

Published in final edited form as:

Diabetologia. 2011 September ; 54(9): 2272–2282. doi:10.1007/s00125-011-2182-9.

The InterAct Project: An Examination of the Interaction of Genetic and Lifestyle Factors on the Incidence of Type 2 Diabetes in the EPIC Study

A full list of authors and affiliations appears at the end of the article.

Abstract

Background—Studying gene-lifestyle interaction may help to identify lifestyle factors that modify genetic susceptibility and uncover genetic loci exerting important subgroup effects. Adequately powered studies with prospective, unbiased, standardised assessment of key behavioural factors for gene-lifestyle studies are lacking.

Objective—To establish a type 2 diabetes case-cohort study designed to investigate how genetic and potentially modifiable lifestyle and behavioral factors, particularly diet and physical activity, interact in their influence on the risk of developing type 2 diabetes.

Methods—Funded by the Sixth European Framework Programme, InterAct consortium partners ascertained and verified incident cases of type 2 diabetes occurring in European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts between 1991 and 2007 from 8 of the 10 EPIC countries. A pragmatic, high sensitivity approach was used for case ascertainment including multiple sources at each EPIC centre, followed by diagnostic verification. Prentice-weighted Cox regression and random effects meta-analyses were used to investigate differences in diabetes incidence by age and sex.

Results—A total of 12,403 verified incident cases of type 2 diabetes occurred during 3.99 million person-years of follow-up of 340,234 EPIC participants eligible for InterAct. We defined a centre stratified subcohort of 16,154 individuals for comparative analyses. Individuals with incident diabetes that were randomly selected into the subcohort ($n=778$) were included as cases in the analyses. All prevalent diabetes cases were excluded from the study. InterAct cases were followed-up for an average of 6.9 years, 49.7% were men. Mean baseline age and age at diagnosis were 55.6 and 62.5 years, mean BMI and waist were 29.4 kg/m² and 102.7 cm in men, and 30.1 kg/m² and 92.8 cm in women, respectively. Risk of type 2 diabetes increased linearly with age, with an overall hazard ratio (95% CI) of 1.56 (1.48; 1.64) for a 10 year age difference, adjusted for sex. A male excess in the risk of incident diabetes was consistently observed across all countries, with a pooled hazard ratio of 1.51 (1.39; 1.64), adjusted for age.

Corresponding author (all authors listed at the end of the manuscript): Dr Claudia Langenberg, MRC Epidemiology Unit, Institute of Metabolic Science, Box 285, Addenbrooke's Hospital, Cambridge CB2 0QQ, United Kingdom, Tel +44 (0)1223 769137, Fax +44 (0)1223 330316, claudia.langenberg@mrc-epid.cam.ac.uk.

Duality of interest

All authors declare that there is no duality of interest associated with this manuscript. IB and her spouse own stock in the companies GlaxoSmithKline (GSK) and Incyte (INCY).

MeSH: type 2 diabetes, genetic epidemiology, gene-lifestyle interaction, cohort studies; epidemiologic research design;

Conclusions—InterAct is a large, well powered, prospective study which will inform our understanding of the interplay between genes and lifestyle factors on the risk of type 2 diabetes development.

Keywords

type 2 diabetes; genetic epidemiology; gene-lifestyle interaction; cohort studies; epidemiologic research design

Introduction

The identification of common genetic variants that are reproducibly associated with type 2 diabetes has accelerated considerably with the availability of results from genome wide association studies (GWAS) of prevalent cases and controls [1-6]. However, the discovered associations only account for a relatively small proportion of the heritable component [7]. Interactions between genetic factors and lifestyle exposures, gene-gene interaction, and genetic variation other than common single nucleotide polymorphisms (SNPs) are likely to be important factors that contribute to the remaining variance [8], but have not been systematically explored.

The existing case-control studies used to identify the genetic loci associated with type 2 diabetes are not optimally designed to investigate gene-environment or lifestyle interaction (GEI) since they do not have standardised assessment of key lifestyle, behavioural factors and do not have a prospective design in which those factors are assessed in an unbiased manner before the onset of disease. An optimal study design for investigating GEI is a case-cohort study nested within a large prospective cohort, as this combines the efficiency of the case-control design with the advantages of the longitudinal cohort approach with extensive prospective assessment of key exposures that are not subject to recall bias. Selecting a random subcohort (nested case-cohort study) rather than matched controls (nested case-control study) has the additional advantage that it facilitates the design and conduct of future case-cohort studies for other diseases occurring in the same background population or cohort.

The InterAct Consortium is funded by the Sixth European Framework Programme. It was initiated to investigate how genetic and lifestyle behavioural factors, particularly diet and physical activity, interact on the risk of developing diabetes and how knowledge about such interactions may be translated into preventative action. As part of the wider InterAct project, consortium partners have established a case-cohort study of incident type 2 diabetes (InterAct Study) based on cases occurring in European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts between 1991 and 2007 in 8 of 10 EPIC countries participating in InterAct.

The principal objectives of this report are to a) describe the InterAct Study design, population, and objectives, b) characterise the random subcohort and compare it to EPIC participants from each of the 8 participating European countries eligible for InterAct, and c) investigate characteristics of incident diabetes cases, including differences in diabetes risk by age and sex.

Methods

Participants and study design

The large prospective InterAct type 2 diabetes case-cohort study is coordinated by the MRC Epidemiology Unit in Cambridge and nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) [9]. EPIC was initiated in the late 1980s and involves collaboration between 23 research institutions across Europe in 10 countries (Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden and the United Kingdom). With the exception of Norway and Greece, all EPIC countries participated in the InterAct project, including a total of 455,680 participants (table 1). The majority of EPIC cohorts were recruited from the general population, with some exceptions [10]. French cohorts included women who were members of a health insurance scheme for school and university employees; Turin and Ragusa (Italy) and the Spanish centres included some blood donors. Participants from Utrecht (Netherlands) and Florence (Italy) were recruited via a breast cancer screening program. The majority of participants recruited by the EPIC Oxford (UK) centre consisted of vegetarian and “health conscious” volunteers from England, Wales, Scotland, and Northern Ireland [10].

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.

Measurements

As part of EPIC, standardised information was collected at baseline on lifestyle exposures. Information on socio-economic status, education, and occupation was collected by questionnaire.

The assessment of diet was undertaken using a self- or interviewer-administered dietary questionnaire, developed and validated within each country to estimate the usual individual food intakes of the study participants [10;11]. Additionally, in a stratified subsample of 36,900 participants, a standardised 24-hr recall of food intake was collected [11;12].

Physical activity was assessed at baseline using a brief questionnaire covering occupation and recreational activity [12;13]. Although an index of physical activity derived from this questionnaire had previously been validated against repeated objective measures of activity in the UK and in the Netherlands [13], no validation study has been conducted in any other country. Therefore, a study to test the validity of the questionnaire in European populations was conducted in InterAct, including 100 men and 100 women comparable to those originally recruited into the EPIC study from each participating InterAct country, with an objective measurement of physical activity by individually calibrated combined heart rate and movement sensing [14]. Results from this validation study will be published as a separate report.

Standard anthropometric data and biological samples (blood plasma, blood serum, white blood cells and erythrocytes) were collected from 385,747 of the 519,978 EPIC study participants and 346,055 of 455,680 individuals participating in 8 of the 10 EPIC cohorts

included in InterAct. Individuals without stored blood (n=109,625) or without information on reported diabetes status (n=5,821) were excluded, leaving a total of 340,234 participants with 3.99 million person-years of follow-up eligible for inclusion in InterAct (figure 1).

Samples were stored from collection at -196°C in liquid nitrogen at the coordinating centre at the International Agency for Research into Cancer (IARC) in Lyon, France, or in liquid nitrogen in local biorepositories with the exception of Umeå where -80°C freezers were used. Follow-up data on mortality and disease status has been ascertained via registries, clinical records, and other sources of clinical information [15;16]. At least one follow-up was conducted in each centre 3-5 years after baseline and questionnaires and telephone-based interviews were administered to repeat exposure measurement and collect self-reported health status data.

