Supplementary appendix

Supplement to: Dietary fat quality and genetic risk of type 2 diabetes: meta-analysis of individual participant data from prospective cohort studies

Jordi Merino, *postdoctoral research fellow*[1,2,3*], Marta Guasch-Ferré, *postdoctoral research fellow*[4*], Christina Ellervik, *associate professor*[5,6*], Hassan S. Dashti, *postdoctoral research fellow*[1,2*], Stephen J. Sharp, *senior statistician*[7], Peitao Wu, *doctoral researcher*[8], Kim Overvad, *professor* [9,10], Chloé Sarnowski, *postdoctoral research fellow*[8], Mikko Kuokkanen, *postdoctoral research fellow*[11,12], Rozenn N. Lemaitre, *associate professor*[13], Anne E. Justice, *assistant professor*[14,15], Ulrika Ericson, *associate professor*[16], Kim V.E. Braun, *postdoctoral research fellow*[17], Yuvaraj Mahendran, *postdoctoral research fellow*[18,19], Alexis C. Frazier-Wood, *associate professor*[20], Dianjianyi Sun, *postdoctoral research fellow*[21], Audrey Y. Chu, *research fellow* [22], Toshiko Tanaka, *staff scientist*[23], Jian'an Luan, *senior statistician*[7], Jaeyoung Hong, *postdoctoral research fellow*[8], Anne Tjønneland, *professor*[24], Ming Ding, *research fellow*[4], Annamari Lundqvist, *adjunct professor*[11], Kenneth Mukamal, *associate professor*[25], Rebecca Rohde, *doctoral researcher*[15], Christina-Alexandra Schulz, *doctoral researcher*[16], Oscar H. Franco, *professor*[17], Niels Grarup, *associate professor*[18], Yii-Der Ida Chen, *research fellow*[26], Lydia Bazzano, *associate professor*[21], Paul W. Franks, *professor*[4,27,28,29], Julie E. Buring, *professor*[30], Claudia Langenberg, *group leader*[7], Ching-Ti Liu, *associate professor*[8], Torben Hansen, *professor*[18], Majken K. Jensen, *associate professor*[4,31], Katri Sääksjärvi, *research fellow*[11], Bruce M. Psaty, *professor*[13], Kristin L. Young, *research assistant professor*[15], George Hindy, *postdoctoral research fellow*[1,2,16], Camilla Helene Sandholt, *research fellow*[18], Paul M. Ridker, *professor*[30,32], Jose M. Ordovas, *professor*[33,34,35], James B. Meigs, *professor*[36,37], Oluf Pedersen, *professor*[18], Peter Kraft, *professor*[38,39], Markus Perola, *professor*[11,12,40], Kari E. North. *professor*[15], Marju Orho-Melander, *professor*[16], Trudy Voortman, *research fellow*[17], Ulla Toft, *associate professor*[41,42], Jerome I. Rotter, *professor*[26], Lu Qi, *professor*[4,21,31], Nita G. Forouhi, *professor*[7], Dariush Mozaffarian, *professor*[43], Thorkild I.A. Sørensen, *professor*[18,44], Meir J. Stampfer, *professor*[4,31,38], Satu Männistö, *research fellow*[11], Elizabeth Selvin, *professor*[45], Fumiaki Imamura, *senior investigator scientist*[7], Veikko Salomaa, *professor*[11], Frank B. Hu, *professor*[4,31], Nick J. Wareham, MBBS, *professor*[7], Josée Dupuis, *professor*[8,46], Caren E. Smith, *research scientist*[33#], Tuomas O. Kilpeläinen, *associate professor*[18#], Daniel I. Chasman, *associate professor* [#2,30,47] Jose C. Florez, *professor*[1,2,3,37#]

* Contributed equally as first authors
\# Contributed equally as senior authors

Address for Correspondence: Jose C. Florez, MD, PhD and Jordi Merino, PhD. Diabetes Unit and Center for Genomic Medicine Massachusetts General Hospital 185 Cambridge Street, CPZN 5.250 Boston, MA 02114. jcflorez@partners.org and jmerino@mgh.harvard.edu

Table of Contents:

Supplementary table 1. Components of the type 2 diabetes genetic risk score

Chr.	Pos (hg38)	SNP	Locus	EA	Other	T2D OR	95%CI
1	119,975,336	rs10923931	*NOTCH2*	T	G	1.08	1.04-1.12
1	50,444,313	rs17106184	*FAF1*	A	G	0.91	0.87-0.95
1	213,981,376	rs2075423	*PROX1*	T	G	0.93	0.91-0.95
2	43,462,891	rs10203174	*THADA*	T	C	0.87	0.84-0.91
2	164,672,366	rs13389219	*GRB14 / COBLL1*	T	C	0.93	0.91-0.95
2	60,341,610	rs243088	*BCL11A*	T	A	1.07	1.04-1.09
2	226,228,869	rs2943640	*IRS1*	A	C	0.91	0.89-0.93
2	160,489,936	rs7569522	*RBMS1*	A	G	1.05	1.03-1.07
2	27,518,370	rs780094	*GCKR*	T	C	0.94	0.92-0.96
3	123,363,551	rs11717195	*ADCY5*	C	T	0.90	0.88-0.92
3	64,104,687	rs12497268	*PSMD6*	C	G	0.97	0.94-0.99
3	23,413,299	rs1496653	*UBE2E2*	G	A	0.92	0.90-0.95
3	186,895,620	rs17301514	*ST6GAL1*	A	G	1.05	1.01-1.09
3	12,351,626	rs1801282	*PPARG*	G	C	0.88	0.86-0.91
3	185,793,899	rs4402960	*IGF2BP2*	T	G	1.13	1.10-1.16
3	64,719,689	rs6795735	*ADAMTS9*	T	C	0.93	0.90-0.95
3	188,022,735	rs6808574	*LPP*	T	C	0.94	0.92-0.69
4	6,288,259	rs4458523	*WFS1*	T	G	0.91	0.89-0.93
4	152,599,323	rs6813195	*TMEM154*	T	C	0.93	0.91-0.96
4	1,299,457	rs6819243	*MAEA*	C	T	0.93	0.87-0.99
5	56,510,924	rs459193	*ANKRD55*	A	G	0.93	0.90-0.95
5	53,976,834	rs4865796	*ARL15*	G	A	0.94	0.92-0.97
5	77,131,486	rs6878122	*ZBED3*	G	A	1.10	1.07-1.13
6	39,316,274	rs1535500	*KCNK16*	T	G	1.09	1.05-1.13
6	7,212,967	rs17762454	*SSR1 / RREB1*	T	C	1.04	1.01-1.07
6	38,209,891	rs4299828	*ZFAND3 / BTBD9*	G	A	0.96	0.94-0.99
6	20,679,478	rs7756992	*CDKAL1*	G	A	1.17	1.14-1.20
7	130,752,930	rs13233731	*KLF14*	G	A	1.05	1.02-1.07
7	14,858,657	rs17168486	*DGKB*	T	C	1.11	1.07-1.14
7	127,356,783	rs17867832	*GCC1 / PAX4 / ZNF800*	G	T	0.92	0.87-0.97
7	44,192,287	rs6975024	*GCK*	C	T	1.07	1.04-1.11
7	28,156,794	rs849135	*JAZF1*	G	A	1.11	1.08-1.13
8	117,172,786	rs3802177	*SLC30A8*	A	G	0.88	0.86-0.9
8	41,661,730	rs516946	*ANK1*	T	C	0.92	0.89-0.94
8	94,925,274	rs7845219	*TP53INP1 / NDUFAF6*	T	C	1.06	1.03-1.08
9	4,292,083	rs10758593	*GLIS3*	A	G	1.06	1.04-1.09
9	8,369,533	rs16927668	*PTPRD*	T	C	1.04	1.01-1.07
9	79,290,675	rs17791513	*TLE4*	G	A	0.89	0.85-0.93
10	119,389,891	rs10886471	*GRK5*	C	T	1.01	0.98-1.04
10	92,703,125	rs1111875	*HHEX / IDE*	T	C	0.90	0.88-0.92

10	12,265,895	rs11257655	*CDC123*	T	C	1.07	1.04-1.10
10	79,182,874	rs12571751	*ZMIZ1*	G	A	0.93	0.91-0.95
10	12,286,011	rs12779790	*CAMK1D*	G	A	1.09	1.06-1.13
10	112,998,590	rs7903146	*TCF7L2*	T	C	1.39	1.35-1.42
11	92,975,544	rs10830963	*MTNR1B*	G	C	1.10	1.07-1.13
11	72,722,053	rs1552224	*ARAP1*	C	A	0.90	0.880.93
11	2,825,839	rs163184	*KCNQ1*	G	T	1.09	1.06-1.11
11	1,675,619	rs2334499	*DUSP8 / MOB2 / FAM99A*	T	C	1.04	1.02-1.06
12	27,812,217	rs10842994	*KLHDC5*	T	C	0.91	0.89-0.94
12	4,265,207	rs11063069	*CCND2*	G	A	1.08	1.05-1.11
12	110,900,990	rs11065756	*CCDC63 / MYL2*	C	T	1.03	0.99-1.07
12	120,989,098	rs12427353	*HNF1A / OASL*	C	G	0.92	0.90-0.95
12	65,818,538	rs2261181	*HMGA2*	T	C	1.13	1.08-1.17
12	71,039,513	rs7955901	*TSPAN8 / CTD-2021H9.2*	C	T	1.07	1.05-1.10
13	80,143,021	rs1359790	*SPRY2*	A	G	0.93	0.91-0.95
15	91,000,846	rs12899811	*PRC1*	G	A	1.08	1.05-1.10
15	89,831,025	rs2028299	*AP3S2*	C	A	1.07	1.04-1.10
15	62,090,956	rs4502156	*C2CD4A / B*	C	T	0.94	0.92-0.97
15	77,540,420	rs7177055	*HMG20A*	G	A	0.93	0.91-0.95
15	38,530,704	rs7403531	*RASGRP1*	T	C	1.02	0.99-1.05
16	75,213,347	rs7202877	*BCAR1 / CTRB1*	G	T	0.89	0.86-0.93
16	53,785,257	rs9936385	*FTO*	C	T	1.13	1.10-1.16
17	2,312,964	rs391300	*SRR*	T	C	1.01	0.99-1.04
18	60,217,517	rs12970134	*MC4R*	A	G	1.08	1.05-1.11
18	7,068,463	rs8090011	*LAMA*	G	C	1.06	1.03-1.09
19	19,296,909	rs10401969	*CLIP2 / SUGP1*	C	T	1.13	1.09-1.18
19	33,418,804	rs8182584	*PEPD*	T	G	1.04	1.01-1.07
20	44,360,627	rs4812829	*HNF4A*	A	G	1.06	1.03-1.09

Table legend; Position based on hg38; SNP, single nucleotide polymorphism; EA, effect allele; MAF, minor allele frequency; OR, odds ratio; T2D, type 2 diabetes. Effect of T2D-raising genetic variants (OR and 95%CI) on T2D odds obtained from publicly-available data from DIAGRAM (34,840 cases and 114,981 controls).

