

Dairy product intake and risk of type 2 diabetes in EPIC-InterAct: A Mendelian randomization study.

Dairy intake and type 2 diabetes: An MR study.

Linda E.T. Vissers¹, Iivonne Sluijs¹, Yvonne T. van der Schouw¹, Nita G. Forouhi², Fumiaki Imamura², Stephen Burgess³, Aurelio Barricarte⁴, Heiner Boeing⁵, Catalina Bonet⁶, Maria-Dolores Chirlaque^{7,8}, Guy Fagherazzi⁹, Paul W. Franks¹⁰, Heinz Freisling¹¹, Marc J. Gunter¹¹, J. Ramón Quirós¹², Daniel B. Ibsen¹³, Rudolf Kaaks¹⁴, Timothy Key¹⁵, Kay T. Khaw³, Tilman Kühn¹⁴, Olatz Mokoroa^{8,16}, Peter M. Nilsson¹⁰, Kim Overvad¹⁷, Valeria Pala¹⁸, Domenico Palli¹⁹, Salvatore Panico²⁰, Carlotta Sacerdote²¹, Annemieke MW Spijkerman²², Anne Tjønneland²³, Rosario Tumino²⁴, Miguel Rodríguez-Barranco^{8,25}, Olov Rolandsson²⁶, Elio Riboli²⁷, Stephen J. Sharp², Claudia Langenberg², Nicholas J. Wareham².

1. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands
2. MRC Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom
3. Department of Public Health and Primary Care, School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom
4. Navarre Public Health Institute (ISPN), Pamplona, Spain
5. German Institute of Human Nutrition Potsdam-Rehbruecke, Germany
6. Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Institut Català d'Oncologia
7. Department of Epidemiology, Regional Health Council, IMIB-Arrixaca, Murcia University, Spain.
8. CIBER in Epidemiology and Public Health (CIBERESP), Madrid, Spain
9. CESP UMR1018, Inserm, Institut Gustave Roussy, Paris South- Paris Saclay University, Villejuif, France
10. Lund University, Malmö, Sweden
11. Section of Nutrition and Metabolism, International Agency for Research on Cancer, Lyon, France.
12. Public Health Directorate, Asturias, Spain
13. Section for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark.

14. German Cancer Research Center (DKFZ), Heidelberg, Germany
15. University of Oxford, United Kingdom
16. Public Health Division of Gipuzkoa, Biodonostia Research Institute, San Sebastian, Spain.
17. Department of Cardiology, Aalborg University Hospital, Aalborg, Denmark.
18. Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan
19. Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network - ISPRO, Florence- ITALY
20. Dipartimento di medicina clinica e chirurgia, federicio II University, Naples, Italy
21. Unit of Cancer Epidemiology, Citta' della Salute e della Scienza Hospital-University of Turin and Center for Cancer Prevention (CPO), Torino, Italy
22. National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
23. Danish Cancer Society Research Center, Copenhagen, Denmark
24. ASP Ragusa, Italy
25. Escuela Andaluza de Salud Pública. Instituto de Investigación Biosanitaria ibs.GRANADA. Hospitales Universitarios de Granada/Universidad de Granada, Granada, Spain
26. Umeå University, Umeå, Sweden
27. School of Public Health, Imperial College London, UK

Corresponding author Ivonne Sluijs, PhD

Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht

STRT 6.131, PO Box 85500, 3508GA Utrecht, The Netherlands

T: 0031 88 7568181

E: I.Sluijs-2@umcutrecht.nl

Word count: 4,000

Number of tables: 13

Number of figures: 2

Abstract

Objective – Estimate the causal association between intake of dairy products and incident type 2 diabetes.

Research design and methods – The analysis included 21,820 European individuals (9,686 diabetes cases) of the EPIC-InterAct case-cohort study. Participants were genotyped and rs4988235 (LCT-12910C>T), a SNP for lactase persistence (LP) which enables digestion of dairy sugar, i.e. lactose, was imputed. Baseline dietary intakes were assessed with diet questionnaires. We investigated the associations between imputed SNP dosage for rs4988235 and intake of dairy products and other foods through linear regression. Mendelian Randomization (MR) estimates for the milk-diabetes relationship were obtained through a two stage least squares regression.

Results – Each additional LP allele was associated with a higher intake of milk (β 17.1 g/day, 95%CI 10.6,23.6) and milk beverages (β 2.8 g/day, 95%CI 1.0,4.5), but not with intake of other dairy products. Other dietary intakes associated with rs4988235 included fruits (β -7.0 g/d, 95%CI -12.4,-1.7 per additional LP allele), non-alcoholic beverages (β -18.0 g/d, 95%CI -34.4,-1.6) and wine (β -4.8 g/d, 95%CI -9.1,-0.6). In IV analysis, LP associated milk intake was not associated with diabetes (HR 0.99_{per 15 g/day} (95%CI: 0.93,1.05).

Conclusions – rs4988235 was associated with milk intake, but not with intake of other dairy products. This MR study does not suggest that milk intake is associated with diabetes, which is consistent with previous observational and genetic associations. LP may be associated with intake of other foods as well, but due to the modest associations we consider it unlikely that this has caused the observed null result.

Introduction

Eating healthily on a daily basis is a major step to prevent development of type 2 diabetes (1). Higher intake of dairy products has been associated with a lower risk of diabetes in a meta-analysis of observational studies (2). Particularly yoghurt and cheese intake were associated with lower diabetes risk, whereas milk intake was not, with substantial heterogeneity for most dairy products. In European populations, a direct association between milk intake and risk of diabetes was suggested, but this was not statistically significant (2). Protective components of dairy products may be whey-proteins, odd-chain fatty acids, and the high nutrient density of dairy (3). Also, interactions within the dairy food matrix may modify the metabolic effects of dairy consumption (4).

However, potential confounding and reverse causation cannot be excluded (5). Due to these limitations, the causal role of dairy products in diabetes prevention remains debatable.

The relationship between dairy products and risk of diabetes could be investigated by applying a Mendelian Randomization (MR) approach (5), using genetic variability in the *MCM6* gene associated to lactase persistence (LP) in adults as an instrumental variable (IV). Lactase is necessary to break down the sugars that are found in dairy products, i.e. lactose. SNPs in the *MCM6* region have been associated to LP (6). rs4988235 (LCT-12910C>T) has been associated to LP in European populations (6; 7), and has been associated to a higher intake of milk in European cohorts (8-12), albeit not in all (13). Previous MR studies reported no association between LP associated milk intake and diabetes (8; 14). However, variation in the *MCM6* gene is likely to lead to population

stratification (6; 10), which would introduce bias to an MR analysis, and previous MR studies (8; 14) did not sufficiently adjust for population substructure. Also, previous studies did not investigate whether rs4988235 was specifically associated with dairy product intake after adjusting for population substructure.

We therefore investigated whether rs4988235 associated with intake of dairy products and other foods in a pan-European study in 8 countries with different dietary habits (15). We adjusted for genetic principal components and study centre to adjust for population substructure (16). Next, we used rs4988235 in an IV analysis to investigate if there is a causal relationship between the LP associated exposure and risk of diabetes.

Research Design and Methods

Study design and population

EPIC-InterAct is a prospective case-cohort study nested within eight European countries of the European Prospective Investigation of Cancer and nutrition (EPIC) study (17). From 340,234 adults of the EPIC study for whom baseline blood samples were available, EPIC-InterAct randomly selected a subcohort of 16,154 participants and identified 12,403 incident cases of type 2 diabetes between 1991 and 2007, including 778 cases from the subcohort by design.

We excluded 5,287 participants that were not successfully genotyped (14% of samples had no DNA, 58% had insufficient DNA, 8% failed initial PCR and 20% failed array-based genotyping). Of the 5,287 participants that failed genotyping, 57% developed diabetes during follow-up, versus 55% of 22,492 participants that were successfully genotyped.

We additionally excluded participants with missing information on dairy product intake (n=592) and one participant per set of relatives (identity by descent > 0.1875, n=80), leading to a total population of 21,820 (Supplemental Figure 1).

Genotyping and GWAS search

DNA was extracted from buffy coat from a citrated blood sample on an automated Autopure LS DNA extraction system (Qiagen, Hilden, Germany) with PUREGENE chemistry (Qiagen). Participant samples were genotyped with the Illumina HumanCore Exome Chip arrays 12v1 or 24v1 (n=12,792) or with the Illumina 660W QuadBeadChip (n=8,955). Sample exclusion criteria were low call rate (threshold <95.4% in Illumina

660, <98% in core exome arrays), discordance between self-reported sex and sex based on X chromosome heterozygosity, outliers for heterozygosity, lack of concordance with previous genotyping results or non-European genetic ancestry.

