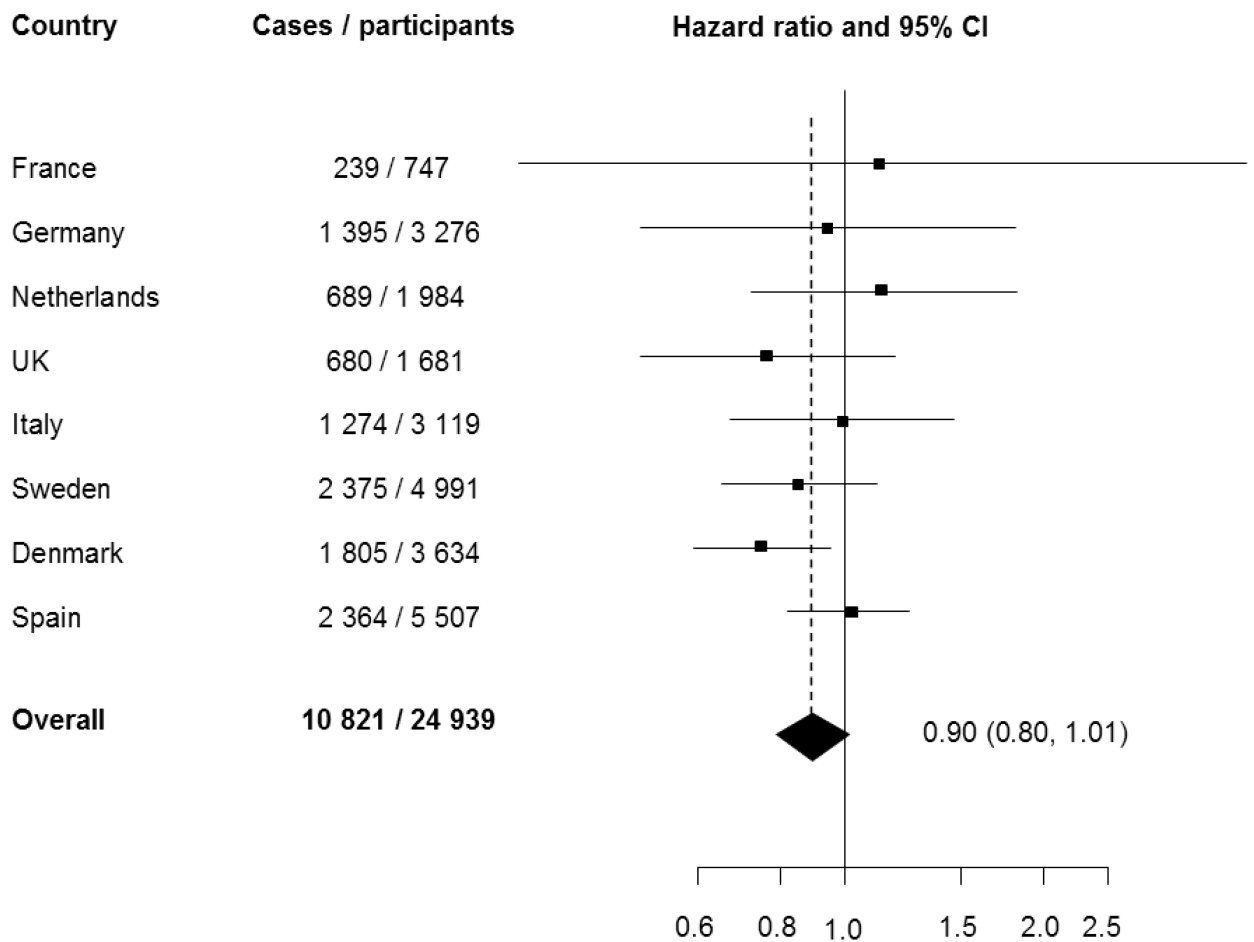


Figure 2.