Supplementary table 2. Genotyping information in participating cohorts

Cohort	Genotyping platform and SNP panel	Genotyping calling algorithm	SNP QC			Imputation stats	
			MAF	HWE	Call rate	Imputation software	Imputation quality metrics
ARIC	Affymetrix Human SNP Array 6.0	Birdseed	1%	1×10^{-5}	95%	IMPUTE2	r^2> 0.3
BHS	Illumina Human610, ExomeBeadChip	Illumina BeadStudio	1%	1×10^{-5}	< 97%	MACH	r^2> 0.3
CHS	Illumina HumanCNV370-Duo BeadChip	Illumina BeadStudio	1%	1×10^{-5}	< 95%	BIMBAM v0.99	r^2> 0.3
DCH	Illumina HumanCoreExome BeadChip	Illumina BeadStudio	1%	1×10^{-5}	< 95%	IMPUTE2	r^2>0.3
EPIC-InterAct	Illumina 660W-Quad BeadChip, Illumina HumanCoreExome-12, Illumina HumanCoreExome-24	Illuminus (Illumina 660) GenCall (Core Exome)	1%	1×10^{-5}	< 95%	IMPUTE2	r^2> 0.3
FHS	Affymetrix GeneChip 500K, MIPS 50K	BRLMM	1%	1×10^{-6}	95%	MACH	r^2> 0.3
FINRISK	Illumina 670K, HumanCoreExome	Illumina BeadStudio Suite/Zcall	1%	1×10^{-6}	95%	IMPUTE2	r^2> 0.3
Health 2000	Illumina 610K	Illumina BeadStudio Suite	1%	1×10^{-6}	95%	IMPUTE2	r^2> 0.3
HPFS	Affymetrix Human SNP Array 6.0	Birdseed	1%	1×10^{-6}	95%	MACH	r^2> 0.3
Inter99	Illumina ExomeBeadChip, Metabochip	Illumina GenCall	5%	1×10^{-4}	>95%	IMPUTE2	r^2> 0.3
MDC-CC	Illumina OmniExpressExome	Illumina GenomeStudio	N/A	1×10^{-6}	95%	IMPUTE2	r^2> 0.3
MESA	Affymetrix Human SNP Array 6.0	Birdseed	5%	1×10^{-5}	95%	IMPUTE2	r^2> 0.3
NHS	Affymetrix Human SNP Array 6.0	Birdseed	2%	1×10^{-6}	98%	MACH	r^2> 0.3
RS-I	Illumina HumanMap 550K	BeadStudio	1%	1×10^{-6}	98%	MACH	r^2> 0.5
WGHS	Illumina HumanHap Duo+	Illumina BeadStudio	1%	1×10^{-6}	97%	MACH	r^2> 0.3

Table legend; Genotyping platform, calling algorithm, quality control and imputation metrics used in each of the participating cohorts. MAF; Minor allele frequency threshold detection, HWE; Hardy-Weinberg equilibrium, N/A; not available

Supplementary table 3. Dietary intake assessment in participating cohorts

Cohort	Baseline (year)	Follow-up administration	Food-items (n)	Reproducibility and validation of dietary assessment tools
ARIC	From 1987 to 1989	Every 4-y	61-item FFQ	Stevens J. et al. Reliability of a food frequency questionnaire by ethnicity, gender, age and education. Nutrition Research. Nutrition Research. 1996; 16:735-45.
BHS	1998-2001	Every 3-y	131-item FFQ	Deshmukh-Taskar P. et al. Does Food Group Consumption Vary by Differences in Socioeconomic, Demographic, and Lifestyle Factors in Young Adults? The Bogalusa Heart Study. Journal of the American Dietetic Association. 2007; 107:223–4. Jago R. et al. Physical activity and health enhancing dietary behaviors in young adults: Bogalusa Heart Study. Preventive Medicine. 2005; 41:194–202.
CHS	1989/90	1995/96	99-item FFQ	Kumanyika S, et al. Picture-sort method for administering a food frequency questionnaire to older adults. J Am Diet Assoc. 1996; 96:137-44.
DCH	From 1993 to 1997	Every 5-y	192-item FFQ	Overvad K. Development of a semi quantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. Int J Epidemiol. 1991; 20: 900–5. Tjonneland A, Validation of a semi quantitative food frequency questionnaire developed in Denmark. Int J Epidemiol 1991; 20:906–12.
EPIC-InterAct	From 1991 to 1997	Dietary intake was measured at baseline	FFQ and dietary histories validated in each country	Riboli E, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr. 2002; 5:1113–24
FHS	1991-1996	Every 4-y	126-item FFQ	Salvini S et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol 1989; 18: 858–67. Rimm EB, et al. Reproducibility and validity of an expanded self-administered semi quantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992; 135:1114-26
FINRISK	2007	Dietary intake was measured at baseline	131-item FFQ	Männistö S, et al. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. J Clin Epidemiol. 1996; 49:401-9. Kaartinen NE, et al. Relative validity of a FFQ in measuring carbohydrate fractions, dietary glycaemic index and load: exploring the effects of subject characteristics. Br J Nutr 2012; 107:1367-75.
Health 2000	2000	Dietary intake was measured at baseline	131 and 128-item FFQ	Männistö S, et al. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. J Clin Epidemiol. 1996; 49:401-9. Paalanen L Validity of a food frequency questionnaire varied by age and body mass index. J Clin Epidemiol. 2006; 59:994-1001

HPFS	1986	Every 2 to 4-y	131-item FFQ	Willett WC, et al. Reproducibility and validity of a semi quantitative food frequency questionnaire. Am J Epidemiol 1985; 122:51-65.
Inter99	1999	Year 1,3,5	198-item FFQ	Toft U, et al. The dietary quality score: validation and association with cardiovascular risk factors: the Inter99 study. Eur. J. Clin. Nutr. 2008; 62:1038–1046
MDC-CC	From 1991-1994	Dietary intake was measured at baseline	Modified diet history including 1/ 168-item FFQ 2/ 7-d record 3/ Diet interview	Callmer E, et al. Dietary assessment methods evaluated in the Malmö food study. J Intern Med 1993; 233,53-7. Elmstahl S, et al. The Malmo Food Study: the relative validity of a modified diet history method and an extensive food frequency questionnaire for measuring food intake. *Eur J Clin Nutr* 1996; 50:143-151. Riboli E, et al. The Malmo Food Study: validity of two dietary assessment methods for measuring nutrient intake. *Int J Epidemiol* 1997; 26 Suppl 1:S161-173. Elmstahl S, et al. The Malmo Food Study: the reproducibility of a novel diet history method and an extensive food frequency questionnaire. *Eur J Clin Nutr* 1996; 50:134-142. Wirfalt E, et al. Methodological report from the Malmo Diet and Cancer study: development and evaluation of altered routines in dietary data processing. Nutr J 2002; 1:3.
MESA	From 2000 to 2002	Baseline and year 2010-2012	128-item FFQ	FFQ adapted from Mayer-Davis EJ, et al. Validity and reproducibility of a food frequency interview in a Multi-Cultural Epidemiology Study. Ann Epidemiol 1999; 9:314-24
NHS	1984	Every 2 to 4-y	131-item FFQ	Willett WC, et al. Reproducibility and validity of a semi quantitative food frequency questionnaire. Am J Epidemiol 1985; 122:51-65.
RS-I	1989	Dietary intake was measured at baseline	170-item FFQ	Klipstein-Grobusch K, et al. Dietary assessment in the elderly: validation of a semi quantitative food frequency questionnaire. Eur J Clin Nutr. 1998; 52:588-96.
WGHS	1991	Dietary intake was measured at baseline	131-item FFQ	Willett WC, et al. Reproducibility and validity of a semi quantitative food frequency questionnaire. Am J Epidemiol 1985; 122:51-65.

Table legend; Dietary assessment in prospective cohort studies included in the present study. FFQ, food frequency questionnaire. Food-items indicates the number of foods assessed by each questionnaire. Baseline year considered for this study. Follow up administration shows the number of dietary points cohorts used for this analysis.

Supplementary table 4. Covariates included in the association and interaction analyses based on the pre-specified analysis plan

Cox regression models	Exposure	Step 1	Step 2
		Association Analysis (AA) covariates	**Interaction Analysis (Exposure variable + PRS + Interaction term + AA covariates)**
Model 1	Total Fat, % energy	Covariates 1 to 9	T2D ~ Total fat + PRS + Total fat*PRS + AA
Model 2	PUFA, % energy	Covariates 1 to 9 + MUFA + SFA	T2D ~ PUFA + PRS + PUFA*PRS + AA
Model 3	MUFA, % energy	Covariates 1 to 9 + PUFA + SFA	T2D ~ MUFA + PRS + MUFA*PRS + AA
Model 4	SFA, % energy	Covariates 1 to 9 + MUFA + PUFA	T2D ~ SFA + PRS + SFA*PRS + AA
Model 5	ω-3 PUFA, g/d	Covariates 1 to 9 + SFA + MUFA	T2D ~ ω-3 PUFA + PRS + ω-3 PUFA *PRS + AA
Model 6	ω-6 PUFA, g/d	Covariates 1 to 9 + SFA + MUFA	T2D ~ ω-6 PUFA + PRS + ω-6 PUFA *PRS + AA
Model 7	*trans* fat, g/d	Covariates 1 to 9 + PUFA + SFA + MUFA	T2D ~ *trans* + PRS + *trans*PRS + AA

Table legend; Covariates definitions were harmonized among cohorts and included; 1) age (years, continuous), 2) sex (men, women), 3) body mass index (BMI) (kg/m2, continuous), 4) smoking (never, former or current; if former was not reported then current vs. non-current smoker), 5) physical activity (categories of metabolic equivalents or study-specific validated scores, or leisure-time activity, as defined in individual studies), 6) family history of diabetes (yes/no, if available), 7) dyslipidaemia (yes/no, defined based on treatment with lipid-lowering drugs or diagnosis of dyslipidaemia), 8) hypertension (yes/no, defined based on treatment with antihypertensive drugs or diagnosis of hypertension), and 9) dietary factors including total energy intake (kcal/d, continuous), protein intake (% of energy intake, continuous), dietary fibre (g/d, continuous), magnesium intake (g/d, continuous), and alcohol intake (g/d, continuous).
Covariates 1 to 9 were included in all models. Different subtypes of fat according to the model were included to estimate differences in the risk of substituting a certain percentage of energy from carbohydrate with total or specific types of fat.
Abbreviations: AA: Association analysis; PRS, polygenic risk score PUFA; polyunsaturated fat, MUFA; monounsaturated fat, SFA; saturated fat, total ω-3 PUFA; total ω-3 polyunsaturated fat, total ω-6 PUFA; total ω-6 polyunsaturated fat.

Supplementary table 5. Pattern of the missing data in the present study

Study	ARIC	BHS	CHS	DCH	EPIC-InterAct	FINRISK	FHS	HPFS	Health 2000	Inter99	MDC-CC	MESA	NHS	RS-I	WGHS
Original size†	6,816	1,233	3,009	9,527	22,492	3,968	7,076	6,078	2,391	5,755	4,002	1,846	9,623	3,032	22,421
Analysed size	6,690	790	2,813	8,788	20,856	3,822	6,710	5,587	1,946	5,607	3,745	1,535	8,832	2,509	22,120
T2D incidence	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
PRS	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Total fat	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
PUFA	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
MUFA	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
SFA	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
ω-3 PUFA	√		√	√		√	√	√	√	√	√	√	√	√	√
ω-6 PUFA	√		√	√		√	√	√	√	√	√	√	√	√	√
Total *trans*	√		√			√	√	√	√	√		√	√	√	√
Age	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Sex	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
HTA	√	√	√	√	√	√	√	√	√		√	√	√	√	√
DLP	√	√	√		√	√	√	√	√		√	√	√	√	√
BMI	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Smoking	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
PA	√		√	√	√	√	√	√	√	√	√	√	√	√	√
Energy	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Protein	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Carbohydrate	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Fiber	√		√	√	√	√	√	√	√	√	√	√	√	√	√
Magnesium	√		√	√	√	√	√	√	√	√	√	√	√	√	√
Alcohol	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√

Table legend; This table shows missing data information for the main outcome, exposures and covariates in studies included in the IPD meta-analysis. A tick denotes that variable was measured in the corresponding study. Analysed sizes correspond to the size of the studies after exclusion observations with sporadically missing data (<5%).

† Original size means the number of eligible participants in each cohort based on inclusion and exclusion criteria for this specific analysis.

Abbreviations: T2D, type 2 diabetes; PRS, polygenic risk score; PUFA, polyunsaturated fat; MUFA, monounsaturated fat; SFA, saturated fat; ω-3 PUFA, total ω-3 polyunsaturated fat; ω-6 PUFA, total ω-6 polyunsaturated fat; HTA hypertension; DLP dyslipidaemia; BMI, body mass index; PA, physical activity.