Type 2 diabetes case ascertainment and verification

We followed a pragmatic, high sensitivity approach for case ascertainment with the aim of identifying a) all potential incident diabetes cases and b) excluding all individuals with prevalent diabetes.

Prevalent diabetes was identified on the basis of baseline self-report of a history of diabetes, doctor diagnosed diabetes, diabetes drug use, or evidence of diabetes after baseline with a date of diagnosis earlier than the baseline recruitment date. All ascertained cases with any evidence of diabetes at baseline were excluded.

Ascertainment of incident type 2 diabetes involved a review of the existing EPIC datasets at each centre using multiple sources of evidence including self-report, linkage to primary care registers, secondary care registers, medication use (drug registers), hospital admissions and mortality data (online appendix; table ST1). Information from any follow-up visit or external evidence with a date later than the baseline visit was used. Cases in Denmark and Sweden were not ascertained by self-report, but identified via local and national diabetes and pharmaceutical registers and hence all ascertained cases were considered to be verified (online appendix; table ST1).

To increase the specificity of the case definition for centres other than those from Denmark and Sweden, we sought further evidence for all cases with information on incident type 2 diabetes from fewer than 2 independent sources at a minimum, including individual medical records review in some centres. Follow-up was censored at the date of diagnosis, the 31st of December 2007 or the date of death, whichever occurred first. In total, 12,403 verified incident cases were identified; there were 471 cases in the first year of follow-up and 587 in the second year. Sample size calculations are included in the online appendix (online appendix; figure SF1).

Dates of diagnosis

The date of diagnosis for incident cases was set as either the date of diagnosis reported by the doctor, the earliest date that diabetes was recorded in medical records, the date of inclusion into the diabetes registry, the date reported by the participant, or the date of the questionnaire in which diabetes was first reported. If the date of diagnosis could not be

ascertained from any of the sources listed above, the midpoint between recruitment and censoring was used (n=421).

Case-cohort design

The case-cohort design of the InterAct study differs from the nested case-control design in that a random subcohort is selected instead of a set of matched controls. Because only a subset of the original cohort is randomly selected into the subcohort, cases are overrepresented in the case-cohort set, which needs to be accounted for in the analysis methods, as outlined in the online appendix (supplementary methods).

Subcohort

A subcohort of 16,835 individuals was randomly selected from those with available stored blood and buffy coat, stratified by centre. We oversampled the number of individuals in the subcohort for the proportion of prevalent diabetes cases in each centre to account for later exclusion of individuals with prevalent diabetes from InterAct analyses. After exclusion of 548 individuals with prevalent diabetes, 129 individuals without information on reported diabetes status, and 4 individuals with post-censoring diabetes, 16,154 subcohort individuals were included in the analysis. Due to the random selection, this subcohort also included a random set of 778 individuals who had developed incident type 2 diabetes during follow-up.

Stored samples, genotyping and biomarker measurement

Details of the quality, quantity and availability of stored samples can be found in the online appendix (supplementary methods), together with a description of the InterAct strategy for genotyping and biomarker measurement.

Statistical analyses

Characteristics of the InterAct incident cases are described using summary statistics (means, standard deviations, frequencies and percentages) separately for men and women, and overall. Characteristics of the randomly selected subcohort are also summarised, alongside summaries from the overall EPIC cohort from which it was sampled, to provide some indication of the representativeness of the subcohort compared with the whole of EPIC. Comparison p-values were not calculated for these two groups, as due to the large sample size, even very small, clinically negligible differences in the distribution of a particular characteristic are likely to be statistically significant. Prentice-weighted Cox regression models and random effects meta-analyses were used, as described in more detail in the online appendix (supplementary methods), to investigate differences in the incidence of diabetes by sex and age. Crude and age-standardised incidence rates were calculated within each country.

Results

A total of 12,403 incident cases of type 2 diabetes were ascertained and verified (figure 1) during 3.99 million person-years of follow-up of 340,234 EPIC participants (mean follow-up 11.7 years), excluding individuals without stored blood (n=109,625) or without information on reported diabetes status (n=5,821). The total number of incident cases in

InterAct further excluded a total of 2,577 verified Danish cases for logistical reasons, as local sample retrieval and DNA extraction of more than the originally anticipated 2,000 cases was not feasible in the required timeframe. A random subcohort of 16,835 individuals was selected; after exclusion of 548 individuals with prevalent diabetes, 129 individuals with unknown and 4 with post-censoring diabetes status, 16,154 individuals were included in the subcohort for InterAct analyses. Due to the random selection, this subcohort also included a random set of 778 individuals who had developed incident diabetes during follow-up.

Characteristics of individuals in the random subcohort

Country specific baseline characteristics of the random subcohort were similar to those in the overall EPIC population eligible for inclusion for each country (online appendix; table ST2). The mean age of participants in the random subcohort was 52.5 years, a total of 38.4% were men (online appendix; table ST2). Average BMI and waist were 26.1 kg/m² and 86.8 cm (26.7 kg/m² and 95.4 cm in men, 25.8 kg/m² and 81.6 cm in women), respectively. A total of 57.0% of participants reported being physically inactive or moderately inactive, 34.6% were educated at secondary level or above, and 46.4% never smoked. Family history of diabetes was not ascertained in Italy, Spain, Heidelberg or Oxford; 8,832 of the 16,835 subcohort participants had information on family history, in whom it was positive in 1,628 individuals (18.4% of those with data, or 9.7% of the subcohort).

Characteristics of type 2 diabetes cases

Characteristics of ascertained and verified incident cases are shown in table 2; country specific information is provided in the online appendix (table ST3). Including the 2,577 of the 4,632 Danish cases with available blood samples that are not part of InterAct, the overall incidence in InterAct was 3.76 per 1000 person-years of follow-up, based on 14,980 verified diabetes cases occurring during 3,989,345 person-years. Crude incidences ranged from 1.4/1000 person-years in French women to 7.4/1000 (men 8.9, women 6.0) person-years in Denmark (online appendix; table ST3). For all analyses other than the calculation of crude and standardised incidence rates the 2,577 additional Danish cases are excluded, leading to a total of 12,403 InterAct cases (14,980-2,577=12,403). Of the 12,403 InterAct cases with an average follow-up of 6.9 years, 49.7% were men (n=6,165). Mean baseline age and age at diagnosis were 55.6 and 62.5 years, mean BMI and waist were 29.7 kg/m² and 97.7 cm (29.4 kg/m² and 102.7 cm in men, 30.1 kg/m² and 92.8 cm in women), respectively (table 2). A total of 30.8% of all InterAct cases reported a positive family history of diabetes (26.3% of men, 35.3% of women), excluding centres which did not obtain this information (table 2).

Associations of age and sex with incident type 2 diabetes

Men showed a significantly greater risk of incident diabetes than women (figure 2), with a pooled hazard ratio (95% confidence interval) of 1.51 (1.39; 1.64). Despite a consistent male excess of diabetes risk across all centres and countries, some heterogeneity in the effect of sex was present (I² 57%). Adjusting the centre-specific effects of sex for waist, but not BMI explained some of the heterogeneity (I² was reduced to 33%). Diabetes incidence increased linearly with age (figure 3), with an overall, pooled hazard ratio of 1.56 (1.48; 1.64) for a 10

year age difference (1.44 (1.35; 1.55) in men and 1.64 (1.55; 1.74) in women); the apparently substantial heterogeneity (I^2 73% overall, 71% in men, 61% in women) was mainly due to the larger than average and statistically significant effect in the Bilthoven cohort (hazard ratio 2.29 in men and 2.28 in women).

Associations of measures of obesity, smoking, alcohol intake, physical activity, socioeconomic status and dietary information with diabetes are each the subject of separate InterAct reports.