Hazard ratio of type 2 diabetes comparing the highest with the lowest quartile of estimated total fruit and vegetable intake across countries: EPIC-InterAct study *

* Model with age as the underlying timescale and adjusted for centre, sex, education level, BMI, physical activity level, smoking status, total energy intake, and alcohol intake. Square sizes are proportional to weight of each country

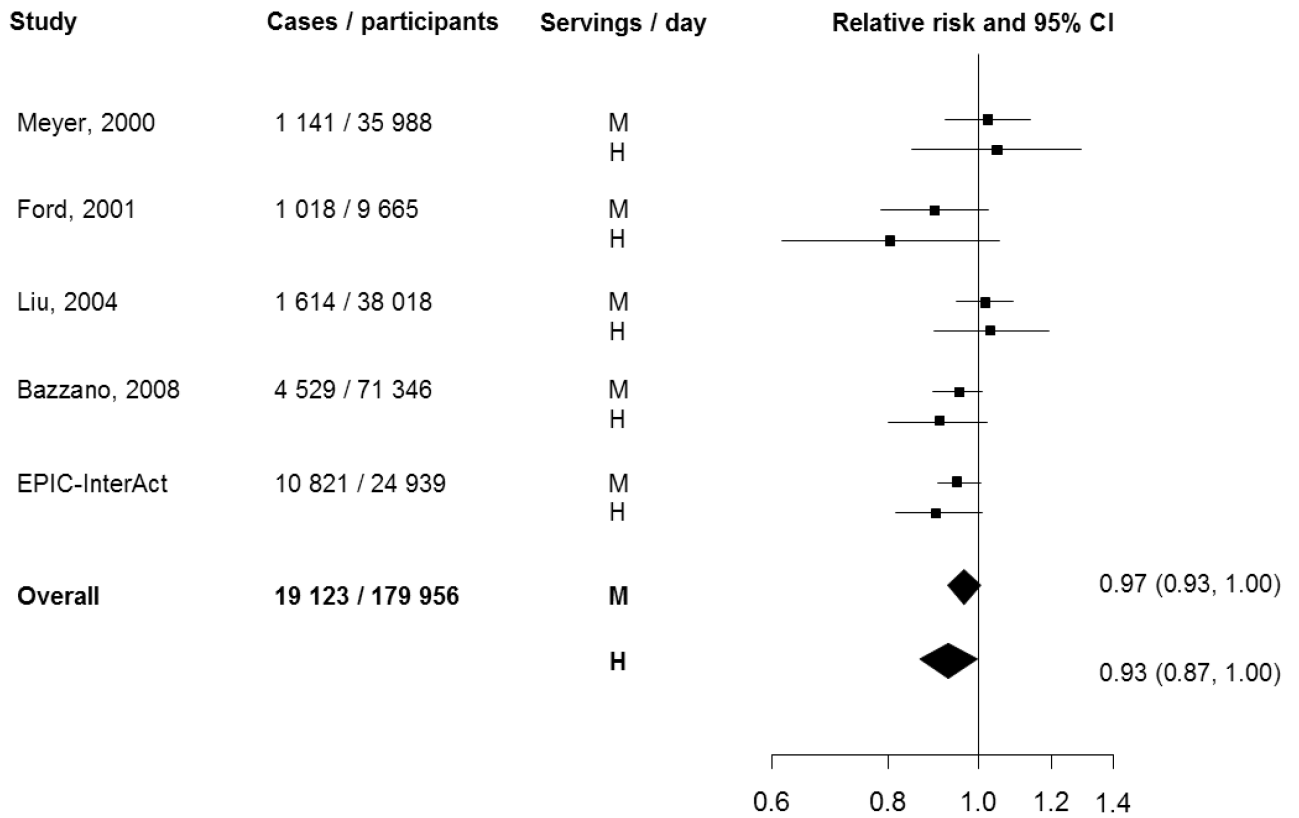


Figure 3. Relative risk of type 2 diabetes for the middle and highest estimated intake categories of fruit and vegetables versus the reference intake category: meta-analysis results
M=medium category relative to the reference category, H=high category relative to the reference category.

Table 1

Classification by major groups and sub-groups of fruits and vegetables in the EPIC-InterAct Study *

Major group	Sub-groups	Selected foods
Fruits	Citrus fruits	Grapefruit, orange, tangerine, lemon
	Fruits (non-citrus)	Apple, pear, grape, apricot, cherry, peach, plum, nectarine, prune, melon, pineapple, strawberry, raspberry, blueberry, banana, kiwi
Vegetables	Green leafy vegetables	Spinach, chard, endive, lettuce, borage, watercress, beet leaves
	Fruiting vegetables	Tomato, pepper, avocado, courgette, artichoke, aubergine, pumpkin, squash
	Root vegetables	Carrot, radish, salsify, beetroot, turnip, celeriac, swede
	Cabbages	Broccoli, cabbage, Brussels sprouts, cauliflower, kale
	Onion, garlic	Onion, shallot, spring onion, garlic
	Stalk vegetables, sprouts	Leek, celery, fennel, asparagus, bean sprouts, bamboo shoots
	Other vegetables	Peas, corn, broad bean, mushroom, champignon, mixed salad, mixed vegetables

* Potatoes, other tubers and legumes were not included in the category of total vegetables since they differ from vegetables regarding energy and carbohydrate content. Fruit and vegetable juices were also not included in this analysis as they differ to their source of origin in terms of added sugars and food matrix.