Before imputation, SNPs were filtered to remove those with minor allele count <2, call rate <95% or Hardy Weinberg p-value <1e-6. Imputation to the Haplotype Reference Consortium v1.0 panel using IMPUTE v2.3.2 was performed at the Wellcome Trust Centre for Human Genetics.

For the present analyses, we used the imputed LP SNP rs4988235 (LCT-12910C>T) (7). Imputation info, a measure that reflects certainty of imputation (18), was 0.89 on the Illumina 660W quad chip and 0.45 on the Illumina HumanCore Exome chip for rs4988235.

Phenoscaner (19) was used to search for phenotypes that have been associated with rs4988235 or its proxies ($R^2 > 0.8$) in previous GWAS. These phenotypes could be mediators of an association between milk intake and diabetes, but could also indicate pleiotropy. Pleiotropy occurs when genetic variation affects two or more phenotypic traits that are seemingly unrelated and can lead to violation of assumptions made in MR studies (5). The effect allele (T) of rs4988235 has been associated (at $p < 5 \times 10^{-8}$) with a larger hip circumference (20), taller height (21) and lower LDL and total cholesterol (22) (Supplemental Table 1). We did not find evidence of an association between rs4988235 and phenotypes that are unrelated to dairy product intake.

Dietary measurement

Dietary intake over the previous twelve months before study inclusion was assessed at baseline through diet questionnaires, which varied per country or study centre. These questionnaires were developed to reflect local eating habits and validated locally (23-25).

Information on intake of milk ((semi-)skimmed or full-fat, regardless of fermentation), yoghurt and thick fermented milk (e.g. sour milk) and cheese was available for the full subcohort (n=12,722). Availability of consumption data for other dairy products differed by country and/or centre, depending on the cohort-specific questionnaire. Information on intake of dairy creams (e.g. whipped cream), curd (e.g. quark, cottage cheese), milk based puddings (e.g. custard), milk beverages (e.g. chocolate milk), and milk for coffee and creamers was available for 11,536, 10,372, 10,372, 6,867 and 2,959 subcohort participants, respectively. Information on consumption of non-milk dairy was calculated by summing intake of all dairy products other than milk.

Measurement of fatty acids has been described in detail previously (26). In short, fatty acids C15:0 and C17:0 were measured in phospholipids in plasma samples that were stored at baseline at -196°C or -150°C , using gas chromatography (Agilent Technologies, CA, USA) equipped with flame ionisation detection .

Covariates

Baseline information on education, lifestyle and medical history was obtained from self-administered questionnaires. Weight and height and hip and waist circumference were measured during a visit to a study centre. Body mass index, (BMI , weight (kg) divided

by height squared (m^2)) and waist-hip ratio (WHR, waist circumference/hip circumference) were calculated, adjusted for clothing by subtracting 1.5 kg for weight and 2.0 cm for circumferences for people that were normally dressed without shoes. Physical activity was classified as inactive, moderately inactive, moderately active, and active, according to the Cambridge Physical Activity Index (27). The lipids HDL-cholesterol, triglycerides, and lipoprotein (a) were measured in serum using a Cobas enzymatic assay (Roche Diagnostics, Mannheim, Germany) on a Roche Hitachi Modular P analyser. LDL-cholesterol was calculated by Friedewald equation. Erythrocyte HbA1c was measured using Tosoh (HLC-723G8) ion exchange high-performance liquid chromatography on a Tosoh G8.

Diabetes

Ascertainment of incident type 2 diabetes has been described previously (17) and involved a review of the existing EPIC datasets at each centre using multiple sources of evidence, including self-report, linkage to primary- and secondary-care registers, drug registers, hospital admissions and mortality data. To increase specificity of case definition, confirmation of type 2 diabetes diagnosis was sought where there was only one source of evidence, including individual medical records review in some centres. Follow-up was censored at date of diagnosis, 31 December 2007 or the date of death, whichever occurred first.

Data analysis

Baseline characteristics and dietary intakes of the subcohort and by most probable rs4988235 genotype were presented as percentages, mean \pm standard deviation (SD) or median (p25-p75). We additionally reported dietary intakes per country. Hardy Weinberg equilibrium (HWE) based on most probable LP genotype was examined in the subcohort, and after stratification to country.

We first checked two IV assumptions, namely if the IV is reliably associated with exposure (rs4988235 and intake of dairy products), and if the IV is independent of confounders of the exposure-outcome relationship (rs4988235 and intake of other foods, and diabetes risk factors). The third IV assumption is that the IV solely influences the outcome via a causal pathway that includes the exposure of interest, and cannot be checked (5). IV assumptions were checked only in subcohort participants, because the subcohort is population-based. We then proceeded to the IV analysis in the full case-cohort to answer the main research question: Is there a causal relationship between dairy product intake and risk of diabetes? (rs4988235 and risk of diabetes and MR analysis)

rs4988235 and intake of dairy products

rs4988235 was modelled continuously with SNP dosage, assuming an additive effect of LP alleles. Linear regression was used to examine the association between LP and dairy products. If rs4988235 was associated with multiple dairy products, the association between the LP SNP and a composite dairy product exposure (e.g. milk and milk beverages) was examined. The model with the highest F statistic was used for

subsequent IV analyses. The F statistic is often used to assess the risk of weak instrument bias in MR studies, but simulation studies have shown that bias when using one IV is very low, even for expected F statistics around 5 (28).

To support results based on diet questionnaire data, we also performed a linear regression between rs4988235 and plasma levels of saturated fats C15:0 and C17:0. These fatty acids are mainly found in dairy products, and plasma measurements of C15:0 and C17:0 have previously been associated to dairy product intake, although there seems to be some endogenous production of these fatty acids as well (29). We adjusted for the first three genetic principal components (PC), study centre, genotyping platform, sex and age in the linear regression models.

rs4988235 and diabetes risk factors

To examine potential pleiotropy of rs4988235 in EPIC-InterAct, we investigated the relationship between rs4988235 and baseline lifestyle factors, anthropometric measurements and blood lipids. These phenotypes could either be intermediates in the causal chain of an association between milk intake and diabetes, or pleiotropic effects of the LP SNP.

Linear regression was used for continuous variables and logistic regression for dichotomous variables. Triglycerides and lipoprotein a were log transformed before analysis. Analyses were adjusted for genotyping platform, sex, age, first three genetic PC and study centre.

The avoidance or consumption of dairy products may have consequences for the whole dietary pattern. More specifically, if a person avoids dairy products, this could lead to a

reduction in total energy intake, or replacement of dairy by other beverages or foods, or a combination of the two scenarios. It is also possible that dairy products are consumed together with other beverages or foods.

We therefore investigated the relationship between rs4988235 and dietary intake of whole foods other than dairy, using linear regression in an identical manner as for dairy products. We did not correct for multiple testing in the analysis of these closely related dietary exposures.

rs4988235 and risk of diabetes and MR analysis

First, we investigated if rs4988235 was related to HbA1c levels in the subcohort through linear regression, adjusting for genotyping platform, sex, age, first three genetic PC and study centre. The relationship between rs4988235 and risk of incident diabetes was examined per country through a Prentice-weighted Cox regression (30), using age as underlying time scale, with adjustment for genotyping platform, sex, age, the first three genetic PC and study centre. Country-specific results were pooled with inverse variance weights in a random effect meta-analysis using restricted maximum likelihood estimation.

We then performed the MR analysis, a two stage least squares (2SLS) IV analysis. LP associated dairy product intake for each participant was calculated by using rs4988235, genotyping platform, sex, age, the first three genetic PC and study centre as predictors in a linear regression model. Next, we investigated the association between predicted dairy product intake and diabetes in a Prentice-weighted Cox regression per country and

subsequently pooled results. We investigated associations per 15 gram of genetically predicted dairy intake to obtain interpretable 95% confidence intervals.

Because a regular 2SLS IV analysis does not take variance of the gene-exposure association into account and this could influence our findings (31), we additionally performed the MR analysis among 10,000 bootstrap samples and obtained a 95% confidence interval through the percentile method. Also, we repeated MR analysis under the assumption of dominance of LP, after exclusion of non-cases with an HbA1c > 6.5% (48 mmol/mol) at baseline, in participants from countries where rs4988235 was in HWE, and in participants genotyped on the Illumina 660W quad chip. All analysis were performed in R (32) version 3.4.1.