Supplementary table 6. Quality assessment of the included prospective cohort studies using the Newcastle-Ottawa scale

Prospective cohort study	Selection†				Comparability‡	Outcome§			Aggregate score
	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Ascertainment of exposure	Demonstration that outcome of interest was not present at start of study	Comparability of cohorts on the basis of the design or analysis	Assessment of outcome	Was follow up long enough for outcomes to occur	Adequacy of follow up of cohorts	
ARIC	*	*	/	*	**	*	*	*	8
BHS	*	*	/	*	*	/	/	*	5
CHS	*	*	/	*	**	/	/	*	6
DCH	*	*	/	*	*	*	*	*	7
EPIC-InterAct	*	*	/	*	**	*	/	*	7
FHS	*	*	/	*	**	*	*	*	8
FINRISK	*	*	/	*	**	*	/	*	7
Health 2000	*	*	/	*	**	*	*	*	8
HPFS	/	*	/	*	**	/	*	*	6
Inter99	*	*	/	*	*	/	/	*	5
MDC-CC	*	*	/	*	**	*	*	*	8
MESA	*	*	/	*	**	*	/	*	7
NHS	/	*	/	*	**	/	*	*	6
RS-I	*	*	/	*	**	*	/	*	7
WGHS	*	*	/	*	**	*	*	*	8

Table legend; Prospective cohort studies quality assessment was performed using the Newcastle–Ottawa Scale. Cohorts were classified as a low risk of bias if they scored: from 2 to 4 stars in selection domain (good or fair quality) AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome/exposure domain. † Maximum 4 stars**;** ‡ Maximum 2 stars**;** § Maximum 3 stars. We allocated the points as following;
†Representativeness of exposed cohort: *given if the cohort was representative of the average population at risk of T2D; / given if the cohort was selected based on convenience (i.e., groups of employees) or if there was no description of the derivation of the cohort. Selection of non-exposed cohort: * given if the non-exposed cohort was drawn from the same community as the exposed cohort; / was given if it was drawn from a different source or there was no description of the cohort derivation. Exposure ascertainment: * given if obtained from secure record (hospital chart); / was given if from a written self-report or no description given. Outcome was not present at start of the study: * given if outcome of interest was not present at start of study.
‡ Comparability: ** given if the study measured all covariates were; * given if any covariate was missing.
§Assessment of outcome: * given if independent blind assessment or evidence of record linkage (i.e., through medical records); / given if through self-report or no description. Follow-up: * given if follow-up was shorted than 12 years. Adequacy of follow-up of cohorts: Adequacy of follow-up: * given if complete follow-up and all participants accounted for or if loss to follow-up was small and unlikely to introduce bias (follow-up rate >90% or description provided of those lost); / given if follow-up rate <90%, no description of those lost, or no statement.
/=study did not fulfil listed criteria; *=study fulfilled listed criteria; NA=criteria not applicable to the study.

Supplementary table 7. Demographic, lifestyle and baseline clinical characteristics

Cohort	Sex, % men (n)	Prevalence of hypertension, % (n)	Prevalence of dyslipidaemia, % (n)	Body Mass Index	Smoking status	Physical activity†	Total energy	Total protein	Total carbohydrate	Total fibre	Alcohol intake
ARIC	45.9 (3,075)	30.4 (2,036)	58.3 (3,900)	26.6 (4.6)	43.2/35.4/21.4	7.2 (1.4)	1636 (597)	17.8 (4.0)	48.5 (9.2)	17.6 (7.9)	6.6 (12.8)
BHS	43.8 (326)	13.4 (100)	66.7 (496)	27.6 (6.4)	N/A/ N/A/29.9	N/A	2100 (810)	14.8 (2.6)	51.2 (6.2)	N/A	2.1 (2.6)
CHS	37.7 (1,062)	34.8 (980)	26 (732)	26.0 (4.3)	47.8/40.7/11.5	1309 (1660)	2018 (648)	19.0 (3.2)	52.4 (7.9)	18.2 (7.0)	1.4 (3.3)
DCH	51.2 (5,496)	56.7 (6,068)	N/A	27.0 (4.5)	36.3/28.9/34.8	65.4 (44.9)	2354 (671)	16.5 (2.5)	43.5 (6.5)	20.7 (7.0)	21.3 (22.7)
EPIC-InterAct	43.2 (9,459)	27 (5,758)	44.7 (9,287)	27.5 (4.8)	43.9/29/27.1	25.8/33.3/ 21.8/19.1	2149 (636)	17.0 (3.0)	43.9 (7.0)	22.7 (7.6)	14.0 (19.9)
FHS	44.1 (3,064)	25.3 (1,761)	39.7 (2756)	26.7 (4.9)	45.7/36.1/16.3	36.2 (7.1)	1973 (669)	17.5 (3.7)	49 (8.9)	19.3 (8.9)	10.8 (15.5)
FINRISK	44.5 (1,910)	53.4 (2,268)	44 (1874)	26.6 (4.6)	57.8/24.4/17.8	81.9 (35.8)	2495 (895)	17.7 (2.5)	49.1 (6.1)	30.5 (12.6)	8.6 (14.2)
Health 2000	48.7 (976)	55.3 (1,174)	66.5 (1,373)	27.3 (4.5)	47.9/22.2/29.8	74.8 (15.7)	2245 (783)	17.2 (2.3)	41.5 (9.4)	24.5 (10.7)	12.5 (21.3)
HPFS	100 (5,399)	19.4 (1,047)	11.7 (631)	25.3 (4.5)	49/42.6/8.4	20.2 (26.9)	2031 (612)	18.4 (3.3)	45.9 (7.0)	20.9 (8.6)	12.3 (15.9)
Inter99	48 (2,569)	N/A	N/A	26.2 (4.5)	35/26/39	66 (33.1)	2337 (849)	15.1 (2.5)	53.8 (6.1)	24.5 (10.4)	15.9 (19)
MDC-CC	39.2 (1,467)	78.5 (2,939)	63.1 (2,364)	25.4 (3.7)	40.2/31.6/28.2	8208 (5875)	2328 (673)	14.9 (2.4)	45.2 (6.6)	21.3 (7.5)	10.5 (12.7)
MESA	47.9 (734)	41.5 (636)	61.8 (947)	27.1 (5.1)	45.7/43.9/10.3	81.7 (58.7)	1769 (709)	15.5 (3.1)	52 (8.8)	20.3 (9)	9.1 (16.9)
NHS	0 (0)	17.8 (1,572)	9.8 (865)	25.2 (4.8)	44.7/35.1/19.9	14 (18.2)	1756 (523)	17.8 (3.3)	46.6 (7.8)	18 (6.9)	7.2 (11.2)
RS-I	37.9 (950)	24.6 (618)	63.1 (1,584)	26.0 (3.4)	35.3/42.5/22.2	89.9 (46.9)	1984 (506)	16.8 (2.9)	43.5 (6.8)	26.6 (7.1)	10.5 (14.6)
WGHS	0 (0)	23.7 (5,237)	29.3 (6,476)	25.8 (4.8)	51/37.4/11.6	14.8 (18.4)	1732 (524)	18.8 (3.3)	51.3 (7.9)	19.1 (8.8)	4.1 (8.0)

Table legend; Definition were standardized and harmonized across participating cohorts including sex (men, women), prevalence of hypertension (yes/no, defined based on treatment with antihypertensive drugs or diagnosis of hypertension), prevalence of dyslipidaemia (yes/no, defined based on treatment with lipid-lowering drugs or diagnosis of dyslipidaemia), body mass index (kg/m^2, continuous), smoking status (never, former or current; if former was not reported then current vs. non-current smoker), †physical activity (presented as metabolic equivalents, physical activity scores, or categories of leisure-time physical activity, as defined by individual studies), and dietary factors including total energy intake (kcal/d), protein intake (% of energy intake), dietary fibre (g/d), alcohol intake (g/d). N/A; not available. Systematically missing covariates were handled by multiple imputation using multiple imputation by chained equations.

Supplementary table 8. Interaction between total and subtypes of fat intake and polygenic risk score on the risk of type 2 diabetes: Sensitivity analysis including only cohorts with repeated measurements of diet

Dietary factor†	β interaction (SE)	P value $_{interaction}$	Direction of interaction in 8 studies	$\tau^{2‡}$	Sample size
Total Fat, % energy	0.042 (0.023)	0.070	+-+-++-+	0	44,539
PUFA, % energy	0.056 (0.096)	0.560	--+++++-	0	44,539
MUFA, % energy	0.056 (0.056)	0.313	+-+--+-+	0	44,539
SFA, % energy	0.014 (0.051)	0.780	+---+--+	0	44,539
ω-3 PUFA, g/d	0.015 (0.046)	0.746	-?--+--+	0.003	43,749
ω-6 PUFA, g/d	0.013 (0.008)	0.097	+?++++-+	0	43,749
Total *trans* fat, g/d	0.014 (0.023)	0.544	-??---++	0	33,961

Table legend: For each dietary factor the combined interaction P value, heterogeneity and sample size are shown. Direction of interaction represents the sign of the beta in each cohort (Cohorts presented in alphabetical order [ARIC, BHS, DCH, FHS, HPFS, Inter99, MESA, NHS]). (? = Not available).
† Combined estimates from inverse variance-weighted random-effects meta-analysis represent an interaction effect on the risk of T2D per increment of 10 risk alleles in the PRS and the isocaloric replacement for each type of fat: Isocaloric replacement of carbohydrate with total fat, PUFA, MUFA, SFA (5% energy) and total ω-3 PUFA, ω-6 PUFA, and *trans* fat (1g/d).
‡ Between-study variance (τ^2) was used to assess heterogeneity.

Supplementary table 9. Interaction between total and subtypes of fat intake and polygenic risk score on the risk of type 2 diabetes: Sensitivity analysis including cohorts classified as a lower risk of bias

Dietary factor†	β interaction (SE)	P value $_{interaction}$	Direction of interaction in 12 studies	$\tau^{2‡}$	Sample size
Total Fat, % energy	0.021 (0.016)	0.177	+++---+--+--	0	93,140
PUFA, % energy	0.033 (0.059)	0.572	-+++++-+--+	0	93,140
MUFA, % energy	0.040 (0.034)	0.240	+++----+-+--	0	93,140
SFA, % energy	-0.004 (0.032)	0.912	+-+---+--+--	0	93,140
ω-3 PUFA, g/d	0.028 (0.034)	0.405	--?++++--+++	0.003	72,284
ω-6 PUFA, g/d	0.005 (0.005)	0.296	++?--++--+-+	0	72,284
Total *trans* fat, g/d	0.016 (0.018)	0.384	-??--+-?-+++	0.001	59,751

Table legend: For each dietary factor the combined interaction P value, heterogeneity and sample size are shown. Direction of interaction represents the sign of the beta in each cohort (Cohorts presented in alphabetical order [ARIC, DCH, EPIC-InterAct, FHS, FINRISK, Health 2000, HPFS, MDC-CC, MESA, NHS, RS-1, WGHS]). (? = Not available).

† Combined estimates from inverse variance-weighted random-effects meta-analysis represent an interaction effect on the risk of T2D per increment of 10 risk alleles in the PRS and the isocaloric replacement for each type of fat: Isocaloric replacement of carbohydrate with total fat, PUFA, MUFA, SFA (5% energy) and total ω-3 PUFA, ω-6 PUFA, and *trans* fat (1g/d).

‡ Between-study variance (τ^2) was used to assess heterogeneity.

Supplementary table 10. Combined effect of the genetic risk and the interaction term on type 2 diabetes risk

Dietary factor†	β interaction	SE interaction	Direction of interaction in included studies	β PRS	SE PRS	Direction of PRS in included studies	P value JMA	P value heterogeneity‡	Sample size
Total Fat, % energy	0.032	0.015	+-+++---++--+--	0.522	0.019	+++++++++++++++	<0.001	0.007	102,350
PUFA, % energy	0.074	0.057	--++++++++-+--+	0.533	0.019	+++++++++++++++	<0.001	0.006	102,350
MUFA, % energy	0.068	0.032	+-+++----++-+--	0.538	0.019	+++++++++++++++	<0.001	0.012	102,350
SFA, % energy[$]	0.094	0.024	+-+-----+---+--	0.527	0.019	+++++++++++++++	<0.001	<0.001	102,350
ω-3 PUFA, g/d	0.037	0.022	-?+-?--+++--+++	0.397	0.035	+?++?++++++++++	<0.001	0.015	80,704
ω-6 PUFA, g/d[$]	0.014	0.003	-?++?--+++--+-+	0.452	0.022	+?++?++++++++++	<0.001	0.017	80,704
trans fat, g/d	0.014	0.017	-?-??--+--?-+++	0.501	0.029	+?+??+++++?++++	<0.001	0.036	68,171

Table legend: For each dietary factor interaction and main effects estimates are shown. Direction of interaction and PRS represents the sign of the beta for the interaction term and main effects in each cohort (Cohorts represented in alphabetical order [ARIC, BHS, CHS, DCH, EPIC-InterAct, FHS, FINRISK, Health 2000, HPFS, Inter99, MDC-CC, MESA, NHS, RS-1, WGHS]). (? = Not available). P value $_{JMA}$ derived from combining the joint 2 *df* results (estimates of genetic effects (β_g), and interaction term (β_{ge}) and their corresponding 2×2 covariance matrix) from cohort-specific centred and uniform estimates. Heterogeneity and sample size are also shown.