Table 2

Median (interquartile range) estimated total fruit and vegetable intake, and fruit intake and vegetable intake separately, by sex and country for the EPIC-InterAct subcohort *

	Subcohort (n)	Median (IQR) fruit & vegetable intake, g/d	Median (IQR) fruit intake, g/d	Median (IQR) vegetable intake, g/d
All				
France	516	521 (391-672)	240 (149-331)	264 (194-367)
Germany	1 932	237 (180-320)	105 (80-184)	115 (88-154)
The Netherlands	1 329	327 (229-434)	191 (120-273)	126 (98-161)
United Kingdom	1 021	447 (327-612)	201 (121-311)	235 (172-314)
Italy	1 899	476 (348-629)	302 (206-419)	156 (105-224)
Sweden	2 766	278 (175-405)	149 (87-241)	111 (63-180)
Denmark	1 956	320 (207-466)	143 (72-244)	163 (105-233)
Spain	3 381	531 (358-730)	276 (156-438)	222 (141-331)
<i>Total</i>	<i>14 800</i>	<i>369 (236-545)</i>	<i>193 (104-316)</i>	<i>155 (101-238)</i>
Men				
France	-	-	-	-
Germany	779	212 (162-295)	98 (67-152)	108 (81-149)
The Netherlands	216	232 (177-310)	121 (65-185)	116 (88-143)
United Kingdom	389	396 (288-526)	164 (101-267)	205 (154-282)
Italy	629	442 (329-606)	291 (195-411)	140 (99-200)
Sweden	1 168	225 (139-355)	123 (64-203)	95 (49-160)
Denmark	1 031	282 (186-415)	116 (52-198)	153 (99-222)
Spain	1 290	541 (349-758)	269 (144-438)	237 (146-355)
<i>Total</i>	<i>5 502</i>	<i>322 (197-501)</i>	<i>154 (80-284)</i>	<i>143 (90-227)</i>
Women				
France	516	521 (391-672)	240 (149-331)	264 (194-367)
Germany	1 153	254 (192-342)	117 (89-197)	119 (92-157)
The Netherlands	1 113	350 (243-447)	230 (124-298)	129 (101-164)
United Kingdom	632	493 (357-656)	228 (140-337)	253 (185-335)
Italy	1 270	485 (362-641)	311 (212-430)	164 (109-233)
Sweden	1 598	312 (207-440)	168 (110-260)	126 (76-189)
Denmark	925	364 (243-507)	171 (100-266)	174 (114-244)
Spain	2 091	528 (363-715)	281 (167-438)	216 (138-313)
<i>Total</i>	<i>9 298</i>	<i>399 (262-562)</i>	<i>217 (122-332)</i>	<i>162 (107-242)</i>

* Estimated intake data from dietary questionnaire

Table 3

Distribution of selected baseline characteristics among subcohort men and women according to quartiles of estimated total fruit and vegetable intake: EPIC-InterAct Study