Results

The subcohort consisted of 12,722 participants with an average age of 52 (SD 9) years and 61.6% was female. In EPIC-InterAct, median milk intake was 162 (p25-p75 27-300) g/day and non-consumption of milk was rare (7.1%) (Table 1). Milk intake differed between countries, ranging from 15 (0-150) g/day in France to 200 (107-307) g/day in the United Kingdom. Non-consumption of milk ranged from 47.6% in France to 0% in Denmark (Supplemental Table 2-5).

Prevalence of homozygous LP (rs4988235 T/T genotype) differed within Europe with a range from 7.4% in Italy to 53.9% in Sweden (Supplemental Table 6). LP SNP genotypes were in HWE in the total subcohort, but deviation from HWE at a $p < 0.05$ significance level was observed in Italy, Spain, the United Kingdom, Germany and Denmark (Supplemental Table 6). Baseline characteristics of the subcohort by rs4988235 genotype showed that with increasing number of LP (T) alleles, participants were older, had a higher systolic blood pressure, were more physically active and highly educated (Supplemental Table 7). Also consumption of dairy products, potatoes, margarine, sugar, non-alcoholic beverages and coffee was higher, whereas intake of pasta/rice, cereal products, fruit, vegetables, legumes and vegetable oils was lower (Supplemental Table 8).

rs4988235 and intake of dairy products and other foods

After adjustment for sex, age, principal components, study centre and genotyping platform, one additional LP allele was associated with a higher intake of milk (β 17.1 g/day, 95%CI 10.6 to 23.6) and milk beverages (β 2.8 g/day, 95%CI 1.0 to 4.5), but not

with intake of other dairy products (Table 2). F statistic of the model predicting milk intake was 74.0, and predicting a composite endpoint of milk and milk beverages intake decreased the F statistic to 64.5.

rs4988235 was associated with higher plasma C15:0 (β 0.003 % of total fats, 95%CI 0.0005 to 0.005). A similar association was suggested for C17:0 levels (β 0.002 % of total fats, 95% CI -0.0007 to 0.005), but this was not statistically significant.

rs4988235 and diabetes risk factors

rs4988235 was associated with a modestly lower HDL cholesterol (β -0.017, 95%CI -0.029 to -0.004). After multivariable adjustment, dietary intakes associated with rs4988235 were potatoes (β 3.0, 95%CI 0.9 to 5.1), cereal (β -3.5, 95% CI -6.7 to -0.3), fruits (β -7.0, 95%CI -12.4 to -1.7), poultry (β -0.8, 95%CI -1.4 to -0.2), non-alcoholic beverages (β -18.0, 95%CI -34.4 to -1.6) and wine (β -4.8, 95%CI -9.1 to -0.6) (Table 2). rs4988235 was not associated with energy intake.

rs4988235 and risk of diabetes and MR analysis

rs4988235 was not associated with baseline HbA1c levels (β -0.07 %, 95%CI -0.22 to 0.07, n=12,519), nor with risk of developing diabetes (HR_{per additional LP allele} 0.99, 95%CI 0.94,1.04). MR analysis suggested no association between milk intake and diabetes (HR 0.99_{per 15 g/day}, 95%CI 0.93 to 1.05, $I^2 = 44\%$) (Table 4, Supplemental Figure 2).

Results did not differ when repeating the IV analysis among 10,000 bootstrap samples (HR_{per 15 g/day} 0.99, 95%CI 0.94 to 1.04) and assuming a dominant effect of LP did not alter conclusions (HR_{per 15 g/day} 0.99, 95%CI 0.92-1.07). Restricting analysis to

participants genotyped on the Illumina Human Quad 660 chip did not change conclusions, nor did excluding 73 participants with an HbA1c > 6.5% (48 mmol/mol), or limiting analysis to countries in HWE (Supplemental Table 9).

Conclusions

In this European case-cohort study including 9,686 incident type 2 diabetes cases, we did not find evidence for a causal relationship between milk intake and diabetes.

The LP SNP rs4988235 was associated with milk intake, but not with intake of other dairy products after multivariable adjustment. rs4988235 was associated with a slightly lower HDL-cholesterol, a higher intake of potatoes and a lower intake of cereal, fruit, poultry, wine and non-alcoholic beverages.

The null association we observed between milk intake and diabetes is in line with observational studies (2), including from EPIC-InterAct (33), and studies investigating the association between variation in the *MCM6* gene and diabetes (8; 12; 14; 34; 35).

Our results are consistent with findings from an MR study in a Danish cohort ($OR_{\text{per } 250 \text{ g/week}} 0.99$, 95%CI 0.93-1.06; assuming LP dominant effect; 1,355 cases) (8), and a two-sample MR study using endpoint data from DIAGRAM ($OR_{\text{per } 66 \text{ gram/day}} 0.92$, 95%CI 0.83, 1.02; assuming LP additive effect; 26,488 cases) (14).

These findings combined support the notion that the inverse association between milk intake and diabetes as seen in some studies (2) is likely caused by residual confounding.

Strengths of this study include the large population-based case-cohort from different regions within Europe with a broad range of LP prevalence. We confirmed the gene-exposure (rs4988235-milk) relationship with odd chain fatty acids measurements, and

investigated the association between the LP SNP and a wide range of dietary intakes.

We showed robustness of our findings in sensitivity analyses.

Despite its strengths, there are study limitations worth noting. First, 19% of EPIC-InterAct participants was excluded from our analyses due to missing genetic information. However, baseline characteristics (17) and the proportion of diabetes cases did not differ substantially between participants with and without genotype information, so it is unlikely that this has led to selection bias.

A second limitation is that intake of dairy products was determined with a diet questionnaire, which would mainly lead to nondifferential measurement error for this analysis. However, local validation studies have shown a reasonable to high relative validity of the diet questionnaire when compared to 24h recalls for milk and milk products, and a fair to reasonable for cheese (36; 37).

Also, we relied on imputed data to ascertain a SNP loading for rs4988235 with suboptimal imputation info scores (18) for participants genotyped to the Illumina HumanCore Exome chip, which might lead to lower IV strength due to misclassification. However, restricting analysis to participants genotyped to the Illumina 660W quad chip did not change results. We also observed deviation from HWE ($p < 0.05$) in most countries, while there was no deviation from HWE in the total cohort. Testing for deviation from HWE at $p < 0.05$ has been proposed to assess whether ascertainment bias may be present (38), which would occur if LP-associated characteristics determine study inclusion. This form of selection bias is unlikely since EPIC-InterAct is comprised of population-based cohorts (17) and we did not find evidence of an association between rs4988235 and early death in our GWAS search (Supplemental Table 1). Also,

deviation from HWE may be explained by genetic mixture of subpopulations with different allele frequencies (38), as is the case for rs4988235 in EPIC-InterAct. If this is the case, appropriate correction for population substructure prevent bias. In EPIC-InterAct, participant characteristics and dietary intake by rs4988235 genotype differed substantially (Supplemental Table 7 and 8), but many of these associations were not found after adjustments for study centre and genetic principal components (Table 2 and 3). Also, a sensitivity analysis in countries without deviation from HWE did not alter conclusions.

Due to aforementioned limitations of this study, we cannot exclude the possibility that a small effect of milk intake on diabetes risk is present, despite our null finding.

In EPIC-InterAct, lactase persistence due to rs4988235 was associated with a higher milk intake, and lower intake of other non-alcoholic beverages. We also observed modest associations between the LP SNP and higher intake of potatoes and a lower intake of cereal, fruit, poultry and wine. We found no association between rs4988235 and total energy intake. Our results suggest that LP may be associated with a composite of nutritional factors, rather than only with milk intake, although the association of the LP SNP with milk is much stronger than with other dietary products. These associations should be interpreted cautiously as they may be due to chance since we did not correct for multiple testing, and we cannot exclude presence of residual population stratification. Repeating these analyses in other cohorts is required before drawing firm conclusions on a potential LP associated dietary pattern.

We also observed that all Danish participants consumed milk, including those that were genetically unable to produce lactase, whereas approximately 40% of the French lactase persistent population did not consume milk. An explanation for this could be that there are cultural habits that lead to milk consumption in lactase non persistent people or milk avoidance in lactase persistent people, as hypothesized previously (8; 39). Given the known heterogeneity of gene-exposure association for LP (8-10; 12), we propose, in line with others (39), to demonstrate the association between an LP SNP and milk intake before using this IV in an MR study. To attribute the association of an LP SNP with disease to the correct exposure, one should know the association between the LP SNP and possible replacements for milk, total energy intake and other (dairy) foods in the population where gene-outcome data was derived from. Otherwise, causal inference on the exposure level may lead to an incorrect conclusion (40). Providing additional information on the LP associated exposure also facilitates comparison between studies. In our GWAS search, we found that each additional effect (T) allele of rs4988235 was associated to lower LDL and total cholesterol (22), which we did not observe in EPIC-InterAct. We cannot exclude that this is due to lower power in our analysis.