Discussion

Type 2 diabetes is an increasingly common, complex disease that clusters in families and is influenced by genetic and lifestyle factors. Despite progress in the identification of common genetic variants through genome-wide meta-analyses of diabetes case-control studies, current studies are lacking the power and prospective design to investigate interaction between genes and lifestyle. Effect sizes of type 2 diabetes loci identified to date are small and explain only a small proportion of familial clustering [17;18].

Heterogeneity in effects across studies exists due to the different design and case ascertainment of cross-sectional studies contributing to earlier GWAS [19]. In addition, heterogeneity in effects of established or yet unidentified loci between population subgroups is likely, due to the multifaceted aetiology of type 2 diabetes. Varying effects between subgroups defined by disease characteristics such as early versus late diabetes onset or with versus without positive family history may reflect different relative contributions of genetic versus lifestyle factors. Not accounting for potential sources of heterogeneity may lead to important subgroup effects being overlooked, genetic variants not being identified and the genetic variance explained being underestimated.

Understanding differences in how genetic susceptibility translates into diabetes risk between subgroups defined by lifestyle or health behavioural factors such as obese versus lean or sedentary versus active has the potential to inform strategies for disease prevention, as previously suggested for a common variant in the *TCF7L2* gene in high risk individuals in the Diabetes Prevention Programme [20]. However, existing general population cohorts with prospective data on type 2 diabetes are underpowered to systematically investigate gene-lifestyle interaction. We have therefore set up the European InterAct Study whose nested case-cohort design combines the advantages of a prospective cohort with the efficiency and power of a large case-control study.

Strength and weaknesses

The use of cases and a random cohort nested in EPIC enabled Interact partners to jointly ascertain and verify a total of 12,403 incident type 2 diabetes cases in a relatively short time frame, exceeding the originally estimated number. Extraction of DNA in cases and in the randomly selected subcohort and storage in a comPOUND® (TTP LabTech, Cambridge, UK) automatic DNA handling system allows rapid genotyping that together with standardised baseline information on participants' clinical characteristics and health behaviour enables investigation of the interaction between genes and lifestyle factors.

Genome-wide data on a stratified InterAct sample of 10,000 participants provides the opportunity for discovery of as yet unidentified variants whose larger subgroup effects may have introduced sufficient heterogeneity in non-stratified, conventional GWAS to prevent them from reaching the stringent significance levels required in this setting.

In addition to the prospective design that minimises systematic error introduced by recall bias, advantages of the InterAct case-cohort design include the time- and cost-efficiency and maximal sharing of resources that can be achieved by sharing of the randomly selected subcohort. In a traditional prospective cohort study such as EPIC, initially disease free individuals are followed up over time to observe the rate of occurrence of binary clinical events (e.g. diabetes) in relation to information obtained at baseline. The major advantage is that this approach avoids the problem of recall bias, but it involves follow-up of a large study population over many years and is time consuming and expensive. This is especially true when the characterisation of possible exposures is expensive and inefficient since it is obtained in the whole cohort, only a small proportion of which go on to become cases. In a case-cohort study, efficiency is optimised by only obtaining additional exposure information for participants experiencing the outcome of interest and for members of a random sample (subcohort) selected from the entire cohort independent of the outcome. An added advantage is that this subcohort can be used as a comparison cohort for different outcomes of interest. Case-cohort studies can be designed within large cohorts with blood samples, DNA or other materials stored at baseline for later exposure measurement.

Subcohort participants were shown to be representative of EPIC participants eligible for inclusion in InterAct within each country, and thus provide an excellent opportunity for data sharing with research groups studying other disease outcomes within EPIC. Sensitivity analyses calculating hazard ratios from unweighted Cox regression using full EPIC cohort data compared to weighted Cox regression using case-subcohort data showed that results were comparable for sex, but differed somewhat for age.

The detailed, standardized evaluation of dietary and lifestyle factors on a Europe-wide scale, including assessment of nutritional biomarkers in addition to dietary self-report, optimisation of dietary data through calibration, and validation of physical activity questionnaires against objective measures, will help to address questions that have not been resolved due to inconclusive results from earlier prospective studies or intervention trials. The study of different European populations with considerable heterogeneity in dietary habits and health behaviours will increase the generalisability of any potential findings. Heterogeneity in the confounding structure across countries will help to minimise the identification of false positive, non-causal associations. The large number of cases in InterAct will be instrumental in being able to investigate dose-response effects of important exposures in more detail and inform decisions about appropriate thresholds with greater precision of estimates.

Although the possibility of examining the consistency of main genetic and interaction effects across the 8 European countries will help to understand factors contributing to any potential heterogeneity, results from our European descent InterAct participants do not allow inference about other ethnic groups with different allele frequencies and distribution and

determinants of health behaviours. We used a clinical definition of type 2 diabetes that did not rely on glucose measurement. This means that although our case definition is specific due to the verification process, InterAct rates reflect the incidence of clinical type 2 diabetes and our case number would be even larger had it been possible to identify undiagnosed diabetes cases. Likewise, we have not ruled out undiagnosed diabetes in the subcohort, which may lead to an underestimation of main effects and reduced power for interaction analyses. Despite great efforts to standardise the clinical case definition of our study, some heterogeneity exists between centres due to differences in the information that was available locally for case ascertainment and verification. In addition, not all EPIC cohorts recruited participants from the general population [9;10], potentially limiting the generalisability of our findings. Centre and country specific absolute incidence rates therefore need to be generalised and compared with caution. However, any bias due to differences in the case definition between countries would need to be consistent across centres and countries to lead a false positive association in our meta-analyses, the results of which also provide information about heterogeneity and hence to some degree generalisability of findings. The prevalence of important diabetes risk factors has changed since EPIC participants were recruited, with an ageing European population and greater prevalence of overweight and obesity; however, whether or not their relative importance for diabetes risk or potential interactions with genetic susceptibility has also changed remains unknown.

Diabetes incidence by country, age and sex

Despite the increasing prevalence of type 2 diabetes and associated economic and public health burden, [21;22] data on incidence rates from population-based European studies are scarce. Previous studies identifying clinical as well as undiagnosed, asymptomatic diabetes, have shown rates varying from 3.03 per 1000 person-years in Sweden [23] to 19.1 in Southern Spain [24], with estimates in-between for other European countries or regions [25-29]. The precision of these estimates from individual studies is naturally limited due to the relatively small cohort sizes and numbers of new cases; in addition, rates are expected to be higher compared to those based on a clinical diagnosis alone, as used in InterAct.

In contrast to many prevalence studies universally showing an increase of type 2 diabetes with age, data on differences in the incidence of type 2 diabetes in Europe are sparse. We observed a linear increase in diabetes risk by age across the 8 countries included in InterAct, with an overall 56% increase in risk for a decade of age (men 45%, women 64%). This trend, in the context of an ageing European population, will lead to a rise in the incidence of diabetes if effective prevention measures are not implemented now. We also found a consistent male excess in diabetes incidence across all countries, with an overall 50% higher risk in men, compared to women. Differences in waist circumference accounted for some of the heterogeneity in this association between countries. The previous literature on sex differences in incident diabetes largely supports this finding, with most [24;26;29-31], but not all studies [27] reporting a male excess in risk. Future InterAct work will identify factors that underlie and contribute to differences in diabetes incidence by sex and age.

The aims of the InterAct Project extend beyond the establishment of the InterAct case-cohort study and investigation of gene-lifestyle interaction using prospective observational

studies. To enable InterAct partners to take gene-lifestyle interactions identified in the observational part of the study forward to a trial setting, a consortium of lifestyle intervention diabetes prevention trials ready for genotyping has been established. To better understand the impact of genetic risk information for preventive action, a Cochrane systematic review of the provision of risk information on emotion, cognition and behaviour has been completed, suggesting that the communication of DNA-based disease risk estimates may have little or no effect on behaviour, but may have a small effect on intentions to change behaviour [32].