† Combined estimates from fixed-effects meta-analysis represent an interaction effect on type 2 diabetes per increment of 10 risk alleles in the PRS and each of the dietary fat variables modelled as an isocaloric replacement with carbohydrates (5% of energy from total fat, PUFA, MUFA, SFA and 1g/d in total ω-3 PUFA, total ω-6 PUFA, and *trans* fat).

Individual cohort analyses were adjusted for demographic, lifestyle and clinical characteristics.

‡ Heterogeneity was assessed using the Q-statistic test and reported as heterogeneity P value.

[$] In the SFA and the total ω-6 PUFA models the computed P value for interaction term was significant. We therefore conducted a stratified analysis for quartiles of the PRS and we showed no significance across quartiles (P-value was the same as in the 1*df* meta-analysis [Table 3; 0.584 and 0.125, respectively]). This suggest that the addition of the covariance in the JMA method was driven the significance of this signal, but was not clinically meaningful.

Supplementary table 11: Combined risk of type 2 diabetes per increment of 10 risk alleles in the genetic risk score without (Model 1) and with (Models 2) additional adjustment for dietary fat types

	HR (95%CI)	P value
Model 1	1.68 (1.62, 1.74)	<0.001
Models 2;		
Mediator of interest;		
Total Fat, % energy	1.65 (1.60-1.70)	<0.001
PUFA, % energy	1.65 (1.60-1.70)	<0.001
MUFA, % energy	1.64 (1.58-1.71)	<0.001
SFA, % energy	1.67 (1.60-1.75)	<0.001
ω-3 PUFA, g/d	1.64 (1.58-1.71)	<0.001
ω-6 PUFA, g/d	1.66 (1.63, 1.70)	<0.001
Total *trans* fat, g/d	1.76 (1.67-1.86)	<0.001

Table legend: Combined T2D risk per increment of 10 risk alleles in the PRS using inverse variance-weighted random-effects meta-analysis after adjusting for demographic, lifestyle, and clinical characteristics (Model 1). Models 2 were further adjusted for specific dietary fat to investigate potential mediation effects.

Supplementary Fig 1. Flow diagram

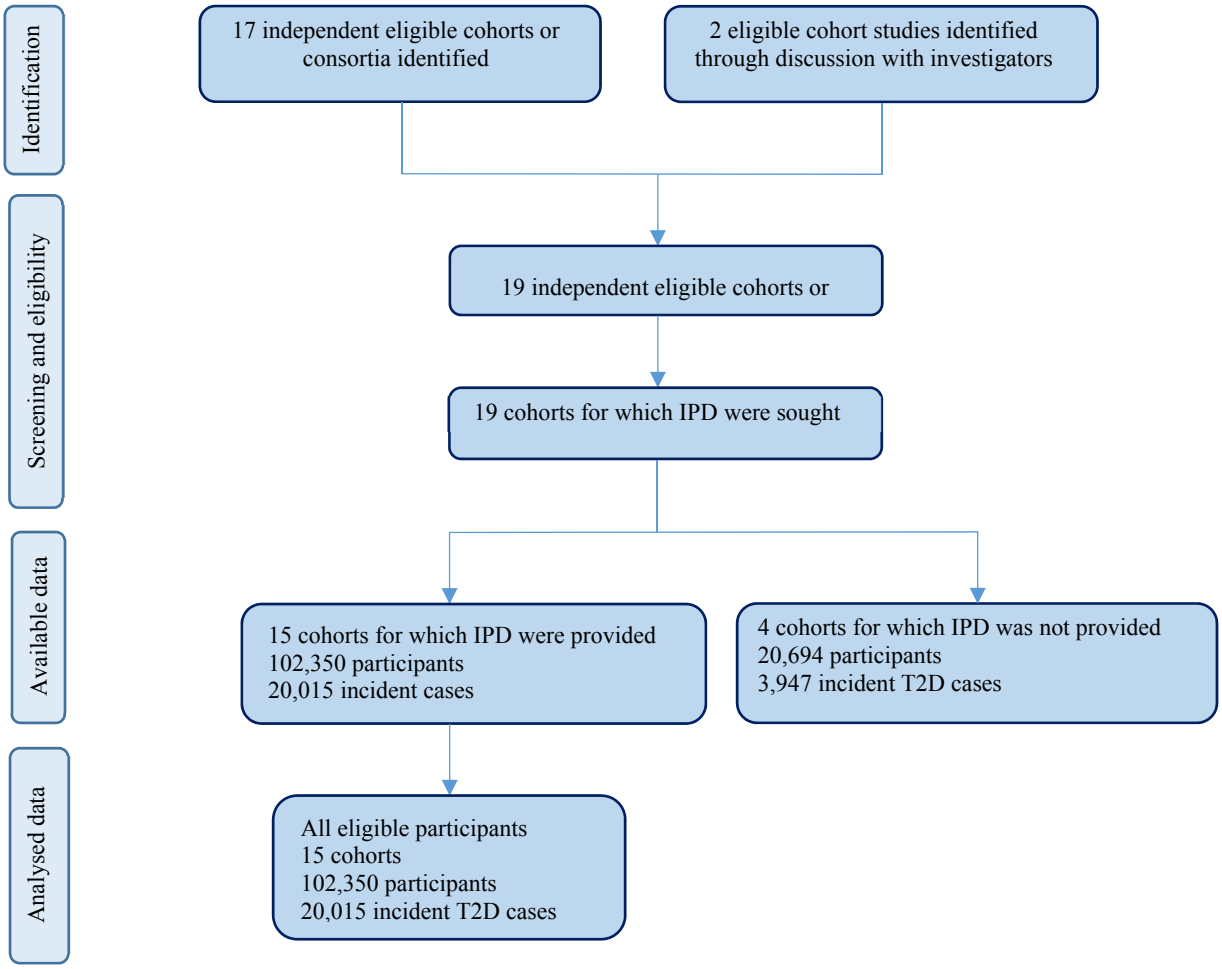


Figure legend: Identification and selection of studies in individual participant data (IPD) meta-analysis of dietary fat quality and genetic risk on type 2 diabetes incidence.

Supplementary fig 2. Association between PUFAs and total ω-3 PUFAs on the risk of type 2 diabetes; Random-effects estimates when one study at a time was removed from the analysis

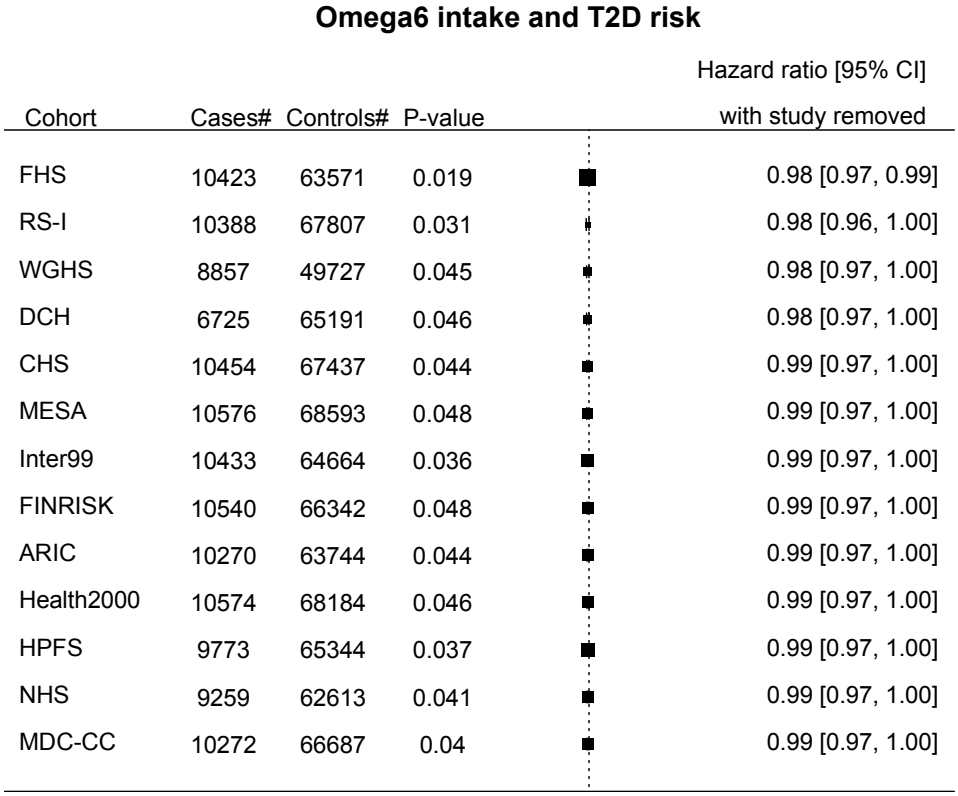


Figure legend: Combined T2D risk estimates using random-effects meta-analysis repeatedly removing one study for each iteration.
# indicates the number of remaining cases and controls after the exclusion of this cohort

Supplementary fig 3. Contour-enhanced funnel plot of the 15 prospective cohort studies included in the individual participant data meta-analysis