In EPIC-InterAct, rs4988235 was associated with intake of milk and milk beverages, but not with intake of other dairy products. The MR analysis provided no evidence for a causal relationship between milk intake and type 2 diabetes, which is in line with previous genetic and observational studies. We consider it unlikely that modest associations between LP and other dietary intakes have caused this null result. No

conclusion can be drawn regarding causality of the relationship of dairy products other than milk with diabetes.

Acknowledgements

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 InterAct Project. 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: NGF, FI, CL and NJW: MRC Epidemiology Unit core support (MC_UU_12015/5 and MC_UU_12015/1); NGF and NJW: National Institute for Health Research Cambridge Biomedical Research Centre [IS-BRC-1215-20014]; IS and YTvdS: Verification of diabetes cases was additionally funded by NL Agency grant IGE05012 and an Incentive Grant from the Board of the UMC Utrecht (The Netherlands); AMWS and DLvdA: 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); FLC: Cancer Research UK; PWF: Swedish Research Council, Novo nordisk, Swedish Heart Lung Foundation, Swedish Diabetes Association; JH, KO and AT: Danish Cancer Society; RK: Deutsche Krebshilfe; SP: Associazione Italiana per la Ricerca sul Cancro; JRQ: Asturias Regional Government; MT: Health Research Fund (FIS) of the Spanish Ministry of Health; the CIBER en Epidemiología y Salud Pública (CIBERESP), Spain; Murcia Regional Government (Nº 6236); RT: AIRE-ONLUS Ragusa, AVIS-Ragusa, Sicilian Regional Government. TK: UK Medical Research Council MR/M012190/1 and Wellcome Trust Our Planet Our Health (Livestock, Environment and People, LEAP 205212/Z/16/Z)

We thank staff from the Technical, Field Epidemiology and Data Functional Group Teams of the MRC Epidemiology Unit in Cambridge, UK, for carrying out sample preparation, DNA provision and quality control, genotyping and data-handling work. We specifically thank S. Dawson for coordinating the sample provision for biomarker measurements, A. Britten for coordinating DNA sample provision and genotyping of candidate markers, N. Kerrison, C. Gillson and A. Britten for data provision and genotyping quality control and M. Sims for writing the technical laboratory specification for the intermediate pathway biomarker measurements and for overseeing the laboratory work.

Author contributions were as follows: Linda E.T. Vissers, Ivonne Sluijs and Yvonne T. van der Schouw had access to all data for this study. Linda E.T. Vissers and Ivonne Sluijs take responsibility for the manuscript contents. Linda E.T. Vissers analysed the data and drafted the manuscript. All authors qualify for authorship according to *Diabetes Care* criteria. They have all 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.

None of the authors declared a conflict of interest.

Table 1. Baseline characteristics and dairy intake of the subcohort.

<i>Baseline characteristics</i>	<i>Subcohort*</i>	<i>Missing</i>
N (persons)	12,722	
Age	52 ± 9	
Male sex	38.4%	
Current smoker	26.2%	
Former smoker	26.7%	
Hypertension	18.4%	2.4%
Systolic blood pressure (mmHg)	133 ± 20	21.2%
Diastolic blood pressure (mmHg)	82 ± 11	21.2%
Total cholesterol (mmol/L)	5.9 ± 1.1	4.3%
HDL cholesterol (mmol/L)	1.5 ± 0.4	4.3%
LDL cholesterol (mmol/L)	3.8 ± 1.0	5.5%
Triglycerides (mmol/L)	1.1 (0.8 – 1.6)	4.3%
Lipoprotein a (mg/L)	384 (200 - 684)	5.6%
Physically inactive	22.6%	1.2%
BMI (kg/m ²)	26.0 ± 4.2	0.5%
Waist-hip ratio	0.85 ± 0.09	7.5%
HbA1c >6.5% (>48 mmol/mol)	1.4%	1.6%
Low level of education †	40.6%	1.6%
Premenopausal status ‡	41.6%	
History of myocardial infarction	1.4%	1.6%
History of stroke	0.9%	8.6%
<i>Dairy product intake</i>		
		<i>No intake</i>
Energy (kcal)	2056 (1675 - 2517)	
Milk (g/day)	162 (37 - 300)	7.1%
Milk beverages (g/day)	1 (0 - 7)	72.2%
Milk for coffee and creamers (g/day)	0 (0 - 14)	89.5%
Dairy creams (g/day)	1 (0 - 4)	36.2%
Milk based puddings (g/day)	3 (0 - 15)	54.0%
Curd (g/day)	0 (0 - 7)	61.3%
Yogurt, thick fermented milk (g/day)	28 (0 - 100)	25.9%
Ice cream (g/day)	3 (0 - 9)	23.9%
Cheese (g/day)	28 (15 - 50)	5.1%
<i>Fatty acid measurements</i>		
		<i>Missing</i>
C15:0 (mol%)	0.21 ± 0.07	0.8%
C17:0 (mol%)	0.40 ± 0.09	0.8%

* Data are expressed as mean ± standard deviation, median (p25 - p75) or percentage of participants with available data for variable. † no education or only primary school education. ‡ Percentage among women

Table 2. Association after multivariable adjustment between rs4988235 and dietary intake among 12,722 subcohort participants.

	β^*	95% CI		p value	n
Energy (kcal/day)	5.2	-12.6	23.0	0.57	12,722
Milk (g/day)	17.1	10.6	23.6	2×10^{-7}	12,722
Non-milk dairy (g/day)	3.8	0.5	7.1	0.02	12,722
Milk beverages (g/day)	2.8	1.0	4.5	2×10^{-3}	6,867
Milk for coffee (g/day)	0.10	-0.72	0.91	0.82	2,959
Dairy creams (g/day)	0.06	-0.13	0.25	0.53	11,536
Cream desserts (g/day)	0.12	-0.70	0.94	0.77	10,372
Curd (g/day)	-0.15	-0.80	0.51	0.66	10,372
Yoghurt (g/day)	2.2	-0.5	4.9	0.11	12,722
Ice cream (g/day)	-0.04	-0.41	0.33	0.82	12,722
Cheese (g/day)	-0.09	-1.05	0.87	0.85	12,722
Potatoes (g/day)	3.0	0.9	5.1	5×10^{-3}	12,722
Pasta/rice (g/day)	-0.2	-1.8	1.3	0.78	12,722
Bread (g/day)	-1.4	-3.7	0.9	0.22	12,722
Cereal (g/day)	-3.5	-6.7	-0.3	0.03	12,722
Vegetables (g/day)	-0.3	-3.6	3.0	0.86	12,722
Legumes (g/day)	-0.14	-0.74	0.46	0.65	12,722
Fruits (g/day)	-7.0	-12.4	-1.7	0.01	12,722
Nuts (g/day)	-0.05	-0.29	0.20	0.72	12,722
Red meat (g/day)	-0.01	-0.96	0.95	0.99	12,722
Poultry (g/day)	-0.8	-1.4	-0.2	6×10^{-3}	12,722
Processed meat (g/day)	0.01	-0.92	0.93	0.99	12,722
Vegetable oils (g/day)	0.07	-0.17	0.31	0.57	12,722
Margarine (g/day)	0.3	-0.1	0.8	0.16	12,722
Butter (g/day)	0.13	-0.12	0.38	0.32	12,722
Sugar (g/day)	0.9	-0.7	2.6	0.25	12,722
Cake/cookies (g/day)	-1.3	-2.8	0.1	0.06	12,722
Beverages (g/day) †	-18.0	-34.3	-1.6	0.03	12,722
Softdrinks (g/day)	-0.3	-5.3	4.6	0.90	12,722
Juice (g/day)	-0.2	-3.6	3.3	0.93	12,722
Coffee (g/day)	4.4	-5.7	14.5	0.39	12,722
Tea (g/day)	-6.5	-14.2	1.2	0.10	12,722
Alcohol (g/day)	-0.3	-0.8	0.3	0.34	12,722
Wine (g/day)	-4.8	-9.1	-0.6	0.03	12,722

* β derived from linear regression model, adjusted for first three genetic principal components, study centre, genotyping platform, sex and age. † Sum of all non-alcoholic beverages, excluding water and milk.