	Men (n=5 502)				Women (n=9 298)			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Fruit and vegetables (g/d)								
Total fruit (g/d)	161 (115-198)	299 (266-335)	442 (405-490)	709 (620-873)	179 (137-207)	302 (269-335)	450 (408-494)	690 (602-824)
Total vegetables (g/d)	65 (32-98)	147 (106-193)	249 (181-305)	433 (324-570)	79 (45-110)	158 (121-201)	256 (205-311)	425 (324-540)
	83 (53-114)	148 (105-191)	195 (136-258)	294 (199-401)	87 (60-113)	138 (106-178)	188 (139-242)	281 (199-370)
Sociodemographic variables								
Age at recruitment (y)	51.9 (9.5)	53.2 (8.9)	53.5 (8.2)	52.1 (7.9)	50.6 (9.7)	52.1 (9.4)	52.8 (8.9)	51.8 (8.8)
Education level, n (%)								
None	41 (2.2)	45 (3.2)	79 (7.2)	144 (12.3)	94 (5.1)	138 (6.0)	229 (8.8)	371 (14.7)
Primary school	596 (32.3)	469 (33.8)	387 (35.3)	419 (35.8)	579 (31.2)	690 (29.9)	826 (31.7)	928 (36.7)
Technical/professional school	499 (27.1)	317 (22.8)	225 (20.5)	203 (17.3)	582 (31.4)	620 (26.8)	616 (23.7)	360 (14.2)
Secondary school	259 (14.1)	167 (12.0)	149 (13.6)	147 (12.6)	283 (15.3)	373 (16.1)	448 (17.2)	431 (17.0)
University	449 (24.4)	391 (28.2)	258 (23.5)	258 (22.0)	318 (17.1)	490 (21.2)	483 (18.6)	439 (17.4)
Anthropometric measures								
Height (cm)	175.5 (7.0)	174.5 (7.5)	173.8 (7.4)	171.6 (7.4)	162.8 (6.7)	162.2 (6.7)	161.2 (6.8)	159.9 (6.5)
Weight (kg)	80.7 (11.7)	80.3 (11.8)	80.6 (11.7)	81.0 (11.3)	66.7 (12.0)	66.5 (11.4)	66.5 (11.3)	67.2 (11.5)
Waist circumference (cm) *	94.4 (10.2)	94.2 (9.8)	95.1 (10.0)	96.6 (9.4)	80.8 (11.6)	80.1 (10.7)	81.0 (10.9)	82.5 (11.3)
BMI (Kg/m ²)	26.2 (3.5)	26.4 (3.6)	26.7 (3.5)	27.5 (3.6)	25.2 (4.5)	25.3 (4.3)	25.6 (4.3)	26.3 (4.6)
Lifestyle								
Physically activity level, n (%)								
Inactive	369 (20.0)	247 (17.8)	185 (16.9)	193 (16.5)	450 (24.3)	506 (21.9)	669 (25.7)	832 (32.9)
Moderately active	578 (31.3)	431 (31.0)	351 (32.0)	347 (29.6)	675 (36.4)	863 (37.3)	892 (34.3)	854 (33.8)
Moderately active	484 (26.3)	359 (25.9)	282 (25.7)	306 (26.1)	431 (23.2)	500 (21.6)	551 (21.2)	467 (18.5)
Active	413 (22.4)	352 (25.3)	280 (25.5)	325 (27.8)	300 (16.2)	442 (19.1)	490 (18.8)	376 (14.9)
Smoking status, n (%)								
Never	560 (30.4)	471 (33.9)	349 (31.2)	403 (34.4)	849 (45.7)	1,219 (52.8)	1,514 (58.2)	1,620 (64.1)
Former	586 (31.8)	497 (35.8)	440 (40.1)	441 (37.7)	371 (20.0)	527 (22.8)	568 (21.8)	525 (20.8)
Current	698 (37.9)	421 (30.3)	309 (28.1)	327 (27.9)	636 (34.3)	565 (24.5)	520 (20.0)	384 (15.2)

Diet	Fruit and vegetable intake							
	Men (n=5 502)				Women (n=9 298)			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Energy (MJ/d)	9.4 (7.9-11.2)	10.1 (8.5-11.8)	10.4 (8.8-12.4)	11.2 (9.4-13.3)	7.0 (5.8-8.3)	7.6 (6.3-8.9)	8.0 (6.7-9.6)	8.6 (7.2-10.3)
Carbohydrate (% energy)	42.3 (37.5-47.5)	42.6 (37.9-47.3)	43.4 (38.6-47.6)	42.6 (37.5-47.6)	43.7 (38.9-48.5)	44.1 (39.9-48.6)	44.9 (40.7-49.1)	45.7 (41.0-50.2)
Protein (% energy)	15.4 (13.8-17.3)	16.4 (14.6-18.3)	16.8 (15.0-18.7)	17.7 (15.7-19.5)	16.0 (14.3-18.2)	16.8 (14.9-18.9)	17.3 (15.3-19.3)	17.8 (15.8-19.8)
Total fat (% energy)	35.3 (31.1-39.2)	34.3 (30.9-38.1)	33.8 (30.5-37.5)	33.9 (29.9-37.7)	36.1 (32.1-39.8)	35.5 (31.8-39.1)	34.7 (30.9-38.4)	34.4 (30.5-38.4)
Alcohol intake (% energy)	4.2 (1.4-10.1)	4.7 (1.6-9.7)	4.4 (1.4-9.1)	4.3 (1.3-9.0)	1.6 (0.3-5.3)	1.8 (0.3-4.7)	1.3 (0.1-4.2)	0.6 (0.0-2.9)
Cereal fiber (g/d)	9.6 (6.5-13.7)	10.1 (6.7-14.1)	9.3 (6.3-13.6)	8.7 (5.6-12.3)	6.9 (4.8-9.7)	7.2 (5.1-10.1)	7.0 (4.7-9.8)	6.4 (4.2-9.4)
Red and processed meat (g/d)	4.2 (1.4-10.1)	4.7 (1.6-9.7)	4.4 (1.4-9.1)	4.3 (1.3-9.0)	1.6 (0.3-5.3)	1.8 (0.3-4.7)	1.3 (0.1-4.2)	0.6 (0.0-2.9)
Coffee (g/d)	485.7 (203.4-801.1)	375.8 (115.0-625.0)	192.4 (74.1-500.0)	100.0 (30.0-222.4)	384.9 (159.9-600.8)	310.7 (120.3-580.2)	249.7 (89.3-500.0)	150.0 (50.0-375.0)
Sugar sweetened beverages (g/d)	42.7 (0.0-203.5)	26.3 (0.0-174.7)	8.2 (0.0-125.1)	0.0 (0.0-67.1)	16.0 (0.0-142.5)	6.6 (0.0-109.7)	0.0 (0.0-94.7)	0.0 (0.0-56.0)