Conclusions

In summary, the InterAct Study is a large scale collaborative endeavour with the potential to improve our understanding of gene-lifestyle interaction by using sufficiently powered, prospective observational data including 12,403 incident type 2 diabetes cases and a random subcohort of 16,835 individuals. The establishment of a well characterised random subcohort will aid the design and conduct of future case-cohort studies nested within EPIC cohorts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

The InterAct Consortium, C Langenberg¹, S Sharp¹, NG Forouhi¹, P Franks², MB Schulze³, N Kerrison¹, U Ekelund¹, I Barroso^{4,45}, S Panico⁵, M Tormo⁶, J Spranger⁷, S Griffin¹, YT van der Schouw⁸, P Amiano^{9,46,47}, E Ardanaz^{10,47}, L Arriola^{9,46,47}, B Balkau^{11,48}, A Barricarte¹², JWJ Beulens⁸, H Boeing¹³, HB Bueno-de-Mesquita¹⁴, BB Buijsse¹⁵, MD Chirlaque Lopez^{16,47}, F Clavel-Chapelon¹⁷, FL Crowe¹⁸, B de Lauzon-Guillan^{17,48}, P Deloukas¹⁹, M Dorronsoro^{9,46,47}, DD Drogan¹⁵, P Froguel²⁰, C Gonzalez²¹, S Grioni²², L Groop^{23,49}, C Groves²⁴, P Hainaut²⁵, J Halkjaer²⁶, G Hallmans²⁷, T Hansen²⁸, R Kaaks²⁹, TJ Key¹⁸, K Khaw³⁰, A Koulman³¹, A Mattiello³², C Navarro^{16,47}, P Nilsson³³, T Norat³⁴, K Overvad³⁵, L Palla¹, D Palli³⁶, O Pedersen^{28,50,54}, PH Peeters⁸, JR Quirós³⁷, A Ramachandran³⁸, L Rodriguez-Suarez³⁷, O Rolandsson²⁷, D Romaguera³⁴, I Romieu²⁵, C Sacerdote^{39,51}, M Sánchez^{40,47}, A Sandbaek⁴¹, N Slimani²⁵, I Sluijs⁸, AMW Spijkerman¹⁴, B Teucher²⁹, A Tjonneland²⁶, R Tumino⁴², DL van der A¹⁴, WMM Verschuren¹⁴, J Tuomilehto^{43,52}, E Feskens⁴⁴, M McCarthy^{24,53,55}, E Riboli³⁴, and NJ Wareham¹

Affiliations

¹MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Box 285, Cambridge, CB2 0QQ, United Kingdom ²Department of Clinical Sciences, Clinical Research Center, Malmö General Hospital (UMAS) Lund University, 20502 Malmö, Sweden ³Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558

Nuthetal, Germany ⁴The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA ⁵Dipartimento di Medicina Clinica e Sperimentale, Università Federico II, Department of Clinical and Experimental Medicine, Federico II University, via Pansini 5 - 80131 Napoli, Italy ⁶Jefa Sección de Estudios Epidemiológicos, C/ Luis Fontes Pagán, nº9, 1a Planta, (Edif. EMI Hospital Universitario "Reina Sofía"), 30003 Murcia, Spain ⁷Charité-Universitätsmedizin Berlin, Dept of Endocrinology, Diabetes and Nutrition, Hindenburgdamm 30, 12200 Berlin, Germany ⁸Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Stratenum 6.131, PO Box 85.500, 3508 GA Utrecht, The Netherlands ⁹Public Health Division of Gipuzkoa, Basque Government, Av. Navarra 4, 20013 San Sebastian, Spain ¹⁰Navarre Public Health Institute, Pamplona, Leyre 15, 31003 Pamplona, Navarra, Spain ¹¹Inserm, CESP Centre for research in Epidemiology and Population Health, U1018: Epidemiology of diabetes, obesity and chronic kidney disease over the lifecourse. 16 Ave Paul Vaillanat Couturier, 94807 Villejuif cedex, France ¹²Epidemiology, Prevention and Promotion Health Service, Leyre 15, 31003 Pamplona, Navarra, Spain ¹³Epidemiology Dept, Arthur-Scheunert-Allee 114-116, 14558 Potsdam, Germany ¹⁴National Institute for Public Health and the Environment (RIVM), PO Box 1, 3720 BA Bilthoven, The Netherlands ¹⁵Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558, Nuthetal, Germany ¹⁶Murcia Regional Health Authority, Ronda de Levante, 11. E30008 Murcia, España ¹⁷Inserm, CESP Centre for research in Epidemiology and Population Health, U1018: Nutrition, hormones and women's health, IGR, 39 rue Camilles Desmouolins, 94800 Villejuif cedex, France ¹⁸Cancer Epidemiology Unit, University of Oxford, Richard Doll Building, Nuffield Department of Clinical Medicine, Roosevelt Drive, Oxford OX3 7LF, United Kingdom ¹⁹Department of Human Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA ²⁰Section of Genetics, Imperial College London Genome Centre, Hammersmith Campus, Du Cane Road, London, W12 0NN ²¹Unit Nutrition, Environment and Cancer, Department of Epidemiology, Gran Via s/n km 2,7, L'Hospitalet de Llobregat, 08907 Barcelona, Spain ²²Fondazione IRCCS Istituto Nazionale Tumori Milan, Via Venezian, 1, 20133 Milan, Italy ²³Department of Clinical Sciences, Diabetes & Endocrinology, University Hospital Scania, 20502 Malmö, Sweden ²⁴Oxford Centre for Diabetes, Endocrinology and Metabolism (OCDEM), University of Oxford, Churchill Hospital, Old Road, Headington, Oxford, OX3 7LJ ²⁵International Agency for Research of Cancer, 150 Cours Albert Thomas, 69372 Cedex 08 Lyon, France ²⁶Danish Cancer Society, Strandboulevarden 49, 2100 Copenhagen, Denmark ²⁷Department of Public Health and Clinical Medicine, Umea University, 90187 Umea, Sweden ²⁸Hagedorn Research Institute, Niels Steensens Vej 2, 2820 Gentofte, Copenhagen, Denmark ²⁹German Cancer Research Centre, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany ³⁰Department of Public Health and Primary Care, University of Cambridge, Addenbrooke's Hospital, Box 251, Cambridge CB2 2QQ, UK ³¹MRC Human Nutrition Research, Elsie Widdowson Laboratory, Fulbourn Road, Cambridge, CB1 9NL, United Kingdom

³²Department of Clinical and Experimental Medicine Federico II University, Via Pansini, 5, 80131 Naples, Italy ³³University of Lund, Department of Clinical Sciences Medicine, University Hospital Scania, 205 02 Malmö, Sweden
³⁴Department of Epidemiology & Biostatistics, School of Public Health, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG ³⁵Department of Epidemiology, School of Public Health, Bartholins Allé 2, 8000 Aarhus C, Denmark ³⁶Molecular&Nutritional Epidemiology Unit, Cancer Research and Prevention Institute (ISPO), Via Cosimo il Vecchio 2, 50139, Florence Italy
³⁷Consejería de Salud y Servicios Sanitarios, C/Ciriaco Miguel Vigil 9, 33006-Oviedo-Asturias ³⁸India Diabetes Research Foundation, No 28, Marshall's Road, Egmore, Chennai – 600008, Tamilnadu, India ³⁹Center for Cancer Prevention (CPO-Piemonte), Via Santena 7, 10126, Torino, Italy. ⁴⁰Andalusian School of Public Health, Cuesta del Observatorio 4, Ap. Correos 2070, 18080 Granada, Spain
⁴¹School of Public Health, Department of General Practice, Bartholins Allé 2, DK-8000 Aarhus C, Denmark ⁴²Cancer Registry and Histopathology Unit, "Civile - M.P. Arezzo" Hospital, ASP 7 Via Dante Nr. 109, 97100 Ragusa, Italy ⁴³Department of Public Health, University of Helsinki, P.O. Box 41, 00014 Helsingin Yliopisto, Finland ⁴⁴Division of Human Nutrition - Section Nutrition and Epidemiology, University of Wageningen, PO Box 8129, 6700 EV Wageningen, The Netherlands
⁴⁵University of Cambridge Metabolic Research Labs, Institute of Metabolic Science Addenbrooke's Hospital CB2 0QQ Cambridge ⁴⁶Instituto BIO-Donostia, Basque Government ⁴⁷CIBER Epidemiología y Salud Pública - CIBERESP, Spain
⁴⁸University Paris Sud 11, UMR5 1018, 94807 Villejuif cedex, France ⁴⁹Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland
⁵⁰Institute of Biomedicine, Faculty of Health, University of Copenhagen ⁵¹Human Genetic Foundation (HuGeF) ⁵²National Public Health Institut, Department of Health Promotion and Chronic Disease Prevention Diabetes Unit, Mannerheimintie 166, 00300 Helsinki, Finland ⁵³Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, UK ⁵⁴Center of Basic Metabolic Research, Faculty of Health Sciences, University of Copenhagen ⁵⁵Oxford NIHR Biomedical Research Centre, Churchill Hospital, Old Road, Headington, Oxford, OX3 7LJ