Table 3. Association between rs4988235 and diabetes risk factors among 12,722 subcohort participants.

	Estimate*	95% CI		p value	n
BMI (kg/m ²)	0.03	-0.09	0.16	0.61	12,652
WHR	0.0017	-0.0003	0.0038	0.10	11,771
Systolic blood pressure (mmHg)	0.41	-0.23	1.06	0.21	10,027
Diastolic blood pressure (mmHg)	0.20	-0.18	0.58	0.30	10,026
Presence of hypertension	0.02	-0.06	0.10	0.63	12,677
Total cholesterol (mmol/L)	-0.014	-0.048	0.021	0.44	12,171
HDL cholesterol (mmol/L)	-0.017	-0.029	-0.004	0.01	12,173
LDL cholesterol (mmol/L)	0.001	-0.031	0.032	0.97	12,028
Triglycerides (mmol/L) †	0.009	-0.007	0.025	0.28	12,172
Lipoprotein a (mg/L) †	-0.01	-0.05	0.03	0.65	12,014
Current or former smoker	1.06	0.99	1.13	0.09	12,722
Physical inactivity	0.99	0.91	1.07	0.81	12,722
Low level of education	1.03	0.95	1.11	0.47	12,722
Peri- or postmenopausal status	1.00	0.85	1.17	0.97	7,841

* Estimate is odds ratio derived from logistic regression model (for smoking, physical inactivity, low level of education and post- or perimenopausal status) or β derived from linear regression (all other variables), adjusted for first three genetic principal components, study centre, genotyping platform, sex and age. † log transformed

Table 4. Association between genetic lactase persistence and diabetes (gene-outcome) and association between genetically predicted milk intake and diabetes among 21,820 EPIC-InterAct participants.

	95% confidence			
	HR*	interval		N
Gene-outcome	0.99	0.94	1.04	21,820
Genetically predicted milk intake (per 15g)	0.99	0.93	1.05	21,820

* HR for diabetes in gene-outcome is expressed per additional lactase persistence allele. HR for diabetes in IV analyses is expressed per 15 gram of genetically predicted milk intake. Analyses are performed per country, using age as underlying time scale and are adjusted for sex, genetic variability (first three principal components), study centre and genotyping platform. The reported overall estimates were obtained by pooling country-specific results in random-effects meta-analysis.

Literature Cited.

1. Franz MJ, Bantle JP, Beebe CA, Brunzell JD, Chiasson JL, Garg A, Holzmeister LA, Hoogwerf B, Mayer-Davis E, Mooradian AD, Purnell JQ, Wheeler M: Evidence-based nutrition principles and recommendations for the treatment and prevention of diabetes and related complications. *Diabetes care* 2002;25:148-198
2. Gijsbers L, Ding EL, Malik VS, de Goede J, Geleijnse JM, Soedamah-Muthu SS: Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies. *The American journal of clinical nutrition* 2016;103:1111-1124
3. Rice BH, Cifelli CJ, Pikosky MA, Miller GD: Dairy components and risk factors for cardiometabolic syndrome: recent evidence and opportunities for future research. *Advances in nutrition (Bethesda, Md)* 2011;2:396-407
4. Thorning TK, Bertram HC, Bonjour JP, de Groot L, Dupont D, Feeney E, Ipsen R, Lecerf JM, Mackie A, McKinley MC, Michalski MC, Remond D, Riserus U, Soedamah-Muthu SS, Tholstrup T, Weaver C, Astrup A, Givens I: Whole dairy matrix or single nutrients in assessment of health effects: current evidence and knowledge gaps. *The American journal of clinical nutrition* 2017;105:1033-1045
5. Davey Smith G, Hemani G: Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human molecular genetics* 2014;23:R89-98
6. Itan Y, Jones BL, Ingram CJ, Swallow DM, Thomas MG: A worldwide correlation of lactase persistence phenotype and genotypes. *BMC evolutionary biology* 2010;10:36
7. Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I: Identification of a variant associated with adult-type hypolactasia. *Nature genetics* 2002;30:233-237
8. Bergholdt HK, Nordestgaard BG, Ellervik C: Milk intake is not associated with low risk of diabetes or overweight-obesity: a Mendelian randomization study in 97,811 Danish individuals. *The American journal of clinical nutrition* 2015;102:487-496
9. Tornaiainen S, Hedelin M, Autio V, Rasinpera H, Balter KA, Klint A, Bellocco R, Wiklund F, Stattin P, Ikonen T, Tammela TL, Schleutker J, Gronberg H, Jarvela I: Lactase persistence, dietary intake of milk, and the risk for prostate cancer in Sweden and Finland. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2007;16:956-961
10. Smith GD, Lawlor DA, Timpson NJ, Baban J, Kiessling M, Day IN, Ebrahim S: Lactase persistence-related genetic variant: population substructure and health outcomes. *European journal of human genetics : EJHG* 2009;17:357-367
11. Travis RC, Appleby PN, Siddiq A, Allen NE, Kaaks R, Canzian F, Feller S, Tjonneland A, Fons Johnsen N, Overvad K, Ramon Quiros J, Gonzalez CA, Sanchez MJ, Larranaga N, Chirlaque MD, Barricarte A, Khaw KT, Wareham N, Trichopoulou A, Valanou E, Oustoglou E, Palli D, Sieri S, Tumino R, Sacerdote C, Bueno-de-Mesquita HB, Stattin P, Ferrari P, Johansson M, Norat T, Riboli E, Key TJ: Genetic variation in the lactase gene, dairy product intake and risk for prostate cancer in the European prospective investigation into cancer and nutrition. *International journal of cancer* 2013;132:1901-1910
12. Lamri A, Poli A, Emery N, Bellili N, Velho G, Lantieri O, Balkau B, Marre M, Fumeron F: The lactase persistence genotype is associated with body mass index and dairy consumption in the D.E.S.I.R. study. *Metabolism: clinical and experimental* 2013;62:1323-1329

13. Corella D, Arregui M, Coltell O, Portoles O, Guillem-Saiz P, Carrasco P, Sorli JV, Ortega-Azorin C, Gonzalez JI, Ordovas JM: Association of the LCT-13910C>T polymorphism with obesity and its modulation by dairy products in a Mediterranean population. *Obesity* (Silver Spring, Md) 2011;19:1707-1714
14. Yang Q, Lin SL, Au Yeung SL, Kwok MK, Xu L, Leung GM, Schooling CM: Genetically predicted milk consumption and bone health, ischemic heart disease and type 2 diabetes: a Mendelian randomization study. *European journal of clinical nutrition* 2017;
15. Hjartaker A, Lagiou A, Slimani N, Lund E, Chirlaque MD, Vasilopoulou E, Zavitsanos X, Berrino F, Sacerdote C, Ocke MC, Peeters PH, Engeset D, Skeie G, Aller A, Amiano P, Berglund G, Nilsson S, McTaggart A, Spencer EA, Overvad K, Tjonneland A, Clavel-Chapelon F, Linseisen J, Schulz M, Hemon B, Riboli E: Consumption of dairy products in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort: data from 35 955 24-hour dietary recalls in 10 European countries. *Public health nutrition* 2002;5:1259-1271
16. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* 2006;38:904-909
17. Langenberg C, Sharp SJ, Franks PW, Scott RA, Deloukas P, Forouhi NG, Froguel P, Groop LC, Hansen T, Palla L, Pedersen O, Schulze MB, Tormo MJ, Wheeler E, Agnoli C, Arriola L, Barricarte A, Boeing H, Clarke GM, Clavel-Chapelon F, Duell EJ, Fagherazzi G, Kaaks R, Kerrison ND, Key TJ, Khaw KT, Kroger J, Lajous M, Morris AP, Navarro C, Nilsson PM, Overvad K, Palli D, Panico S, Quiros JR, Rolandsson O, Sacerdote C, Sanchez MJ, Slimani N, Spijkerman AM, Tumino R, van der AD, van der Schouw YT, Barroso I, McCarthy MI, Riboli E, Wareham NJ: Gene-lifestyle interaction and type 2 diabetes: the EPIC interact case-cohort study. *PLoS medicine* 2014;11:e1001647
18. Marchini J, Howie B: Genotype imputation for genome-wide association studies. *Nature reviews Genetics* 2010;11:499-511
19. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, Paul DS, Freitag D, Burgess S, Danesh J, Young R, Butterworth AS: PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* (Oxford, England) 2016;32:3207-3209
20. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, Strawbridge RJ, Pers TH, Fischer K, Justice AE, Workalemahu T, Wu JMW, Buchkovich ML, Heard-Costa NL, Roman TS, Drong AW, Song C, Gustafsson S, Day FR, Esko T, Fall T, Kutalik Z, Luan J, Randall JC, Scherag A, Vedantam S, Wood AR, Chen J, Fehrmann R, Karjalainen J, Kahali B, Liu CT, Schmidt EM, Absher D, Amin N, Anderson D, Beekman M, Bragg-Gresham JL, Buyske S, Demirkan A, Ehret GB, Feitosa MF, Goel A, Jackson AU, Johnson T, Kleber ME, Kristiansson K, Mangino M, Leach IM, Medina-Gomez C, Palmer CD, Pasko D, Pechlivanis S, Peters MJ, Prokopenko I, Stancakova A, Sung YJ, Tanaka T, Teumer A, Van Vliet-Ostaptchouk JV, Yengo L, Zhang W, Albrecht E, Arnlov J, Arscott GM, Bandinelli S, Barrett A, Bellis C, Bennett AJ, Berne C, Bluher M, Bohringer S, Bonnet F, Bottcher Y, Bruinenberg M, Carba DB, Caspersen IH, Clarke R, Daw EW, Deelen J, Deelman E, Delgado G, Doney AS, Eklund N, Erdos MR, Estrada K, Eury E, Friedrich N, Garcia ME, Giedraitis V, Gigante B, Go AS, Golay A, Grallert H, Grammer TB, Grassler J, Grewal J, Groves CJ, Haller T, Hallmans G, Hartman CA, Hassinen M, Hayward C, Heikkila K, Herzig KH, Helmer Q, Hillege HL, Holmen O, Hunt SC, Isaacs A, Ittermann T, James AL, Johansson I, Juliusdottir T, Kalafati IP, Kinnunen L, Koenig W, Kooner IK, Kratzer W, Lamina C, Leander K, Lee NR, Lichtner P, Lind L,