Data are means \pm SD, medians (IQR) for dietary variables, and % for categorical variables.

Umea (Sweden) excluded from analysis as information on waist circumference was not available.

Table 4

Hazard ratios (95% CI) for incident diabetes comparing quartiles of estimated FVI, total fruit, fruit sub-types, total vegetables, and vegetable sub-types: EPIC-InterAct Study

Range (g/d)	Quartiles of intake				P for trend
	Q1	Q2	Q3	Q4	
Total fruit and vegetables					
Range (g/d)	<235.7	235.7 - <369.1	369.1 - <544.8	544.8	
Total No. of cases	3 031	2 632	2 545	2 613	
Model A *	1.00	0.89 (0.81-0.97)	0.87 (0.80-0.95)	0.84 (0.73-0.96)	0.05
Model B †	1.00	0.92 (0.83-1.03)	0.93 (0.84-1.03)	0.90 (0.80-1.01)	0.42
Total fruit ¹					
Range (g/d)	<103.7	103.7 - <193.4	193.4 - <315.9	315.9	
Total No. of cases	2 989	2 695	2 521	2 616	
Model A *	1.00	0.91 (0.85-0.98)	0.88 (0.78-0.99)	0.87 (0.80-0.96)	0.01
Model B †	1.00	0.92 (0.83-1.03)	0.94 (0.83-1.05)	0.89 (0.76-1.04)	0.30
Citrus fruit ²					
Range (g/d)	<10.1	10.1 - <35.9	35.9 - <79.4	79.4	
Total No. of cases	3 002	2 627	2 536	2 656	
Model A *	1.00	0.97 (0.90-1.05)	0.96 (0.88-1.04)	0.93 (0.84-1.03)	0.46
Model B †	1.00	0.96 (0.86-1.07)	1.00 (0.90-1.10)	1.01 (0.86-1.19)	0.86
Non-citrus fruit ^{2, 3}					
Range (g/d)	<53.0	53.0 - <120.9	120.9 - <213.5	213.5	
Total No. of cases	2 374	2 552	2 504	2 601	
Model A *	1.00	0.98 (0.88-1.08)	0.93 (0.83-1.04)	0.94 (0.86-1.03)	0.10
Model B †	1.00	1.02 (0.92-1.13)	0.97 (0.87-1.08)	0.94 (0.79-1.13)	0.46
Total vegetables ⁴					
Range (g/d)	<100.5	100.5 - <154.8	154.8 - <237.6	237.6	
Total No. of cases	3 027	2 619	2 572	2 603	
Model A *	1.00	0.94 (0.87-1.01)	0.91 (0.81-1.03)	0.88 (0.72-1.07)	0.33
Model B †	1.00	0.92 (0.84-1.01)	0.93 (0.83-1.05)	0.94 (0.84-1.05)	0.30
Green leafy vegetables ^{5, 6}					
Range (g/d)	<3.2	3.2 - <14.1	14.1 - <37.7	37.7	
Total No. of cases	1 274	2 182	2 241	2 529	
Model A *	1.00	0.74 (0.65-0.84)	0.66 (0.58-0.75)	0.79 (0.68-0.92)	0.04
Model B †	1.00	0.74 (0.65-0.84)	0.75 (0.65-0.86)	0.84 (0.65-1.07)	0.03
Fruiting vegetables					
Range (g/d)	<28.6	28.6 - <50.5	50.5 - <87.1	87.1	
Total No. of cases	3 098	2 651	2 446	2 626	
Model A *	1.00	0.93 (0.86-1.01)	0.91 (0.83-1.01)	0.98 (0.85-1.13)	0.45