Acknowledgments

We thank all EPIC participants and staff for their contribution to the study. We thank staff from the Technical, Field Epidemiology and Data Functional Group Teams of the Medical Research Council Epidemiology Unit in Cambridge, UK, for carrying out sample preparation, DNA provision and quality control, genotyping and data handling work. We specifically thank Sarah Dawson for coordinating the sample provision for biomarker measurements, Abigail Britten for coordinating DNA sample provision and genotyping of candidate markers, Nicola Kerrison, Chris Gillson and Abigail Britten for data provision and genotyping quality control, Matt Sims for writing the technical laboratory specification for the intermediate pathway biomarker measurements and for overseeing the laboratory work.

The InterAct study received funding from the European Union (Integrated Project LSHM-CT-2006-037197 in the Framework Programme 6 of the European Community). In addition, InterAct investigators acknowledge funding from the following agencies: MT: Health Research Fund (FIS) of the Spanish Ministry of Health; the CIBER en Epidemiología y Salud Pública (CIBERESP), Spain; Murcia Regional Government (N° 6236); JS: JS was

supported by a Heisenberg-Professorship (SP716/2-1), a Clinical Research Group (KFO218/1) and a research group (Molecular Nutrition to JS) of the Bundesministerium für Bildung und Forschung (BMBF); YTvds, JWJB, PHP, IS: Verification of diabetes cases was additionally funded by NL Agency grant IGE05012 and an Incentive Grant from the Board of the UMC Utrecht; HBBdM: 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 (The Netherlands); MDCL: Health Research Fund (FIS) of the Spanish Ministry of Health; Murcia Regional Government (N° 6236); FLC: Cancer Research UK; PD: Wellcome Trust; LG: Swedish Research Council; GH: The county of Västerbotten; RK: Deutsche Krebshilfe; TJK: Cancer Research UK; KK: Medical Research Council UK, Cancer Research UK; AK: Medical Research Council (Cambridge Lipidomics Biomarker Research Initiative); CN: Health Research Fund (FIS) of the Spanish Ministry of Health; Murcia Regional Government (N° 6236); KO: Danish Cancer Society; OP: Faculty of Health Science, University of Aarhus, Denmark; JRQ: Asturias Regional Government; LRS: Asturias Regional Government; AT: Danish Cancer Society; RT: AIRE-ONLUS Ragusa, AVIS-Ragusa, Sicilian Regional Government; DLvdA, WMMV: 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 (The Netherlands); MMC: Wellcome Trust (083270/Z/07/Z), MRC (G0601261).

Abbreviations

| | |
|-------------|--|
| SD | standard deviation |
| HR | hazard ratio |
| CI | confidence interval |
| EPIC | European Prospective Investigation into Cancer and Nutrition |

References

- Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. Feb 22; 2007 445(7130):881–5. [PubMed: 17293876]
- Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. Jun 1; 2007 316(5829):1331–6. [PubMed: 17463246]
- Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science*. Jun 1; 2007 316(5829):1341–5. [PubMed: 17463248]
- Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. Jun 1; 2007 316(5829):1336–41. [PubMed: 17463249]
- Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet*. Jun; 2007 39(6):770–5. [PubMed: 17460697]
- Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*. May; 2008 40(5):638–45. [PubMed: 18372903]
- Altshuler D, Daly M. Guilt beyond a reasonable doubt. *Nature genetics*. Jul; 2007 39(7):813–5. [PubMed: 17597768]
- Manolio TA. Cohort studies and the genetics of complex disease. *Nature genetics*. Jan; 2009 41(1):5–6. [PubMed: 19112455]
- Bingham S, Riboli E. Diet and cancer--the European Prospective Investigation into Cancer and Nutrition. *Nat Rev Cancer*. Mar; 2004 4(3):206–15. [PubMed: 14993902]
- Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr*. Dec; 2002 5(6B):1113–24. [PubMed: 12639222]

- (11). Slimani N, Deharveng G, CharrondiFre RU, van Kappel AL, Ockq MC, Welch A, 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. *Computer Methods and Programs in Biomedicine*. Mar; 1999 58(3):251–66. [PubMed: 10094230]
- (12). Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A, et al. EPIC-Norfolk: study design and characteristics of the cohort. *European Prospective Investigation of Cancer*. *Br J Cancer*. Jul; 1999 80(Suppl 1):95–103. [PubMed: 10466767]
- (13). Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, 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*. Jun; 2003 6(4):407–13. [PubMed: 12795830]
- (14). Brage S, Ekelund U, Brage N, Hennings MA, Froberg K, Franks PW, et al. Hierarchy of individual calibration levels for heart rate and accelerometry to measure physical activity. *J Appl Physiol*. Aug; 2007 103(2):682–92. [PubMed: 17463305]
- (15). Martyn CN, Barker DJ, Osmond C. Mothers' pelvic size, fetal growth, and death from stroke and coronary heart disease in men in the UK. *Lancet*. Nov 9; 1996 348(9037):1264–8. [PubMed: 8909378]
- (16). Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K, et al. General and abdominal adiposity and risk of death in Europe. *N Engl J Med*. Nov 13; 2008 359(20):2105–20. [PubMed: 19005195]
- (17). McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. *Hum Mol Genet*. Oct 15; 2008 17(R2):R156–R165. [PubMed: 18852205]
- (18). McCarthy MI, Zeggini E. Genome-wide association studies in type 2 diabetes. *Curr Diab Rep*. Apr; 2009 9(2):164–71. [PubMed: 19323962]
- (19). Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, Strachan DP, et al. Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes*. Feb; 2009 58(2):505–10. [PubMed: 19056611]
- (20). Florez JC, Jablonski KA, Bayley N, Pollin TI, de Bakker PI, Shuldiner AR, et al. TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. *N Engl J Med*. Jul 20; 2006 355(3):241–50. [PubMed: 16855264]
- (21). Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care*. May; 2004 27(5):1047–53. [PubMed: 15111519]
- (22). Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract*. Jan; 2010 87(1):4–14. [PubMed: 19896746]
- (23). Jansson SP, Andersson DK, Svardsudd K. Prevalence and incidence rate of diabetes mellitus in a Swedish community during 30 years of follow-up. *Diabetologia*. Apr; 2007 50(4):703–10. [PubMed: 17268796]
- (24). Soriguer F, Rojo-Martinez G, Almaraz MC, Esteva I, Ruiz de Adana MS, Morcillo S, et al. Incidence of type 2 diabetes in southern Spain (Pizarra Study). *Eur J Clin Invest*. Feb; 2008 38(2):126–33. [PubMed: 18226046]
- (25). Forouhi NG, Luan J, Hennings S, Wareham NJ. Incidence of Type 2 diabetes in England and its association with baseline impaired fasting glucose: the Ely study 1990-2000. *Diabet Med*. Feb; 2007 24(2):200–7. [PubMed: 17257284]
- (26). Rathmann W, Strassburger K, Heier M, Holle R, Thorand B, Giani G, et al. Incidence of Type 2 diabetes in the elderly German population and the effect of clinical and lifestyle risk factors: KORA S4/F4 cohort study. *Diabet Med*. Dec; 2009 26(12):1212–9. [PubMed: 20002472]
- (27). Bonora E, Kiechl S, Willeit J, Oberhollenzer F, Egger G, Meigs JB, et al. Population-based incidence rates and risk factors for type 2 diabetes in white individuals: the Bruneck study. *Diabetes*. Jul; 2004 53(7):1782–9. [PubMed: 15220202]
- (28). Arteagoitia JM, Larranaga MI, Rodriguez JL, Fernandez I, Pinies JA. Incidence, prevalence and coronary heart disease risk level in known Type 2 diabetes: a sentinel practice network study in the Basque Country, Spain. *Diabetologia*. Jul; 2003 46(7):899–909. [PubMed: 12830379]