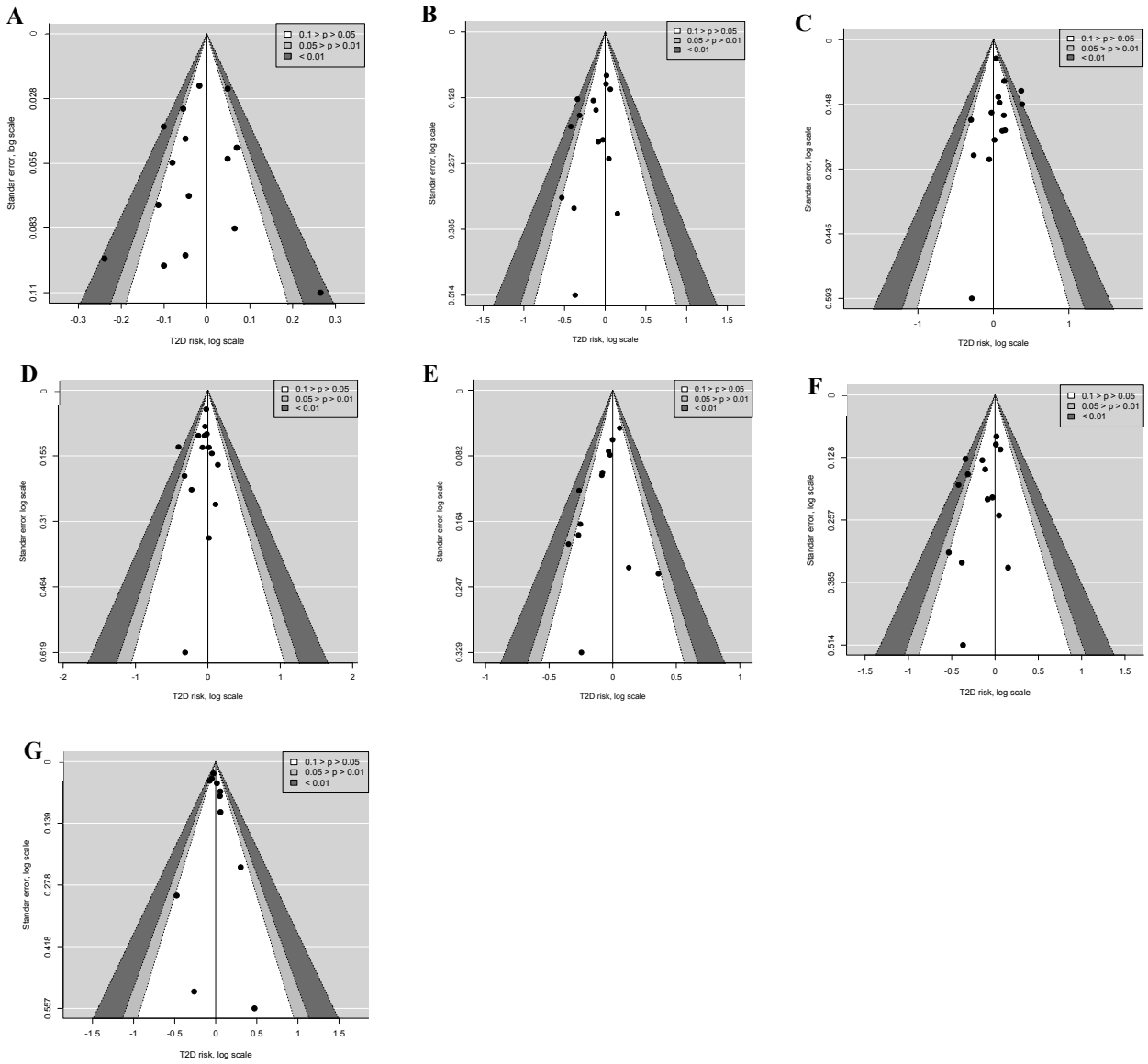
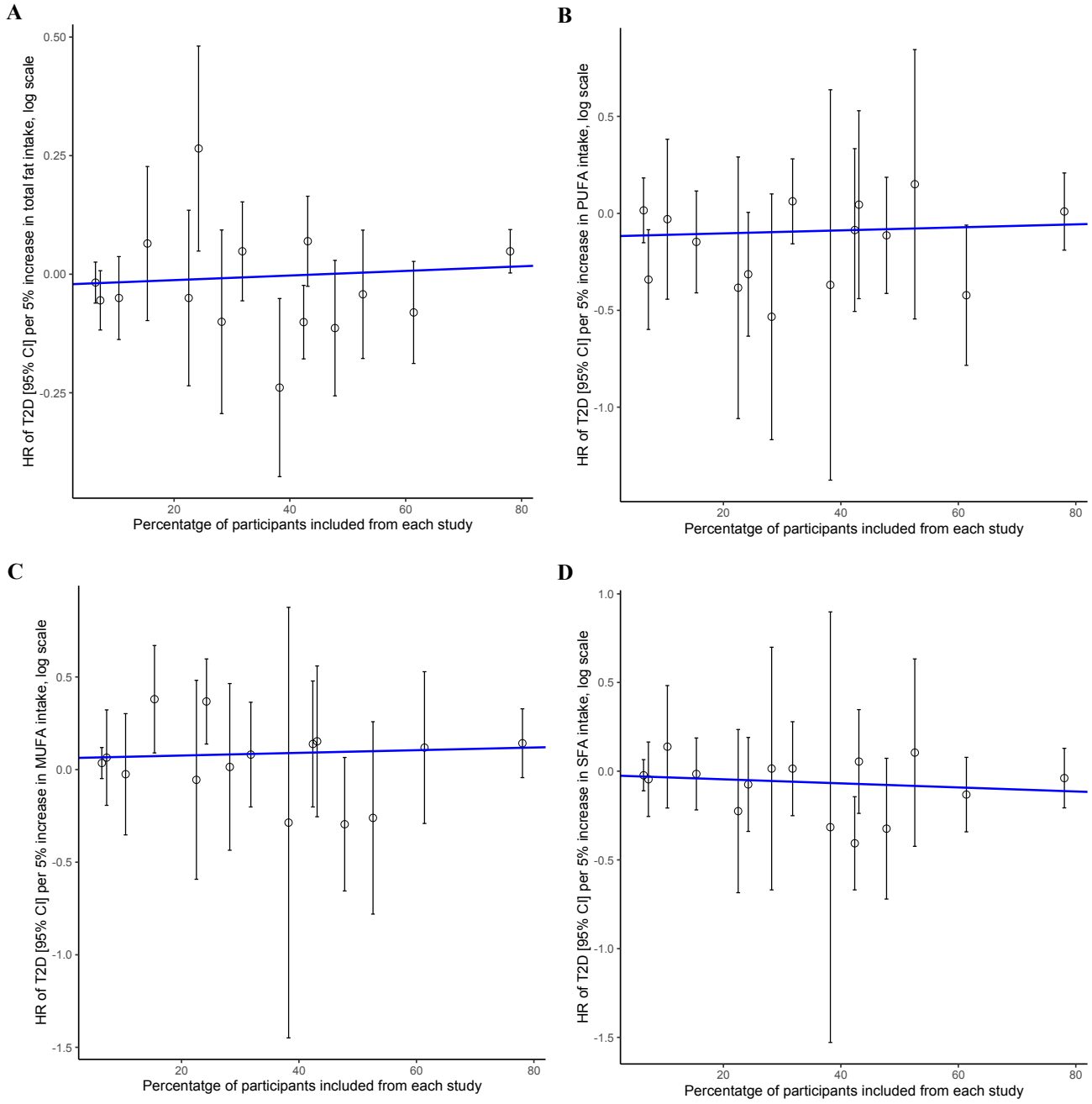


Figure legend: Contour-enhanced funnel plots of the effect estimates from individual studies included for the association between the risk of T2D and total fat (A), PUFA (B), MUFA (c), SFA (D), total ω-3 PUFA (E), total ω-6 PUFA (F), and *trans* fat (G). The y axis represents the standard error and x axes the log transformed T2D risk. The unshaded region in the middle corresponds to p-values greater than .10, the grey-shaded region corresponds to P-values between 0.10 and 0.05, the dark grey-shaded region corresponds to P-values between 0.05 and 0.01, and the region outside of the funnel corresponds to P-values below 0.01. The funnel plot is centred at 0 to denote the value under the null hypothesis of no effect. Debray's test was used to test for funnel plots asymmetry. Among the significant association with T2D reported in this IPD study, the funnel plots for the association of PUFA with T2D risk displayed statistical suggestive significant asymmetry (Debray's test P=0.051), while the association of total ω-6 PUFA and MUFA with the hazard ratio of T2D showed no evidence of asymmetry (Debray's test P=0.699 and P=0.642, respectively). For non-significant associations, P-values were 0.537 for total fat (A), 0.479 for SFA (D), 0.008 for total ω-3 PUFA (E), and 0.071 for trans fat intake (G).

Supplementary fig 4. Combined risk of type 2 diabetes with isocaloric replacement (5% energy or 1g/d) of carbohydrate with total or subtypes of fat plotted against the percentage of individuals from each study with



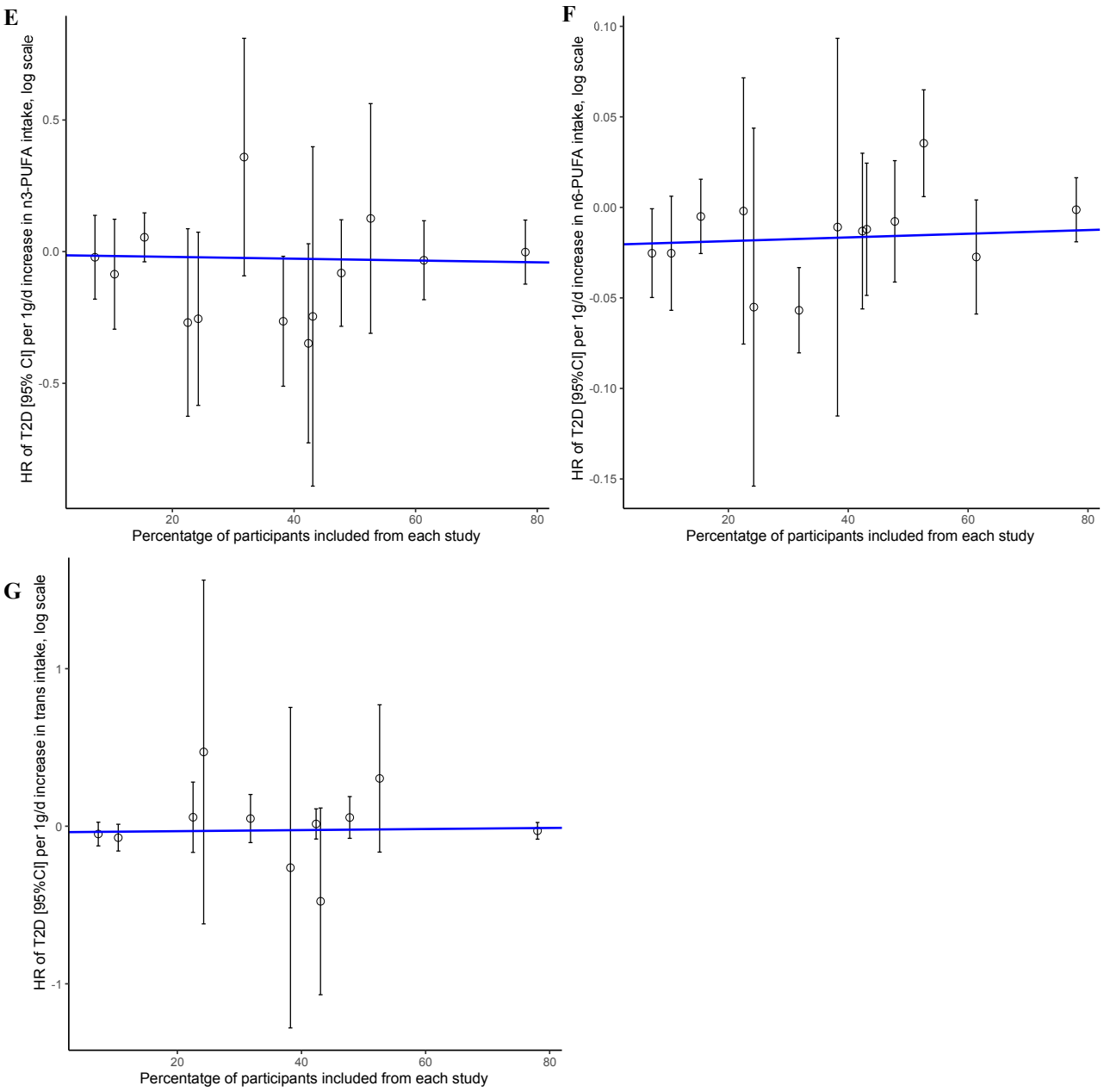
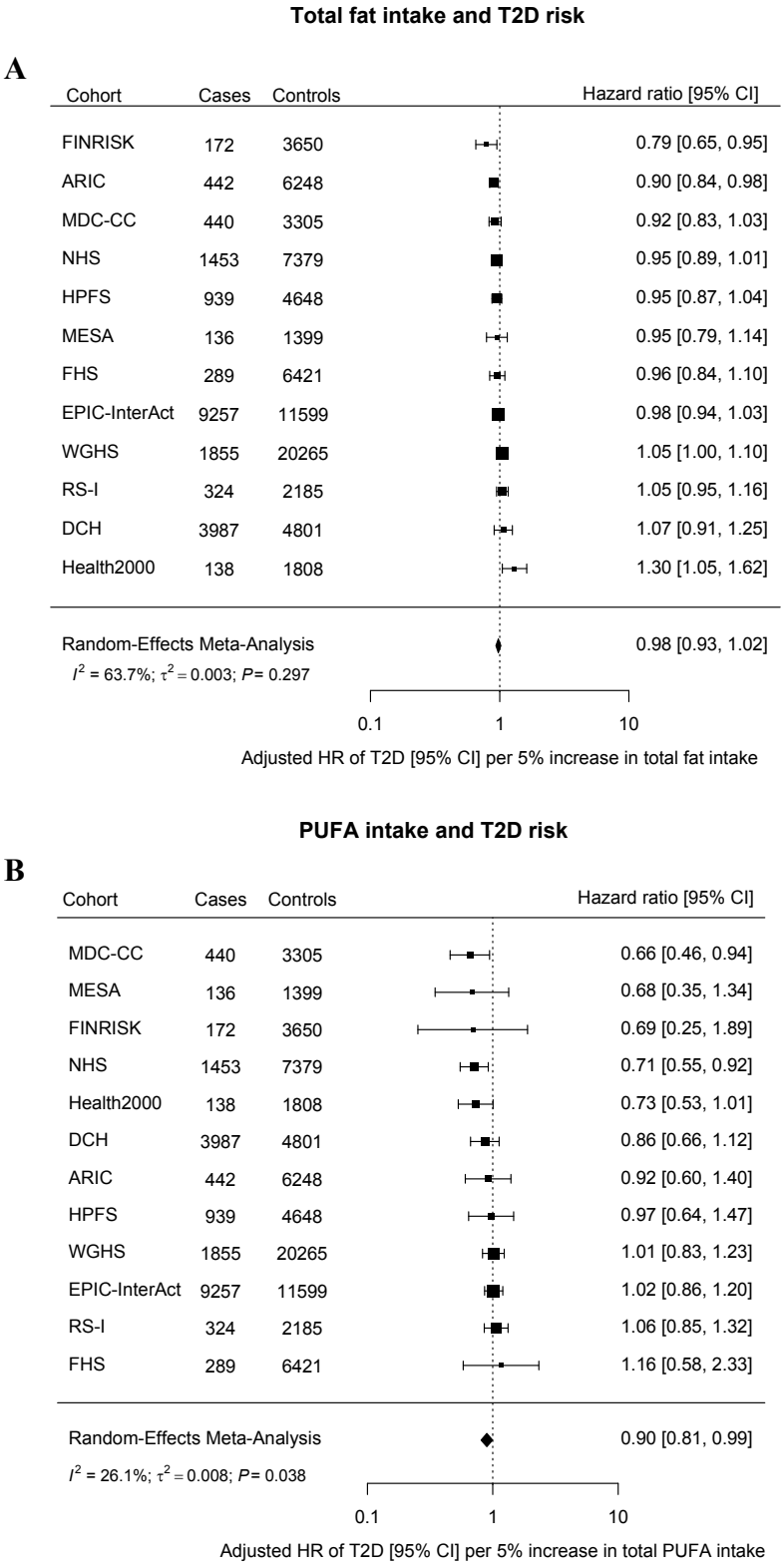