Lindstrom J, Lobbens S, Lorentzon M, Mach F, Magnusson PK, Mahajan A, McArdle WL, Menni C, Merger S, Mihailov E, Milani L, Mills R, Moayyeri A, Monda KL, Mooijaart SP, Muhleisen TW, Mulas A, Muller G, Muller-Nurasyid M, Nagaraja R, Nalls MA, Narisu N, Glorioso N, Nolte IM, Olden M, Rayner NW, Renstrom F, Ried JS, Robertson NR, Rose LM, Sanna S, Scharnagl H, Scholtens S, Sennblad B, Seufferlein T, Sitlani CM, Smith AV, Stirrups K, Stringham HM, Sundstrom J, Swertz MA, Swift AJ, Syvanen AC, Tayo BO, Thorand B, Thorleifsson G, Tomaschitz A, Troffa C, van Oort FV, Verweij N, Vonk JM, Waite LL, Wennauer R, Wilsgaard T, Wojczynski MK, Wong A, Zhang Q, Zhao JH, Brennan EP, Choi M, Eriksson P, Folkersen L, Franco-Cereceda A, Gharavi AG, Hedman AK, Hivert MF, Huang J, Kanoni S, Karpe F, Keildson S, Kiryluk K, Liang L, Lifton RP, Ma B, McKnight AJ, McPherson R, Metspalu A, Min JL, Moffatt MF, Montgomery GW, Murabito JM, Nicholson G, Nyholt DR, Olsson C, Perry JR, Reinmaa E, Salem RM, Sandholm N, Schadt EE, Scott RA, Stolk L, Vallejo EE, Westra HJ, Zondervan KT, Consortium AD, Consortium CAD, Consortium CK, Consortium G, Consortium G, Glgc, Icbp, International Endogene C, LifeLines Cohort S, Investigators M, Mu TC, Consortium P, ReproGen C, Amouyel P, Arveiler D, Bakker SJ, Beilby J, Bergman RN, Blangero J, Brown MJ, Burnier M, Campbell H, Chakravarti A, Chines PS, Claudi-Boehm S, Collins FS, Crawford DC, Danesh J, de Faire U, de Geus EJ, Dorr M, Erbel R, Eriksson JG, Farrall M, Ferrannini E, Ferrieres J, Forouhi NG, Forrester T, Franco OH, Gansevoort RT, Gieger C, Gudnason V, Haiman CA, Harris TB, Hattersley AT, Heliovaara M, Hicks AA, Hingorani AD, Hoffmann W, Hofman A, Homuth G, Humphries SE, Hypponen E, Illig T, Jarvelin MR, Johansen B, Jousilahti P, Jula AM, Kaprio J, Kee F, Keinanen-Kiukaanniemi SM, Kooner JS, Kooperberg C, Kovacs P, Kraja AT, Kumari M, Kuulasmaa K, Kuusisto J, Lakka TA, Langenberg C, Le Marchand L, Lehtimäki T, Lyssenko V, Mannisto S, Marette A, Matise TC, McKenzie CA, McKnight B, Musk AW, Mohlenkamp S, Morris AD, Nelis M, Ohlsson C, Oldehinkel AJ, Ong KK, Palmer LJ, Penninx BW, Peters A, Pramstaller PP, Raitakari OT, Rankinen T, Rao DC, Rice TK, Ridker PM, Ritchie MD, Rudan I, Salomaa V, Samani NJ, Saramies J, Sarzynski MA, Schwarz PE, Shuldiner AR, Staessen JA, Steinthorsdottir V, Stolk RP, Strauch K, Tonjes A, Tremblay A, Tremoli E, Vohl MC, Volker U, Vollenweider P, Wilson JF, Witteman JC, Adair LS, Bochud M, Boehm BO, Bornstein SR, Bouchard C, Cauchi S, Caulfield MJ, Chambers JC, Chasman DI, Cooper RS, Dedoussis G, Ferrucci L, Froguel P, Grabe HJ, Hamsten A, Hui J, Hveem K, Jockel KH, Kivimäki M, Kuh D, Laakso M, Liu Y, Marz W, Munroe PB, Njolstad I, Oostra BA, Palmer CN, Pedersen NL, Perola M, Perusse L, Peters U, Power C, Quertermous T, Rauramaa R, Rivadeneira F, Saaristo TE, Saleheen D, Sinisalo J, Slagboom PE, Snieder H, Spector TD, Stefansson K, Stumvoll M, Tuomilehto J, Uitterlinden AG, Uusitupa M, van der Harst P, Veronesi G, Walker M, Wareham NJ, Watkins H, Wichmann HE, Abecasis GR, Assimes TL, Berndt SI, Boehnke M, Borecki IB, Deloukas P, Franke L, Frayling TM, Groop LC, Hunter DJ, Kaplan RC, O'Connell JR, Qi L, Schlessinger D, Strachan DP, Thorsteinsdottir U, van Duijn CM, Willer CJ, Visscher PM, Yang J, Hirschhorn JN, Zillikens MC, McCarthy MI, Speliotes EK, North KE, Fox CS, Barroso I, Franks PW, Ingelsson E, Heid IM, Loos RJ, Cupples LA, Morris AP, Lindgren CM, Mohlke KL: New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015;518:187-196

21. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z, Amin N, Buchkovich ML, Croteau-Chonka DC, Day FR, Duan Y, Fall T, Fehrmann R, Ferreira T, Jackson AU, Karjalainen J, Lo KS, Locke AE, Magi R, Mihailov E, Porcu E, Randall JC, Scherag A, Vinkhuyzen AA, Westra HJ, Winkler TW, Workalemahu T, Zhao JH, Absher D, Albrecht E, Anderson D, Baron J, Beekman M, Demirkan A, Ehret GB, Feenstra B, Feitosa MF, Fischer K,