- (29). Valdes S, Botas P, Delgado E, Alvarez F, Cadorniga FD. Population-based incidence of type 2 diabetes in northern Spain: the Asturias Study. *Diabetes Care*. Sep; 2007 30(9):2258–63. [PubMed: 17536076]
- (30). Meisinger C, Thorand B, Schneider A, Stieber J, Doring A, Lowel H. Sex differences in risk factors for incident type 2 diabetes mellitus: the MONICA Augsburg cohort study. *Arch Intern Med*. Jan 4; 2002 162(1):82–9. [PubMed: 11784224]
- (31). Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. *BMJ*. 2009; 338:b880. [PubMed: 19297312]
- (32). Marteau TM, French DP, Griffin SJ, Prevost AT, Sutton S, Watkinson C, et al. Effects of communicating DNA-based disease risk estimates on risk-reducing behaviours. *Cochrane Database of Systematic Reviews*. 2010; (10)

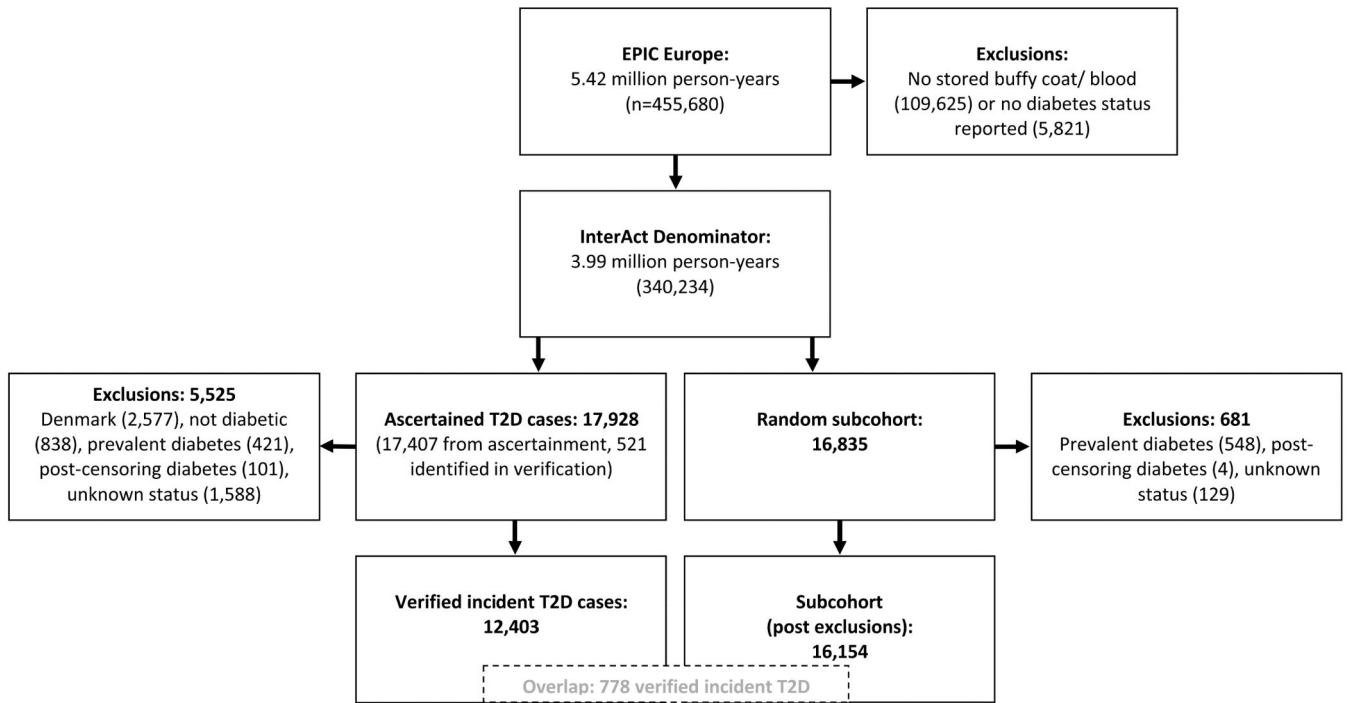


Figure 1. Overview of the InterAct type 2 diabetes (T2D) case-cohort study nested within 8 of the 10 EPIC Europe countries.