Range (g/d)	Quartiles of intake				P for trend
	Q1	Q2	Q3	Q4	
Model B [†]	1.00	0.94 (0.86-1.04)	0.96 (0.86-1.06)	0.97 (0.85-1.12)	0.36
Root vegetables					
Range (g/d)	<3.9	3.9 - <11.1	11.1 - <27.3	27.3	
Total No. of cases	2 962	2 825	2 517	2 517	
Model A [*]	1.00	0.91 (0.84-0.98)	0.82 (0.75-0.89)	0.80 (0.73-0.88)	<0.001
Model B [†]	1.00	0.98 (0.88-1.08)	0.85 (0.76-0.95)	0.87 (0.77-0.99)	0.001
Cabbages ⁵					
Range (g/d)	<1.5	1.5 - <8.5	8.5 - <21.4	21.4	
Total No. of cases	2 461	2 459	2 566	2 545	
Model A [*]	1.00	0.93 (0.83-1.05)	0.91 (0.77-1.06)	0.91 (0.74-1.13)	0.23
Model B [†]	1.00	0.94 (0.74-1.19)	0.93 (0.80-1.07)	0.90 (0.75-1.09)	0.44
Onion & garlic ^{5, 7}					
Range (g/d)	<2.6	2.6 - <7.0	7.0 - <17.7	17.7	
Total No. of cases	1 883	2 288	2 721	2 900	
Model A [*]	1.00	0.92 (0.78-1.10)	0.92 (0.75-1.14)	0.97 (0.73-1.30)	0.97
Model B [†]	1.00	0.94 (0.75-1.18)	0.88 (0.71-1.10)	0.92 (0.63-1.33)	0.68
Stalk vegetables, sprouts ⁵					
Range (g/d)	<0.2	0.2 - <3.8	3.8 - <9.8	9.8	
Total No. of cases	2 110	2 671	2 600	2 650	
Model A [*]	1.00	0.94 (0.83-1.07)	0.79 (0.70-0.90)	0.92 (0.78-1.08)	0.38
Model B [†]	1.00	0.91 (0.70-1.18)	0.78 (0.68-0.91)	0.82 (0.63-1.07)	0.24
Other vegetables ⁵					
Range (g/d)	<3.4	3.4 - <10.2	10.2 - <23.0	23.0	
Total No. of cases	2 548	2 161	2 513	2 809	
Model A [*]	1.00	0.96 (0.88-1.05)	0.96 (0.85-1.08)	1.04 (0.86-1.27)	0.92
Model B [†]	1.00	1.01 (0.87-1.19)	0.90 (0.78-1.04)	0.96 (0.76-1.22)	0.36

HRs (and 95% CI) estimated within each country using Prentice-weighted Cox regression, with age as the underlying timescale, and combined across countries using random effects meta-analysis. Adjustment for covariates was performed using multivariable Prentice-weighted Cox regression.

* Model A: stratified by country and adjusted for centre and sex.

† Model B: stratified by country and adjusted for centre, sex, education level, BMI, physical activity level, smoking status, total energy intake, and alcohol intake.

¹ Models additionally adjusted for total vegetable intake.

² In analyses of fruit sub-types other fruit sub-types were adjusted for.

³ Umea (Sweden) excluded from analysis as there was no information on non-citrus fruit.

⁴ Models additionally adjusted for total fruit intake.

⁵Umea (Sweden) excluded from analysis as there was no information on leafy vegetables, cabbages, onion & garlic, stalk vegetables and sprouts, and the category 'other vegetables'.

⁶Denmark excluded from analysis as there was not enough information on leafy vegetables to calculate HRs and 95% CIs.

⁷France excluded from analysis as there was no information on onion & garlic.