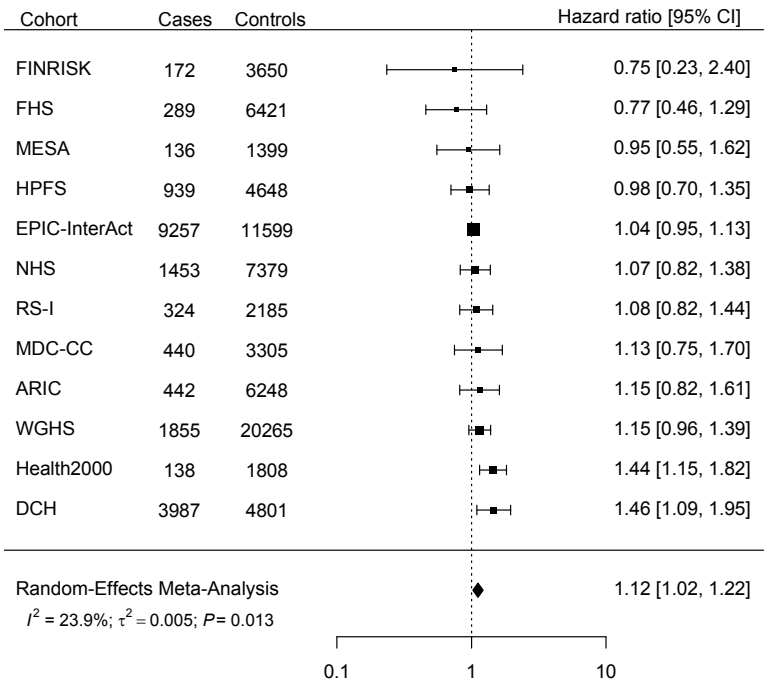
Figure legend: Combined risk of T2D associated with isocaloric replacement (5% energy or 1g/d) of carbohydrate with total or subtypes of fat were estimated by random-effects inverse-variance weighted regression of the pooled log HR on T2D risk.
Each bar represents one study. The blue lines have been fitted with meta-regression.
P-values testing for non-zero slope were; 0.206 for total fat (A); 0.622 for PUFA (B); 0.531 for MUFA (C); 0.316 for SFA (D); 0.735 for total ω-3 PUFA (E); 0.061 for total ω-6 PUFA (F); and 0.554 for total *trans* fat (G).

Supplementary fig 5. Risk of type 2 diabetes associated with isocaloric replacement (5% energy or 1g/d) of carbohydrate with total or subtypes of fat: Sensitivity analysis including cohorts classified as a low risk of bias



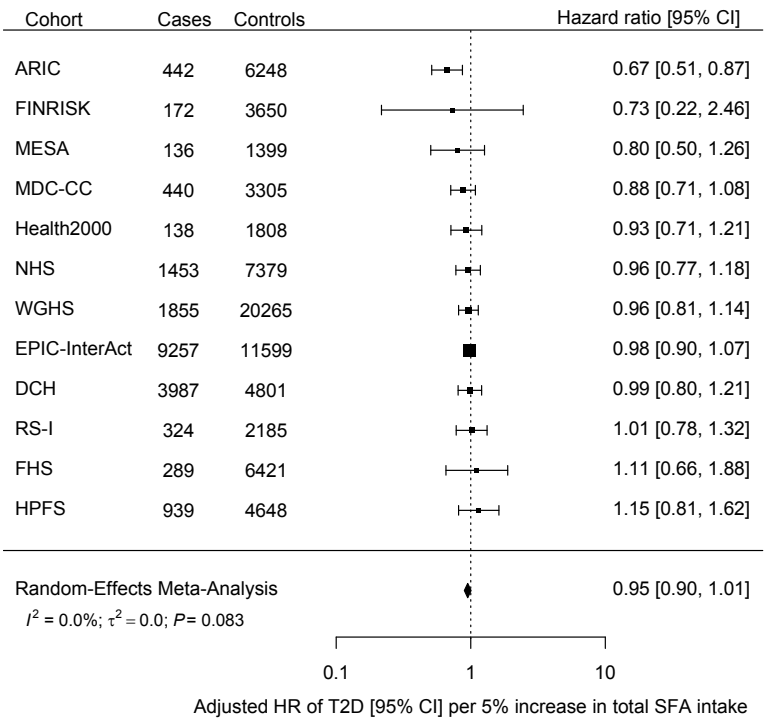
MUFA intake and T2D risk

C



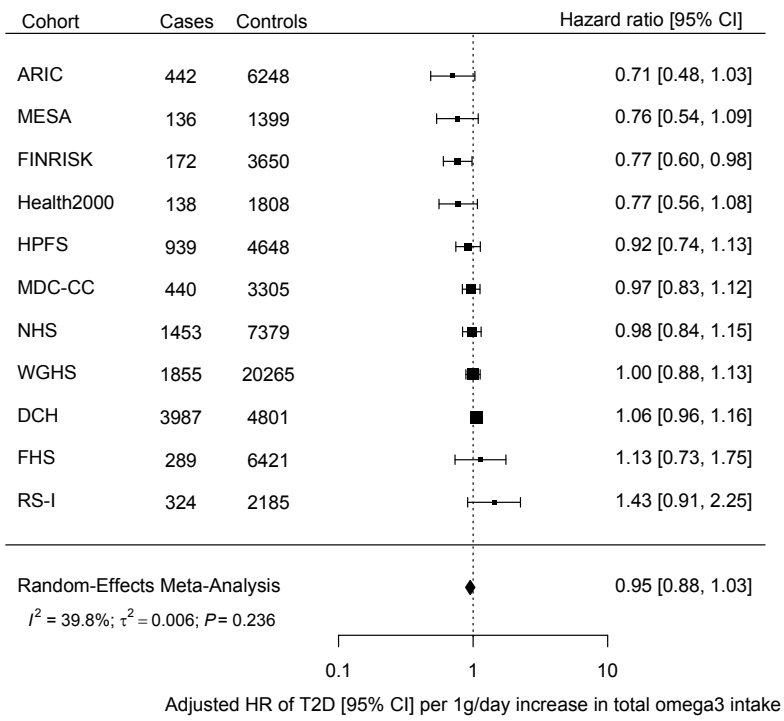
SFA intake and T2D risk

D



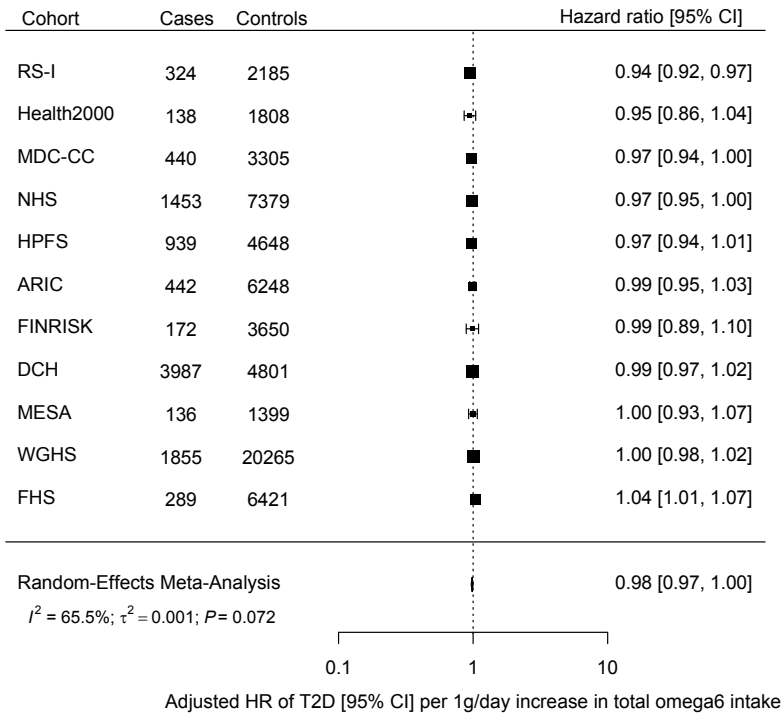
Omega3 intake and T2D risk

E



Omega6 intake and T2D risk

F



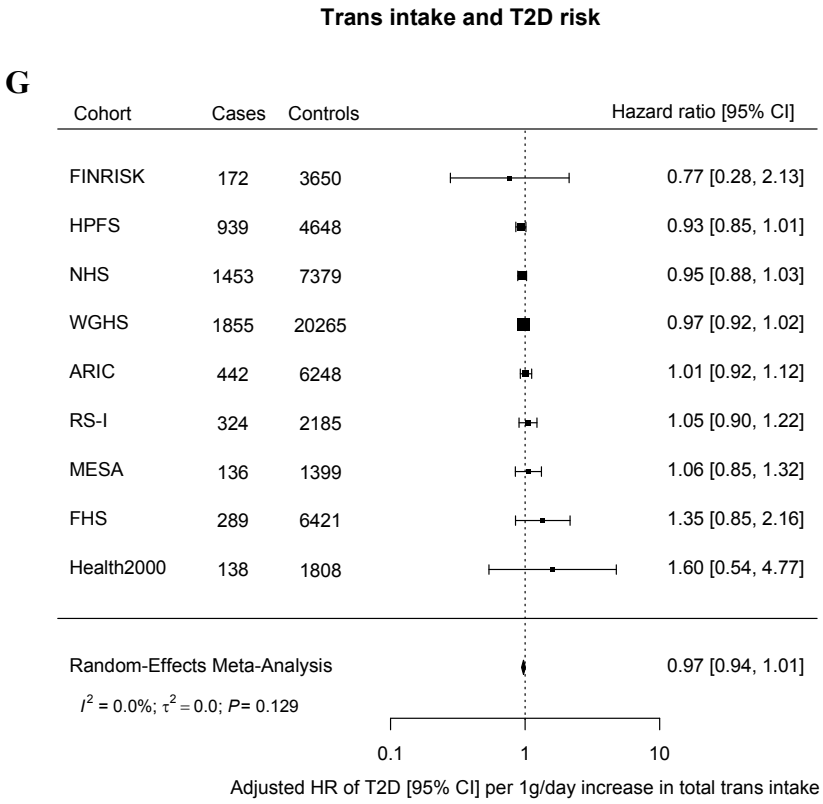
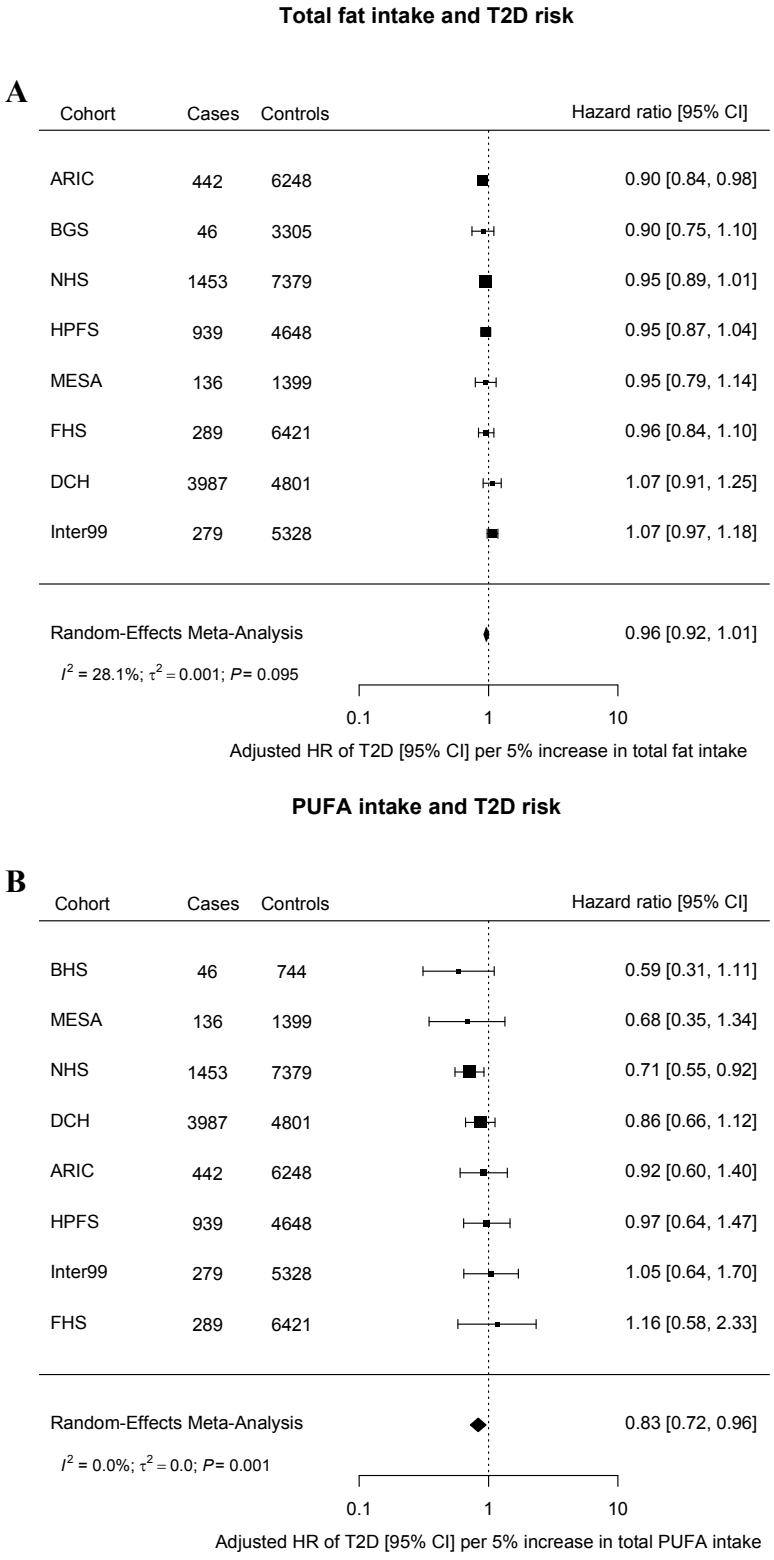