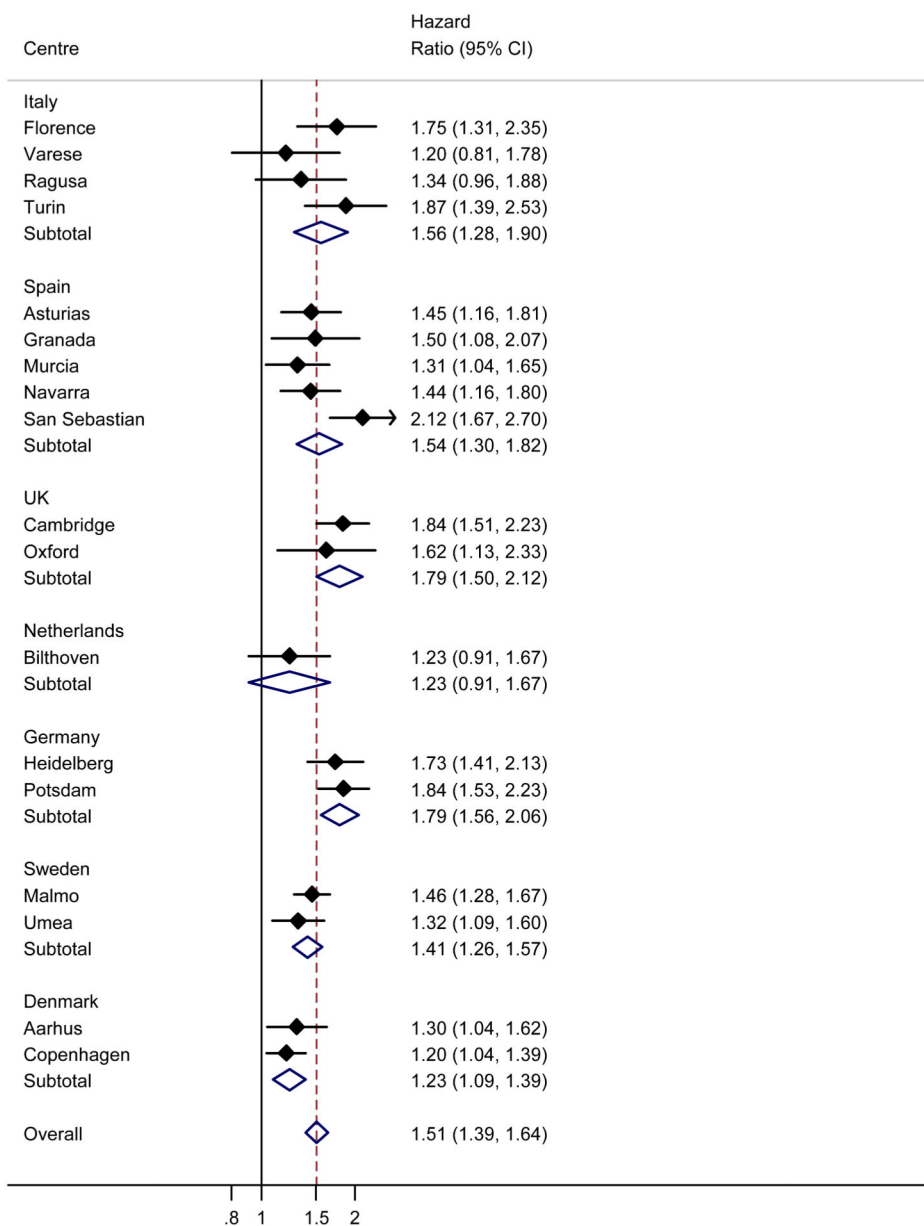
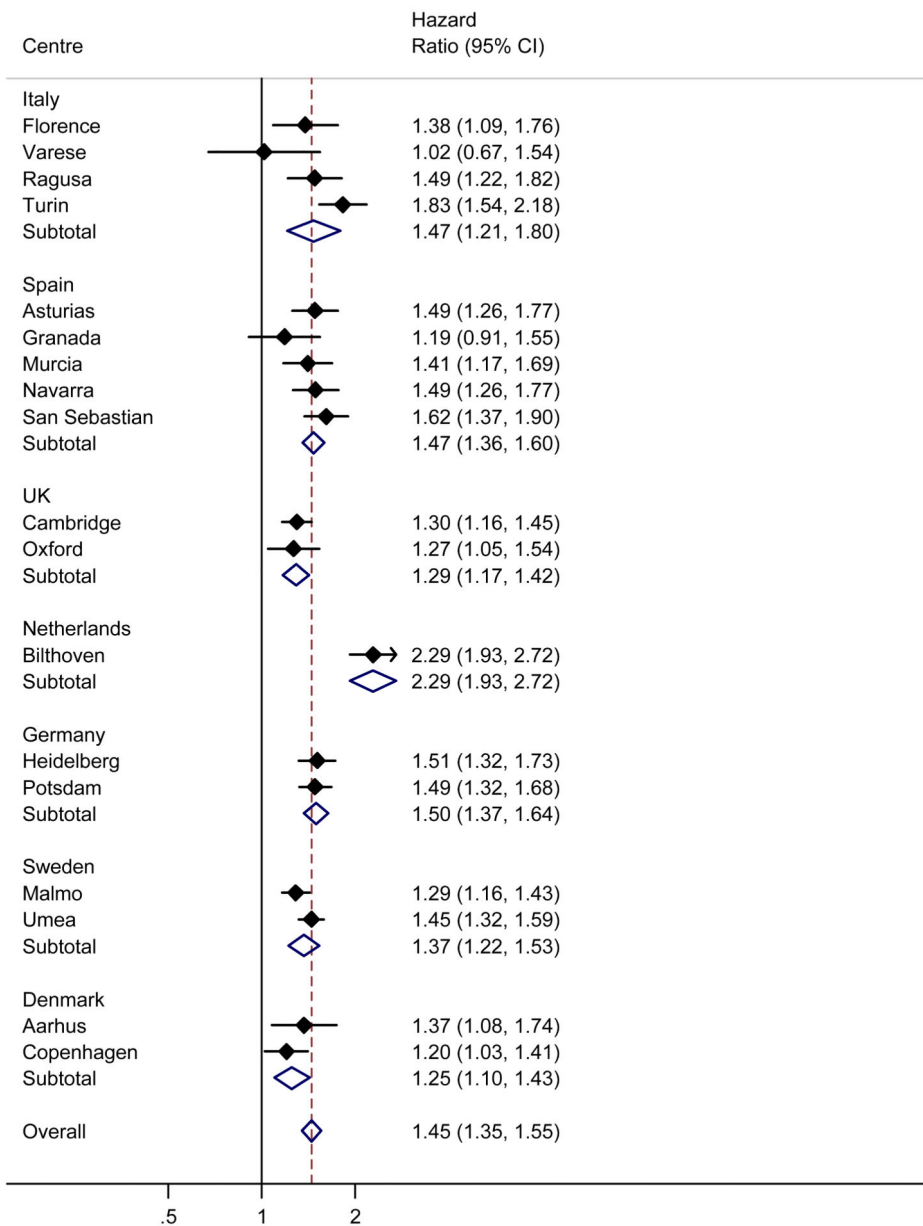


Figure 2. Hazard ratios for incident type 2 diabetes in men compared to women across InterAct centres and countries (I^2 57%). France, Naples and Utrecht included women only and are excluded.

a. Men



b. Women

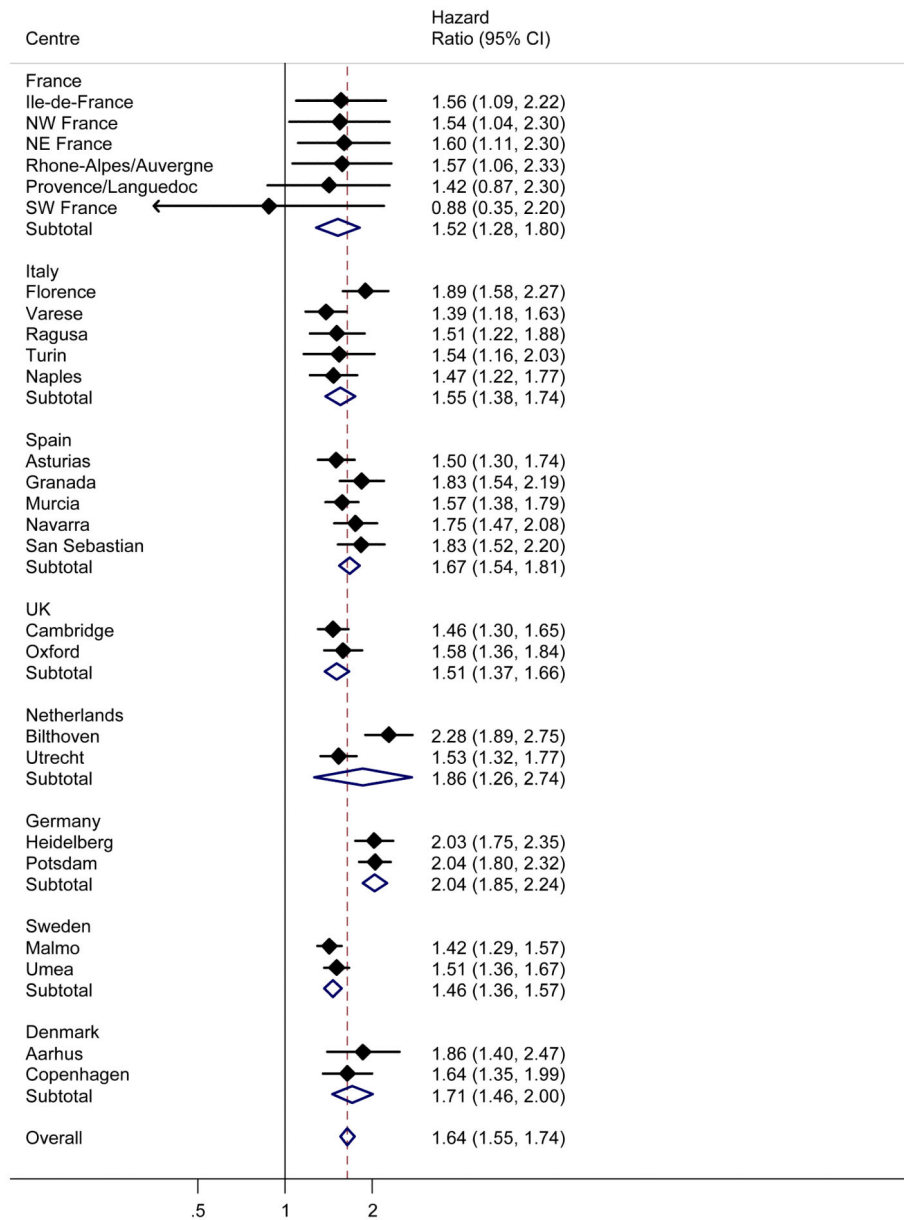


Figure 3. a-b. Hazard ratios for incident type 2 diabetes per 10 years of age across InterAct centres and countries (I^2 71% in men, 61% in women; analysis with calendar time as the underlying timescale).