Table 5
 Characteristics of studies identified by systematic review and included in the meta-analysis

Author	Country/ countries	Age (yrs)/sex/ years of follow-up	Exposure assessment	Highest intake category/ reference category *	Number of diabetes cases/ diabetes/ assessment criteria (if stated)	Exclusion criteria	Variables adjusted for in analyses [†]
Current results from EPIC- InterAct study	France, Germany, The Netherlands, UK, Italy, Sweden, Denmark, and Spain	40-79/ men & women/ 11	Country specific dietary questionnaires	FV: 8.71/2.13 Fruit: 5.39/0.75 Veg: 3.94/0.88 GLV: 5.93/0.05 RV: 3.87/0.12	10 821/ Self-report (self- report of doctor diagnosed diabetes, anti- diabetic drug use), linkage to primary and secondary care registers, medication use (drug registers), hospital admissions and mortality data with a coding for diabetes, local and national diabetes and pharmaceutical registers	Incomplete dietary information; extreme values for total energy intake (top or bottom 1% of the original EPIC study sample); missing information on potential confounding variables; diabetes, myocardial infarction or stroke at baseline.	Age, sex, centre, country, education level, BMI, physical activity level, smoking status, total energy intake, alcohol intake
Villegas et al., 2008 ³⁰	China (Shanghai Women's Health Study)	40-70/ women/ 4.6	77 item food frequency questionnaire. Mean of intake between baseline and follow- up used.	Fruit: 5.57/1.09 Veg: 5.35/1.52 GLV: 11.91/2.45 RV: 1.51/0.04	1 608/ Self-report (self-report of doctor diagnosed diabetes, glucose test history and/or use of hypoglycaemic medication)/ ADA 1997 criteria	<40 yrs or > 70 yrs; type 2 diabetes, cancer, or cardiovascular disease present at baseline; extreme total energy intake (<2,090 or >14,630 KJ/d).	Age, EI, meat intake, BMI, WHR, smoking, alcohol, PA, income level, education level, occupational status, hypertension
Bazzano et al., 2008 ⁷	USA (Nurses' Health Study)	30-55/ women/ 18	61 item food frequency questionnaire	F&V: 7.66/2.35 Fruit: 2.64/0.46 Veg: 5.40/1.61 GLV: 10.36/1.75	4 529/ Self-report with questionnaire (asking about symptoms, diagnostic tests, and treatments) / NDDG criteria before 1997, ADA criteria after 1997	Extreme values for total energy intake (<600 or >3,500 Kcal/d); 12 FFQ items left blank; diabetes or missing data for diabetes, cancer, or cardiovascular disease present at baseline; missing date for diabetes at follow-up	Age, BMI, PA, family history of diabetes, postmenopausal hormone use, alcohol use, smoking, EI
Montonen et al., 2005 ²⁹	Finland (Finnish Mobile Clinic Health Examination Survey)	40-69/ men & women/ 23	Diet history interview	Fruit: 2.17/0.0 Veg: 1.94/0.30 GLV: 4.55/0.39	383/ Social Insurance Institution nationwide register of persons receiving drug reimbursement (for diabetes)	<40 yrs or > 69 yrs; extreme values for total energy intake (<800 or >6,000 Kcal/d); pregnant women; diabetes at baseline.	Age, sex, BMI, EI, smoking, family history of diabetes, geographic area
Liu et al., 2004 ¹⁰	USA (Women's Health Study)	45/ women/ 8.8	131 item food frequency questionnaire	F&V: 10.16/2.54 Fruit: 3.91/0.62 Veg: 6.84/1.47	1 614/ Self-report of doctor diagnosed diabetes (with	Diabetes present at baseline	Age, smoking, EI, alcohol use, BMI, PA, history of

Author	Country/ countries	Age (yrs)/sex/ years of follow-up	Exposure assessment	Highest intake category/ reference category *	Number of diabetes cases/ diabetes/ assessment criteria (if stated)	Exclusion criteria	Variables adjusted for in analyses †
Ford and Mokdad, 2001 ⁸	USA (NHANES I)	25-74/ men & women/ 20	24-hour recall	GLV: 9.94/0.98 RV: 7.00/0.49	confirmation from different sources including: bloods, telephone interview, supplementary questionnaire, contact with doctor)	Not white or African American; insufficient information of diabetes status; prevalent diabetes; pregnant data for covariates	hypertension, history of high cholesterol, family history of diabetes
Meyer et al, 2000 11	USA (Iowa Women's Health Study)	55-69/ women/ 6	127 item food frequency questionnaire	F&V: 8.86/2.57 Fruit: 3.36/0.57 Veg: 5.93/1.57	1 018/ Self-report of doctor diagnosed diabetes, hospitalisation record ICD-9-CM; 250, death certificate ICD-9: 250	Extreme values for total energy intake (<2,510 or >20,920 KJ/d); 30 FFQ items left blank; diabetes at baseline	Age, sex, smoking status, systolic BP, cholesterol concentration, antihypertensive medication, recreational exercise, non-recreational activity, alcohol use, BMI, education

* Veg=vegetables; GLV=green leafy vegetables; RV=root vegetables; F&V, fruit and vegetables are in portions/d; GLV and RV are in portions/week.