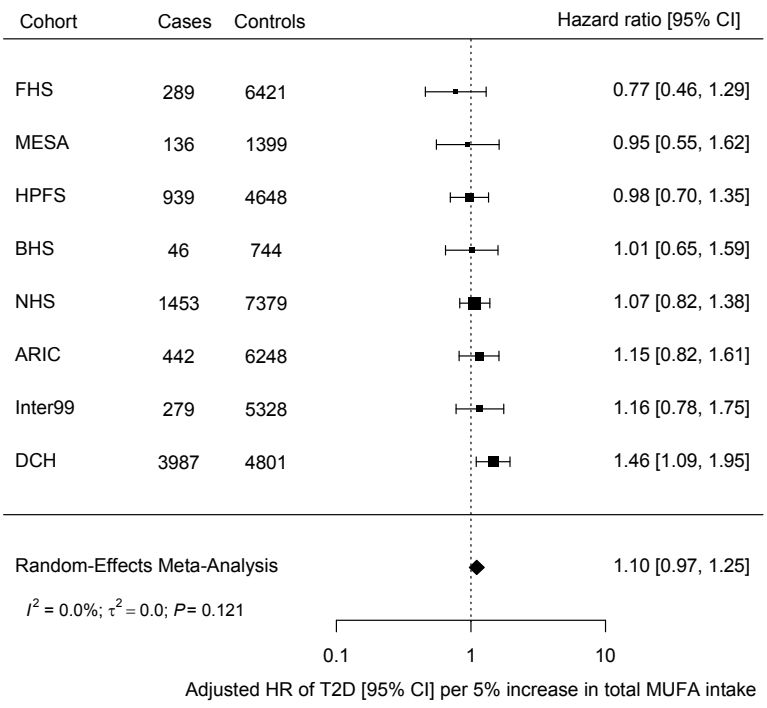
Figure legend: Combined T2D risk estimates in cohorts classified as a lower risk of bias per isocaloric substitution of carbohydrate with total fat (A), polyunsaturated fat (B), monounsaturated fat (C), saturated fat (D), total ω-3 polyunsaturated fat (E), total ω-6 polyunsaturated fat (F), and total *trans* fat (G) on T2D risk (black diamond) using inverse variance-weighted random-effects meta-analysis. Models were adjusted for demographics, lifestyle and clinical characteristics. Also, shown for each cohort the estimate of the association and the 95% confidence interval of the estimate. Cohorts were sorted by increasing T2D risk. Between-study variance (τ^2) was used to assess heterogeneity.

Supplementary fig 6. Risk of type 2 diabetes associated with isocaloric replacement (5% energy or 1g/d) of carbohydrate with total or subtypes of fat: Sensitivity analysis including cohorts with repeated measurements of diet



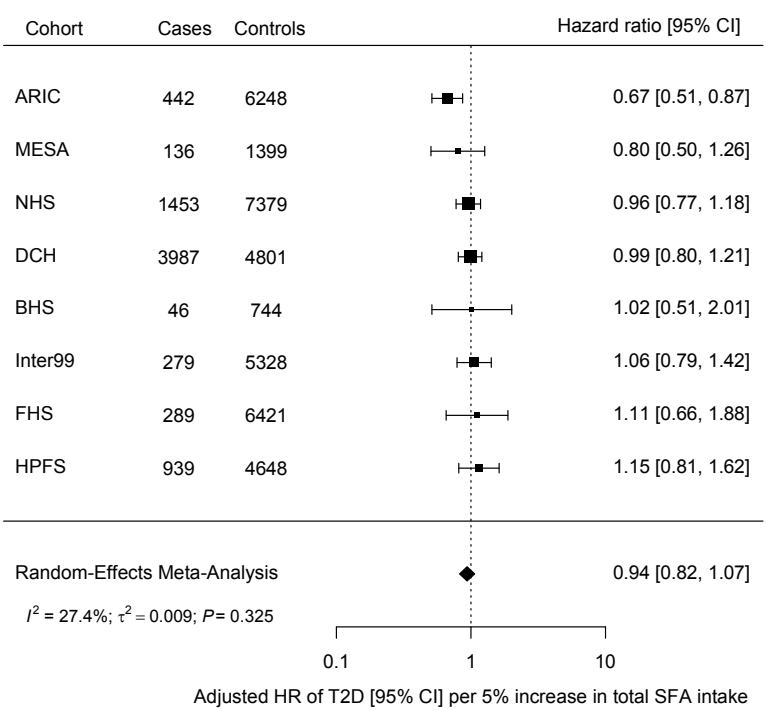
MUFA intake and T2D risk

C



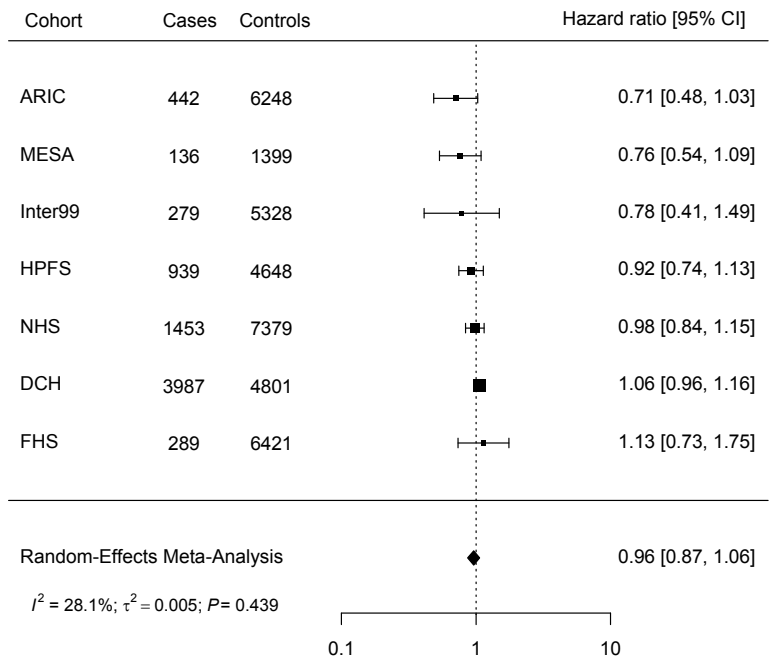
SFA intake and T2D risk

D



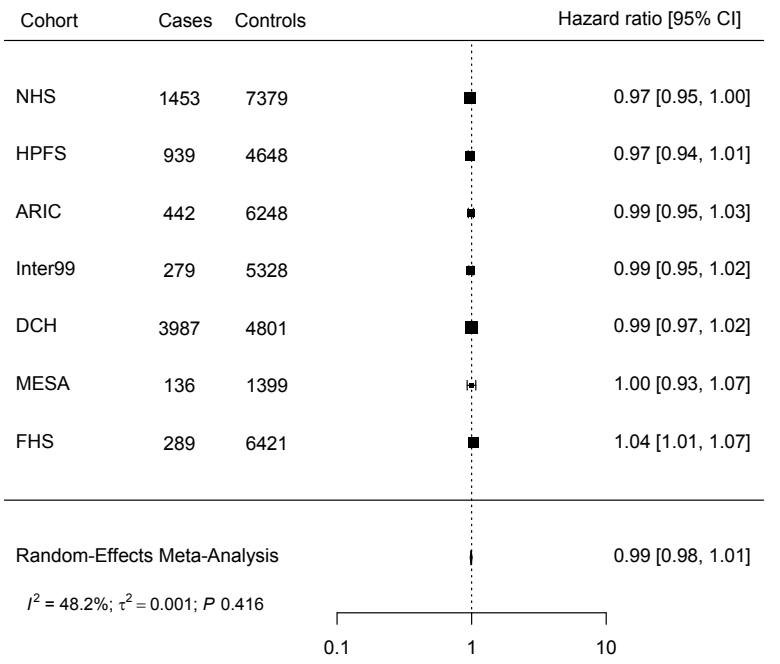
Omega3 intake and T2D risk

E



Omega6 intake and T2D risk

F



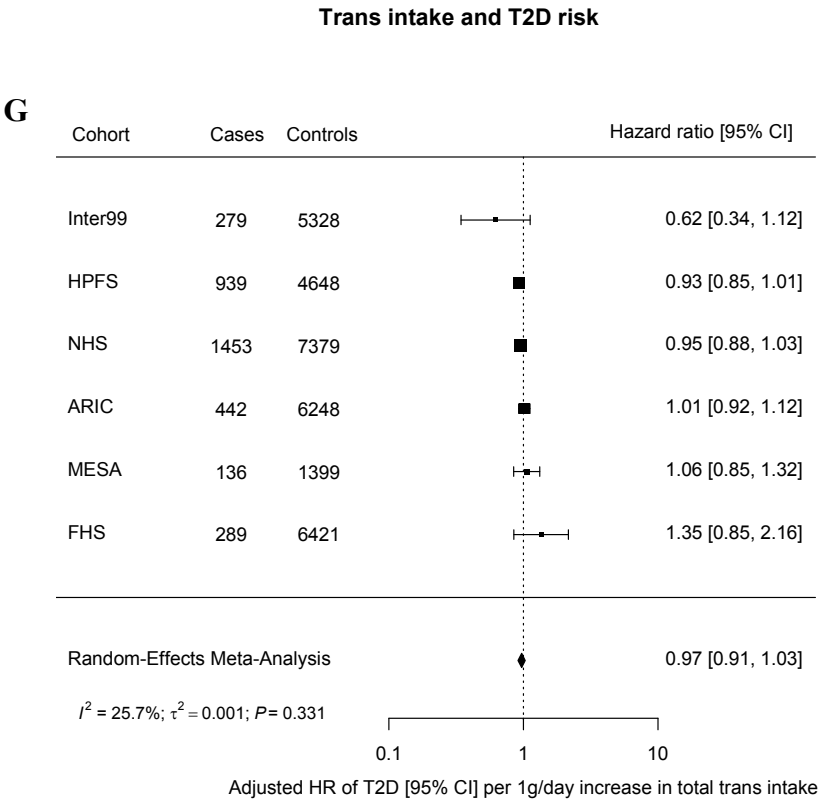


Figure legend: Combined T2D risk estimates in cohorts with repeated measurements of diet per isocaloric substitution of carbohydrate with total fat (A), polyunsaturated fat (B), monounsaturated fat (C), saturated fat (D), total ω-3 polyunsaturated fat (E), total ω-6 polyunsaturated fat (F), and total *trans* fat (G) on T2D risk (black diamond) using inverse variance-weighted random-effects meta-analysis. Models were adjusted for demographics, lifestyle and clinical characteristics. Also, shown for each cohort the estimate of the association and the 95% confidence interval of the estimate. Cohorts were sorted by increasing T2D risk. Between-study variance (τ^2) was used to assess heterogeneity.