Table 1

Overview of the EPIC cohorts contributing to the InterAct type 2 diabetes case-cohort study.

| Country | Centre | Baseline collection | | Stored samples | | |
|--------------|----------------------|---------------------|----------------|----------------|---------|--|
| | | Period | N | N | % Women | 5 th and 95 th Age Percentiles |
| France | Ile-de-France | 1993-1996 | 14,196 | 5,202 | 100 | 44-65 |
| | North-West France | 1993-1996 | 13,073 | 3,880 | 100 | 44-65 |
| | North-East France | 1993-1996 | 16,244 | 3,919 | 100 | 44-65 |
| | Rhone-Alpes/Auvergne | 1993-1996 | 9,966 | 3,319 | 100 | 44-65 |
| | Provence/Languedoc | 1993-1997 | 10,126 | 2,690 | 100 | 44-65 |
| | South-West France | 1993-1996 | 10,919 | 2,076 | 100 | 44-65 |
| Italy | Florence | 1992-1998 | 13,597 | 13,597 | 74 | 37-63 |
| | Varese | 1993-1997 | 12,083 | 12,073 | 79 | 39-64 |
| | Ragusa | 1993-1997 | 6,403 | 6,397 | 52 | 36-61 |
| | Turin | 1993-1998 | 10,604 | 10,604 | 43 | 37-62 |
| | Naples | 1993-1997 | 5,062 | 5,057 | 100 | 38-63 |
| Spain | Asturias | 1992-1995 | 8,542 | 8,422 | 64 | 36-62 |
| | Granada | 1992-1996 | 7,879 | 6,892 | 77 | 36-64 |
| | Murcia | 1992-1996 | 8,516 | 8,146 | 68 | 36-62 |
| | Navarra | 1992-1995 | 8,084 | 8,030 | 52 | 37-62 |
| | San Sebastian | 1992-1995 | 8,417 | 8,338 | 51 | 37-62 |
| UK | Cambridge | 1993-1998 | 30,441 | 24,036 | 55 | 45-74 |
| | Oxford | 1993-1998 | 57,489 | 19,241 | 77 | 24-70 |
| Netherlands | Bilthoven | 1993-1997 | 22,715 | 19,388 | 55 | 23-58 |
| | Utrecht | 1993-1997 | 17,357 | 16,930 | 100 | 49-68 |
| Germany | Heidelberg | 1994-1998 | 25,540 | 24,236 | 53 | 37-63 |
| | Potsdam | 1994-1998 | 27,548 | 26,444 | 60 | 36-64 |
| Sweden | Malmö | 1991-1996 | 28,098 | 28,053 | 61 | 47-71 |
| | Umeå | 1992-1996 | 25,728 | 25,728 | 52 | 30-60 |
| Denmark | Aarhus | 1995-1997 | 17,154 | 17,094 | 51 | 50-64 |
| | Copenhagen | 1993-1997 | 39,899 | 39,036 | 53 | 50-64 |
| Total | | | 455,680 | 348,828 | | |

Table 2

Characteristics (mean (standard deviation) for continuous and % (N) for categorical variables) of 12,403 InterAct incident type 2 diabetes cases.

| | Total Population | | | Men | | | Women | | |
|--|------------------|---------|--------|-----------|---------|--------|-----------|---------|--------|
| | % missing | Mean/ % | SD / N | % missing | Mean/ % | SD / N | % missing | Mean/ % | SD / N |
| | N=12,403 | | | N=6,165 | | | N=6,238 | | |
| Age (yrs) | 0.0 | 55.6 | 7.7 | 0.0 | 55.4 | 7.5 | 0.0 | 55.8 | 7.8 |
| Age at diagnosis (yrs) ^a | 0.0 | 62.5 | 8.0 | 0.0 | 62.3 | 7.8 | 0.0 | 62.6 | 8.2 |
| BMI (kg/m ²) | 0.8 | 29.7 | 4.7 | 0.8 | 29.4 | 4.0 | 0.8 | 30.1 | 5.3 |
| Waist (cm) | 7.7 | 97.7 | 12.5 | 8.7 | 102.7 | 10.5 | 6.6 | 92.8 | 12.3 |
| Waist hip ratio (%) | 7.7 | 91.6 | 8.9 | 8.8 | 98.0 | 6.0 | 6.6 | 85.4 | 6.6 |
| Follow-up (yrs) | 0.0 | 6.9 | 3.3 | 0.0 | 6.9 | 3.3 | 0.0 | 6.8 | 3.3 |
| Family history of diabetes ^b | 1.5 | | | 1.9 | | | 1.1 | | |
| | <i>No</i> | 67.7 | 4,998 | | 71.8 | 2,641 | | 63.6 | 2,357 |
| | <i>Yes</i> | 30.8 | 2,277 | | 26.3 | 968 | | 35.3 | 1,309 |
| Self reported hypertension ^c | 3.1 | | | 4.0 | | | 2.3 | | |
| | <i>No</i> | 60.0 | 7,438 | | 61.1 | 3,764 | | 58.9 | 3,674 |
| | <i>Yes</i> | 36.9 | 4,577 | | 35.0 | 2,156 | | 38.8 | 2,421 |
| Self reported hyperlipidaemia ^c | 29.1 | | | 31.6 | | | 26.8 | | |
| | <i>No</i> | 50.2 | 6,231 | | 45.2 | 2,784 | | 55.3 | 3,447 |
| | <i>Yes</i> | 20.6 | 2,557 | | 23.3 | 1,435 | | 18.0 | 1,122 |
| Self reported diabetes ^d | 4.2 | | | 6.3 | | | 2.4 | | |
| | <i>No</i> | 33.9 | 2,620 | | 35.2 | 1,261 | | 32.8 | 1,359 |
| | <i>Yes</i> | 61.9 | 4,779 | | 58.4 | 2,090 | | 64.8 | 2,689 |
| Antidiabetic drug use ^d | 0.1 | | | 0.1 | | | 0.1 | | |
| | <i>No</i> | 47.7 | 3,682 | | 47.8 | 1,710 | | 47.5 | 1,972 |
| | <i>Yes</i> | 52.3 | 4,037 | | 52.1 | 1,865 | | 52.4 | 2,172 |
| Self reported diabetes & drug use ^d | 4.3 | | | 6.4 | | | 2.5 | | |
| | <i>No</i> | 63.3 | 4,887 | | 61.6 | 2,204 | | 64.7 | 2,683 |
| | <i>Yes</i> | 32.4 | 2,505 | | 32.0 | 1,144 | | 32.8 | 1,361 |

^aIn 421/12,403 (3.3%) cases (235 men, 186 women), date of diagnosis was missing; for these cases date of diagnosis was imputed to be halfway between date of recruitment and end of follow-up, and this imputed date of diagnosis was used to calculate age at diagnosis.

^bFamily history data were not collected in Italy, Spain, Heidelberg and Oxford (excluded from these summaries).

^cHypertension is "Yes" if either "treatment for hypertension" or "self-reported hypertension" is positive; missing if both variables are missing (also hyperlipidaemia).

^dInformation on self reported diabetes/drug use was not available in Sweden and Denmark (excluded from these summaries).