† BMI=body mass index; EI=energy intake; WHR=waist-to-hip ratio; PA=physical activity; BP=blood pressure.

Note: Harding et al. paper not included in systematic review as those data are included within EPIC-InterAct.

Table 6

Pooled relative risks of diabetes (95% CI) for highest versus lowest category of estimated intake of total fruit and vegetables, fruit, vegetables, green leafy vegetables, and root vegetables in meta-analysis according to various study level characteristics

	Fruit & vegetables		Fruit		Vegetables		Green leafy vegetables		Root vegetables	
	No. of studies	RR (95% CI)	No. of studies	RR (95% CI)	No. of studies	RR (95% CI)	No. of studies	RR (95% CI)	No. of studies	RR (95% CI)
Sex										
Men & women	2	0.89 (0.80, 0.98)	2	0.91 (0.82, 1.01)	2	0.80 (0.48, 1.12)	2	0.66 (0.40, 0.93)	1	0.86 (0.76-0.99)
Women only	3	0.97 (0.88, 1.06)	4	0.94 (0.78, 1.09)	4	0.92 (0.74, 1.11)	3	0.88 (0.81, 0.95)	2	0.55 (0.00-1.19)
Duration of follow-up										
<10 years	2	1.04 (0.91, 1.16)	3	0.99 (0.90, 1.09)	3	0.90 (0.63, 1.18)	2	0.88 (0.74, 0.97)	2	0.55 (0.00-1.19)
10 years	3	0.90 (0.83, 0.97)	3	0.83 (0.69, 0.97)	3	0.91 (0.77, 1.05)	3	0.76 (0.56, 0.97)	1	0.86 (0.76-0.99)
Location										
USA	4	0.95 (0.85, 1.04)	3	0.94 (0.72, 1.16)	3	1.01 (0.92, 1.09)	2	0.90 (0.83, 0.97)	1	0.88 (0.78-0.99)
Europe	1	0.91 (0.82, 1.01)	2	0.91 (0.82, 1.01)	2	0.80 (0.48, 1.12)	2	0.66 (0.40, 0.93)	1	0.86 (0.76-0.99)
China	0	---	1	0.96 (0.82, 1.10)	1	0.64 (0.51, 0.80)	1	0.81 (0.69, 0.97)	1	0.23 (0.13-0.39)
Dietary assessment methods										
FFQ	4	0.94 (0.88, 1.01)	5	0.93 (0.82, 1.04)	5	0.92 (0.78, 1.06)	4	0.88 (0.81, 0.94)	3	0.66 (0.26-1.05)
Others	1	0.80 (0.61, 1.06)	1	0.76 (0.50, 1.16)	1	0.58 (0.31, 1.06)	1	0.52 (0.30, 0.88)	0	---

Table 7

Meta-analysis of medium and highest versus lowest intake categories of fruits and vegetables and risk of diabetes

	No. of studies	Type 2 diabetes Cases	Reference category	Pooled RR (95% CI), medium category	Pooled RR (95% CI), high category	Heterogeneity (I ²), %
Total fruit and vegetables ¹	5	19 123	1.00	0.97 (0.93, 1.00)	0.93 (0.87, 1.00)	10.1
Total fruit ²	6	20 096	1.00	0.93 (0.84, 1.02)	0.92 (0.81, 1.02)	62.5
Total vegetables ³	6	20 096	1.00	0.95 (0.88, 1.02)	0.89 (0.75, 1.03)	77.3
GLV ⁴	5	18 955	1.00	0.92 (0.87, 0.97)	0.84 (0.74, 0.94)	50.4
Root vegetables ⁵	3	14 043	1.00	0.80 (0.61, 0.99)	0.66 (0.26, 1.05)	97.1

¹Reference category= 2.0 portions/day; medium category= 5.0 portions/day; high category=8.0 portions/day

²Reference category= 1.0 portions/day; medium category= 2.5 portions/day; high category=4.0 portions/day

³Reference category= 1.0 portions/day; medium category= 3.0 portions/day; high category=5.0 portions/day

⁴Reference category= 1.0 portions/week; medium category= 5.0 portions/week; high category=9.0 portions/week

⁵Reference category= 0.0 portions/week; medium category= 2.0 portions/week; high category=4.0 portions/week