Fraser RM, Goel A, Gong J, Justice AE, Kanoni S, Kleber ME, Kristiansson K, Lim U, Lotay V, Lui JC, Mangino M, Mateo Leach I, Medina-Gomez C, Nalls MA, Nyholt DR, Palmer CD, Pasko D, Pechlivanis S, Prokopenko I, Ried JS, Ripke S, Shungin D, Stancakova A, Strawbridge RJ, Sung YJ, Tanaka T, Teumer A, Trompet S, van der Laan SW, van Setten J, Van Vliet-Ostaptchouk JV, Wang Z, Yengo L, Zhang W, Afzal U, Arnlov J, Arscott GM, Bandinelli S, Barrett A, Bellis C, Bennett AJ, Berne C, Bluher M, Bolton JL, Bottcher Y, Boyd HA, Bruinenberg M, Buckley BM, Buyske S, Caspersen IH, Chines PS, Clarke R, Claudi-Boehm S, Cooper M, Daw EW, De Jong PA, Deelen J, Delgado G, Denny JC, Dhonukshe-Rutten R, Dimitriou M, Doney AS, Dorr M, Eklund N, Eury E, Folkersen L, Garcia ME, Geller F, Giedraitis V, Go AS, Grallert H, Grammer TB, Grassler J, Gronberg H, de Groot LC, Groves CJ, Haessler J, Hall P, Haller T, Hallmans G, Hannemann A, Hartman CA, Hassinen M, Hayward C, Heard-Costa NL, Helmer Q, Hemani G, Henders AK, Hillege HL, Hlatky MA, Hoffmann W, Hoffmann P, Holmen O, Houwing-Duistermaat JJ, Illig T, Isaacs A, James AL, Jeff J, Johansen B, Johansson A, Jolley J, Juliusdottir T, Junttila J, Kho AN, Kinnunen L, Klopp N, Kocher T, Kratzer W, Lichtner P, Lind L, Lindstrom J, Lobbens S, Lorentzon M, Lu Y, Lyssenko V, Magnusson PK, Mahajan A, Maillard M, McArdle WL, McKenzie CA, McLachlan S, McLaren PJ, Menni C, Merger S, Milani L, Moayyeri A, Monda KL, Morken MA, Muller G, Muller-Nurasyid M, Musk AW, Narisu N, Nauck M, Nolte IM, Nothen MM, Oozageer L, Pilz S, Rayner NW, Renstrom F, Robertson NR, Rose LM, Roussel R, Sanna S, Scharnagl H, Scholtens S, Schumacher FR, Schunkert H, Scott RA, Sehmi J, Seufferlein T, Shi J, Silventoinen K, Smit JH, Smith AV, Smolonska J, Stanton AV, Stirrups K, Stott DJ, Stringham HM, Sundstrom J, Swertz MA, Syvanen AC, Tayo BO, Thorleifsson G, Tyrer JP, van Dijk S, van Schoor NM, van der Velde N, van Heemst D, van Oort FV, Vermeulen SH, Verweij N, Vonk JM, Waite LL, Waldenberger M, Wennauer R, Wilkens LR, Willenborg C, Wilsgaard T, Wojczynski MK, Wong A, Wright AF, Zhang Q, Arveiler D, Bakker SJ, Beilby J, Bergman RN, Bergmann S, Biffar R, Blangero J, Boomsma DI, Bornstein SR, Bovet P, Brambilla P, Brown MJ, Campbell H, Caulfield MJ, Chakravarti A, Collins R, Collins FS, Crawford DC, Cupples LA, Danesh J, de Faire U, den Ruijter HM, Erbel R, Erdmann J, Eriksson JG, Farrall M, Ferrannini E, Ferrieres J, Ford I, Forouhi NG, Forrester T, Gansevoort RT, Gejman PV, Gieger C, Golay A, Gottesman O, Gudnason V, Gyllenstein U, Haas DW, Hall AS, Harris TB, Hattersley AT, Heath AC, Hengstenberg C, Hicks AA, Hindorff LA, Hingorani AD, Hofman A, Hovingh GK, Humphries SE, Hunt SC, Hypponen E, Jacobs KB, Jarvelin MR, Jousilahti P, Jula AM, Kaprio J, Kastelein JJ, Kayser M, Kee F, Keinanen-Kiukkaanniemi SM, Kiemeny LA, Kooner JS, Kooperberg C, Koskinen S, Kovacs P, Kraja AT, Kumari M, Kuusisto J, Lakka TA, Langenberg C, Le Marchand L, Lehtimäki T, Lupoli S, Madden PA, Mannisto S, Manunta P, Marette A, Matise TC, McKnight B, Meitinger T, Moll FL, Montgomery GW, Morris AD, Morris AP, Murray JC, Nelis M, Ohlsson C, Oldehinkel AJ, Ong KK, Ouwehand WH, Pasterkamp G, Peters A, Pramstaller PP, Price JF, Qi L, Raitakari OT, Rankinen T, Rao DC, Rice TK, Ritchie M, Rudan I, Salomaa V, Samani NJ, Saramies J, Sarzynski MA, Schwarz PE, Sebert S, Sever P, Shuldiner AR, Sinisalo J, Steinthorsdottir V, Stolk RP, Tardif JC, Tonjes A, Tremblay A, Tremoli E, Virtamo J, Vohl MC, Electronic Medical R, Genomics C, Consortium MI, Consortium P, LifeLines Cohort S, Amouyel P, Asselbergs FW, Assimes TL, Bochud M, Boehm BO, Boerwinkle E, Bottinger EP, Bouchard C, Cauchi S, Chambers JC, Chanock SJ, Cooper RS, de Bakker PI, Dedoussis G, Ferrucci L, Franks PW, Froguel P, Groop LC, Haiman CA, Hamsten A, Hayes MG, Hui J, Hunter DJ, Hveem K, Jukema JW, Kaplan RC, Kivimäki M, Kuh D, Laakso M, Liu Y, Martin NG, Marz W, Melbye M, Moebus S, Munroe PB, Njolstad I, Oostra BA, Palmer CN, Pedersen NL, Perola M, Perusse L, Peters U, Powell JE, Power C, Quertermous T,

Rauramaa R, Reinmaa E, Ridker PM, Rivadeneira F, Rotter JI, Saaristo TE, Saleheen D, Schlessinger D, Slagboom PE, Snieder H, Spector TD, Strauch K, Stumvoll M, Tuomilehto J, Uusitupa M, van der Harst P, Volzke H, Walker M, Wareham NJ, Watkins H, Wichmann HE, Wilson JF, Zanen P, Deloukas P, Heid IM, Lindgren CM, Mohlke KL, Speliotes EK, Thorsteinsdottir U, Barroso I, Fox CS, North KE, Strachan DP, Beckmann JS, Berndt SI, Boehnke M, Borecki IB, McCarthy MI, Metspalu A, Stefansson K, Uitterlinden AG, van Duijn CM, Franke L, Willer CJ, Price AL, Lettre G, Loos RJ, Weedon MN, Ingelsson E, O'Connell JR, Abecasis GR, Chasman DI, Goddard ME, Visscher PM, Hirschhorn JN, Frayling TM: Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature genetics* 2014;46:1173-1186

22. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkila K, Hypponen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikainen LP, Magnusson PKE, Mangino M, Mihailov E, Montasser ME, Muller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney ASF, Doring A, Elliott P, Epstein SE, Ingi Eyjolfsson G, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJP, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindstrom J, Loos RJF, Mach F, McArdle WL, Meisinger C, Mitchell BD, Muller G, Nagaraja R, Narisu N, Nieminen TVM, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruukonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stancakova A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YI, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrieres J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllenstein U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Jarvelin MR, Jula A, Kahonen M, Kaprio J, Kesaniemi A, Kivimäki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, Marz W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njolstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PEH, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BHR, Ordovas JM, Boerwinkle E, Palmer CNA, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR, Global Lipids Genetics C: Discovery and refinement of loci associated with lipid levels. *Nature genetics* 2013;45:1274-1283

23. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J, Overvad K, Tjonneland A, Clavel-Chapelon F, Thiebaut A, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-De-Mesquita HB, Peeters PH, Lund E, Engeset D, Gonzalez CA, Barricarte A, Berglund G, Hallmans G, Day NE, Key TJ, Kaaks R, Saracci R: European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public health nutrition* 2002;5:1113-1124
24. Kroke A, Klipstein-Grobusch K, Voss S, Moseneder J, Thielecke F, Noack R, Boeing H: Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *The American journal of clinical nutrition* 1999;70:439-447
25. Margetts BM, Pietinen P: European Prospective Investigation into Cancer and Nutrition: validity studies on dietary assessment methods. *International journal of epidemiology* 1997;26 Suppl 1:S1-5
26. Forouhi NG, Koulman A, Sharp SJ, Imamura F, Kroger J, Schulze MB, Crowe FL, Huerta JM, Guevara M, Beulens JW, van Woudenberg GJ, Wang L, Summerhill K, Griffin JL, Feskens EJ, Amiano P, Boeing H, Clavel-Chapelon F, Dartois L, Fagherazzi G, Franks PW, Gonzalez C, Jakobsen MU, Kaaks R, Key TJ, Khaw KT, Kuhn T, Mattiello A, Nilsson PM, Overvad K, Pala V, Palli D, Quiros JR, Rolandsson O, Roswall N, Sacerdote C, Sanchez MJ, Slimani N, Spijkerman AM, Tjonneland A, Tormo MJ, Tumino R, van der AD, van der Schouw YT, Langenberg C, Riboli E, Wareham NJ: Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct case-cohort study. *The lancet Diabetes & endocrinology* 2014;2:810-818
27. Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, Day NE: 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 nutrition* 2003;6:407-413
28. Burgess S, Thompson SG: Avoiding bias from weak instruments in Mendelian randomization studies. *International journal of epidemiology* 2011;40:755-764
29. Riserus U, Marklund M: Milk fat biomarkers and cardiometabolic disease. *Current opinion in lipidology* 2017;28:46-51
30. Prentice RL: A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika* 1986;73:1-11
31. Boef AG, Dekkers OM, le Cessie S: Mendelian randomization studies: a review of the approaches used and the quality of reporting. *International journal of epidemiology* 2015;44:496-511
32. R: A language and environment for statistical computing. [article online], 2017. Available from <https://www.R-project.org/>.
33. Sluijs I, Forouhi NG, Beulens JW, van der Schouw YT, Agnoli C, Arriola L, Balkau B, Barricarte A, Boeing H, Bueno-de-Mesquita HB, Clavel-Chapelon F, Crowe FL, de Lauzon-Guillain B, Drogan D, Franks PW, Gavrila D, Gonzalez C, Halkjaer J, Kaaks R, Moskal A, Nilsson P, Overvad K, Palli D, Panico S, Quiros JR, Ricceri F, Rinaldi S, Rolandsson O, Sacerdote C, Sanchez MJ, Slimani N, Spijkerman AM, Teucher B, Tjonneland A, Tormo MJ, Tumino R, van der AD, Sharp SJ, Langenberg C, Feskens EJ, Riboli E, Wareham NJ: The amount and type of dairy product intake

and incident type 2 diabetes: results from the EPIC-InterAct Study. *The American journal of clinical nutrition* 2012;96:382-390

34. Mahajan A, Go MJ, Zhang W, Below JE, Gaulton KJ, Ferreira T, Horikoshi M, Johnson AD, Ng MC, Prokopenko I, Saleheen D, Wang X, Zeggini E, Abecasis GR, Adair LS, Almgren P, Atalay M, Aung T, Baldassarre D, Balkau B, Bao Y, Barnett AH, Barroso I, Basit A, Been LF, Beilby J, Bell GI, Benediktsson R, Bergman RN, Boehm BO, Boerwinkle E, Bonnycastle LL, Burt N, Cai Q, Campbell H, Carey J, Cauchi S, Caulfield M, Chan JC, Chang LC, Chang TJ, Chang YC, Charpentier G, Chen CH, Chen H, Chen YT, Chia KS, Chidambaram M, Chines PS, Cho NH, Cho YM, Chuang LM, Collins FS, Cornelis MC, Couper DJ, Crenshaw AT, van Dam RM, Danesh J, Das D, de Faire U, Dedoussis G, Deloukas P, Dimas AS, Dina C, Doney AS, Donnelly PJ, Dorkhan M, van Duijn C, Dupuis J, Edkins S, Elliott P, Emilsson V, Erbel R, Eriksson JG, Escobedo J, Esko T, Eury E, Florez JC, Fontanillas P, Forouhi NG, Forsen T, Fox C, Fraser RM, Frayling TM, Froguel P, Frossard P, Gao Y, Gertow K, Gieger C, Gigante B, Grallert H, Grant GB, Grrop LC, Groves CJ, Grundberg E, Guiducci C, Hamsten A, Han BG, Hara K, Hassanali N, Hattersley AT, Hayward C, Hedman AK, Herder C, Hofman A, Holmen OL, Hovingh K, Hreidarsson AB, Hu C, Hu FB, Hui J, Humphries SE, Hunt SE, Hunter DJ, Hveem K, Hydrie ZI, Ikegami H, Illig T, Ingelsson E, Islam M, Isomaa B, Jackson AU, Jafar T, James A, Jia W, Jockel KH, Jonsson A, Jowett JB, Kadowaki T, Kang HM, Kanoni S, Kao WH, Kathiresan S, Kato N, Katulanda P, Keinanen-Kiukaanniemi KM, Kelly AM, Khan H, Khaw KT, Khor CC, Kim HL, Kim S, Kim YJ, Kinnunen L, Klopp N, Kong A, Korpi-Hyovalti E, Kowlessur S, Kraft P, Kravic J, Kristensen MM, Krithika S, Kumar A, Kumate J, Kuusisto J, Kwak SH, Laakso M, Lagou V, Lakka TA, Langenberg C, Langford C, Lawrence R, Leander K, Lee JM, Lee NR, Li M, Li X, Li Y, Liang J, Liju S, Lim WY, Lind L, Lindgren CM, Lindholm E, Liu CT, Liu JJ, Lobbens S, Long J, Loos RJ, Lu W, Luan J, Lyssenko V, Ma RC, Maeda S, Magi R, Mannisto S, Matthews DR, Meigs JB, Melander O, Metspalu A, Meyer J, Mirza G, Mihailov E, Moebus S, Mohan V, Mohlke KL, Morris AD, Muhleisen TW, Muller-Nurasyid M, Musk B, Nakamura J, Nakashima E, Navarro P, Ng PK, Nica AC, Nilsson PM, Njolstad I, Nothen MM, Ohnaka K, Ong TH, Owen KR, Palmer CN, Pankow JS, Park KS, Parkin M, Pechlivanis S, Pedersen NL, Peltonen L, Perry JR, Peters A, Pinidiyapathirage JM, Platou CG, Potter S, Price JF, Qi L, Radha V, Rallidis L, Rasheed A, Rathman W, Rauramaa R, Raychaudhuri S, Rayner NW, Rees SD, Rehnberg E, Ripatti S, Robertson N, Roden M, Rossin EJ, Rudan I, Rybin D, Saaristo TE, Salomaa V, Saltevo J, Samuel M, Sanghera DK, Saramies J, Scott J, Scott LJ, Scott RA, Segre AV, Sehmi J, Sennblad B, Shah N, Shah S, Shera AS, Shu XO, Shuldiner AR, Sigurdsson G, Sijbrands E, Silveira A, Sim X, Sivapalaratnam S, Small KS, So WY, Stancakova A, Stefansson K, Steinbach G, Steinthorsdottir V, Stirrups K, Strawbridge RJ, Stringham HM, Sun Q, Suo C, Syvanen AC, Takayanagi R, Takeuchi F, Tay WT, Teslovich TM, Thorand B, Thorleifsson G, Thorsteinsdottir U, Tikkanen E, Trakalo J, Tremoli E, Trip MD, Tsai FJ, Tuomi T, Tuomilehto J, Uitterlinden AG, Valladares-Salgado A, Vedantam S, Veglia F, Voight BF, Wang C, Wareham NJ, Wennauer R, Wickremasinghe AR, Wilsgaard T, Wilson JF, Wiltshire S, Winckler W, Wong TY, Wood AR, Wu JY, Wu Y, Yamamoto K, Yamauchi T, Yang M, Yengo L, Yokota M, Young R, Zabaneh D, Zhang F, Zhang R, Zheng W, Zimmet PZ, Altshuler D, Bowden DW, Cho YS, Cox NJ, Cruz M, Hanis CL, Kooner J, Lee JY, Seielstad M, Teo YY, Boehnke M, Parra EJ, Chambers JC, Tai ES, McCarthy MI, Morris AP: Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics* 2014;46:234-244

35. Enattah NS, Forsblom C, Rasinpera H, Tuomi T, Groop PH, Jarvela I: The genetic variant of lactase persistence C (-13910) T as a risk factor for type I and II diabetes in the Finnish population. *European journal of clinical nutrition* 2004;58:1319-1322
36. Ocke MC, Bueno-de-Mesquita HB, Goddijn HE, Jansen A, Pols MA, van Staveren WA, Kromhout D: The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *International journal of epidemiology* 1997;26 Suppl 1:S37-48
37. Pisani P, Faggiano F, Krogh V, Palli D, Vineis P, Berrino F: Relative validity and reproducibility of a food frequency dietary questionnaire for use in the Italian EPIC centres. *International journal of epidemiology* 1997;26 Suppl 1:S152-160
38. Rodriguez S, Gaunt TR, Day IN: Hardy-Weinberg equilibrium testing of biological ascertainment for Mendelian randomization studies. *American journal of epidemiology* 2009;169:505-514
39. Smith GD, Ebrahim S: Mendelian randomization: prospects, potentials, and limitations. *International journal of epidemiology* 2004;33:30-42
40. Holmes MV, Ala-Korpela M, Smith GD: Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nature reviews Cardiology* 2017;