Supplementary fig 7. Combined risk of type 2 diabetes with isocaloric replacement (5% energy) of saturated fat with polyunsaturated fat

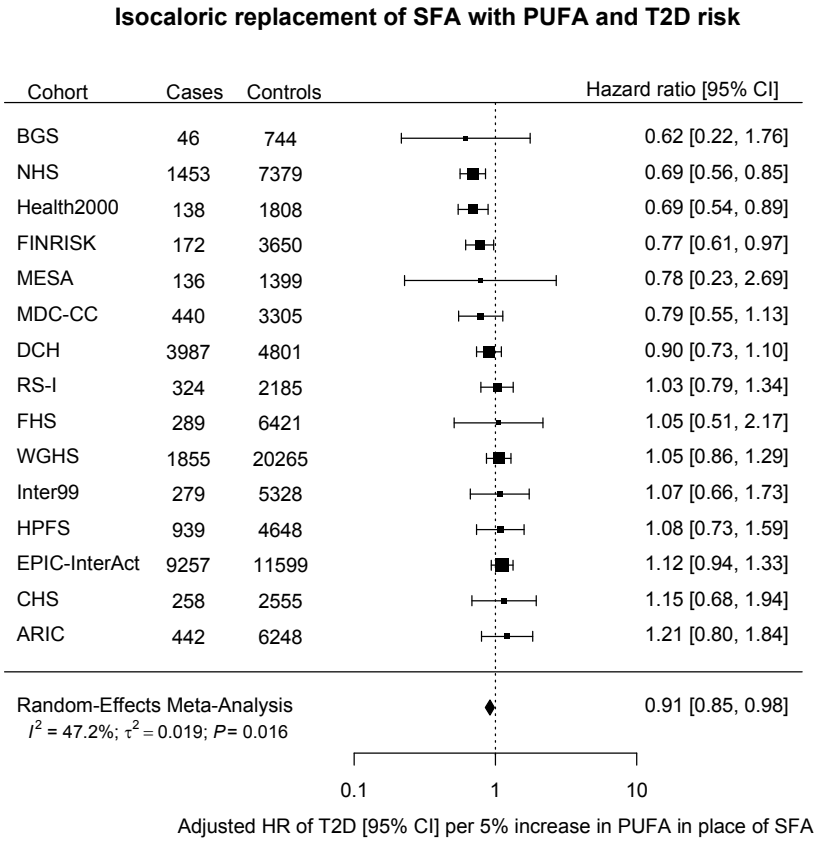


Figure legend: Combined T2D risk estimates per isocaloric substitution of SFA with PUFA (black diamond) using inverse variance-weighted random-effects meta-analysis. Models were adjusted for demographics, lifestyle and clinical characteristics. Also, shown for each cohort the estimate of the association and the 95% confidence interval of the estimate. Cohorts were sorted by increasing T2D risk. Between-study variance (τ^2) was used to assess heterogeneity.

Supplementary fig 8. Interaction between total and subtypes of fat intake and polygenic risk score on the risk of type 2 diabetes

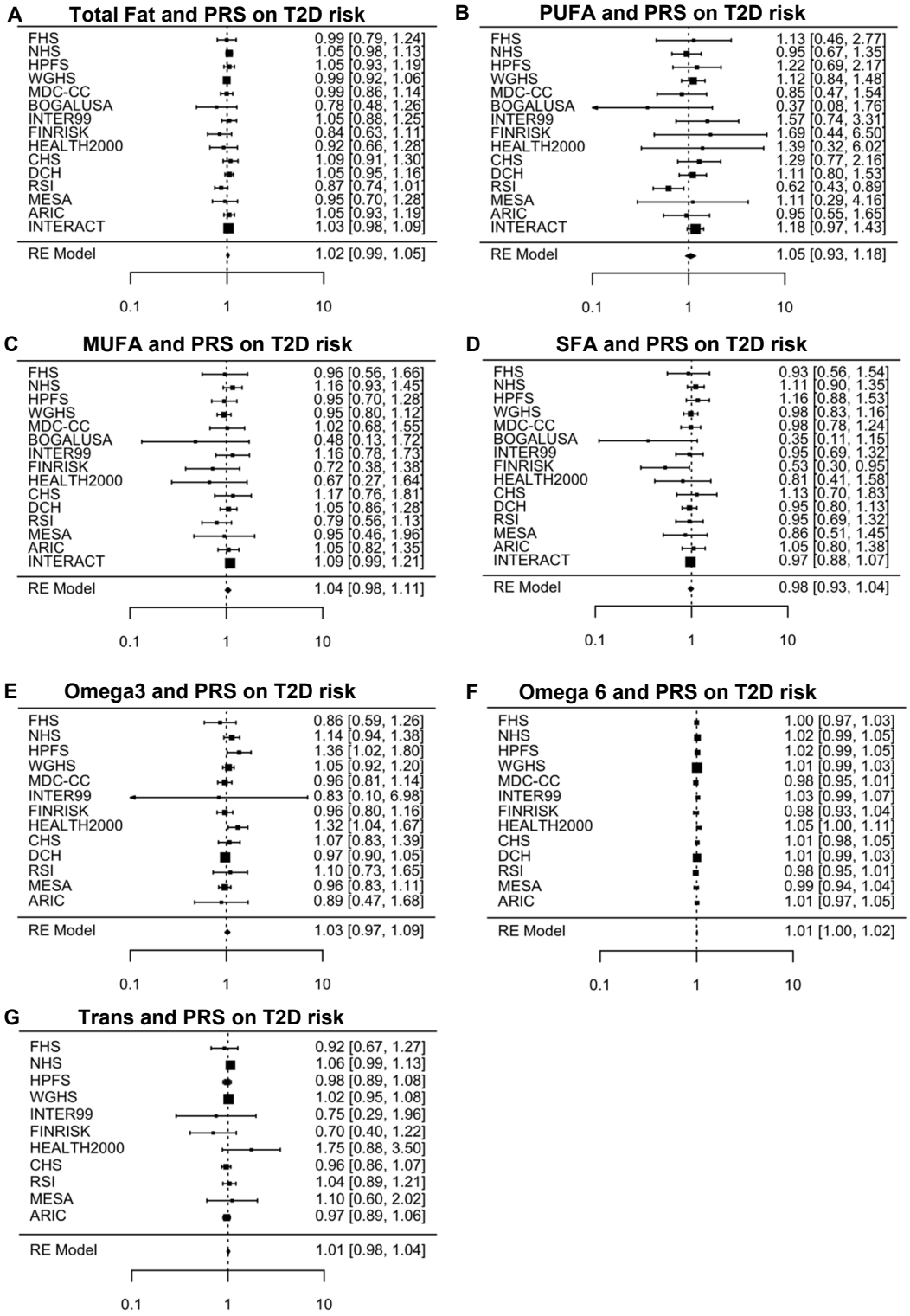


Figure legend: Combined estimates from random-effects inverse-variance weighted meta-analysis represent an interaction effect between dietary fat subtypes and PRS on the risk of T2D per increment of 10 risk alleles in the PRS and the isocaloric replacement for each type of fat: Isocaloric replacement of carbohydrate with A) total fat

(HR= 1.02, 0.99 to 1.05, P=0.200; I^2=0%, τ^2=0), B) PUFA (HR= 1.05, 0.93 to 1.18, P=0.200; I^2=6.1%, τ^2=0.003), C) MUFA (HR= 1.04, 0.98 to 1.11, P=0.218; I^2=0%, τ^2=0), D) SFA (HR= 0.98, 0.93 to 1.04, P=0.584; I^2=0%, τ^2=0), E) total ω-3 PUFA (HR= 1.03, 0.97 to 1.09, P=0.404; I^2=14.2%, τ^2=0.002), F) ω-6 PUFA (HR= 1.01, 1.00 to 1.02, P=0.125; I^2=15.7%, τ^2=0), and G) *trans* fat (HR= 1.01, 0.98 to 1.04, P=0.562; I^2=0%, τ^2=0). Models were adjusted for demographics, lifestyle and clinical characteristics. Also, shown for each cohort the estimate of the association and the 95% confidence interval of the estimate.

Appendix 1: Pre-specified analysis plan

1. General methodological considerations

Participating prospective cohort studies that had agreed to collaborate in this effort are listed below.

Main exclusion: Participants of non-European ancestry with prevalent diabetes, cancer or cardiovascular disease at baseline will be excluded. In addition, individuals who reported implausible baseline energy intake (<500 or >4.500 kcal/d), participants of ages ≥20 years and ≤ 80 years at baseline, or participants who had missing genome-wide genetic data or low call rate for genotyping.

1.1 Primary outcome and exposures

Outcome:

Incident T2D;

Definition: ≥7mmol/l (126mg/dl), random plasma glucose ≥11.1mmol/L (200mg/dl), being on diabetes medications or self-reported diagnosis.

Exposures:

Genetic risk score:

We will build a PRS using summary-level statistics of T2D genetic risk variants from the largest published DIAGRAM consortium (n=89). We will harmonize the PRS among cohorts to create an aggregate weighted PRS for T2D prediction based on the assumption of an additive genetic effect, by assigning 1 point for each risk allele (low-risk homozygotes = 0 points; heterozygotes = 1 point; high-risk homozygotes = 2 points). Each study participant will be assigned a quantitative PRS based on the number of risk alleles and their β-estimates present at the SNPs under investigation (see supplementary PRS formula). To build the PRS we will use successfully genotyped variants from each participating cohort with genotyping success rate ≥ 0.95 and Hardy-Weinberg equilibrium > 1×10^{-4}. When not directly genotyped, we will use variants that passed the quality control (imputation quality > 0.7 and MAF >0.001).

Dietary fat quality:

Dietary fat variables are defined as a cumulative averages of total dietary fat intake and the following subtypes of fat when possible. If not baseline values will be used.

1. Total dietary fat (Percentage of energy intake, continuous)
2. Subtypes of dietary fat
    a. PUFA (Percentage of energy intake, continuous)
    b. MUFA (Percentage of energy intake, continuous)
    c. SFA (Percentage of energy intake, continuous)
    d. Total ω-3 PUFA; g/d; continuous and excluding supplementation
    e. Total *trans* fat; g/d; continuous
    f. Total ω-6 PUFA (g/d; continuous and excluding supplementation). This new exposure has been included posteriori as requested by Reviewers.

1.2 Covariates

Models will be adjusted for 1) age (years, continuous), 2) sex (male, female), 3) body mass index (BMI) (kg/m2, continuous), 4) smoking (never, former or current; if former was not reported then current vs. non-current smoker), 5) physical activity (categorical, quintiles of physical activity), 6) family history of diabetes (yes/no), 7) dyslipidaemia (yes/no, defined based on treatment with lipid-lowering drugs or diagnosis of dyslipidaemia), 8) hypertension (yes/no, defined based on treatment with antihypertensive drugs or diagnosis of hypertension), and 9) dietary factors including total energy intake (kcal/d, continuous), protein intake (% of energy intake, continuous), dietary fibre (g/d, continuous), alcohol intake (g/d, continuous), magnesium (g/d, continuous) and subtypes of fat specific to each statistical model (see below).

1.3 Missing data

As suggested during the revision of the manuscript, missing data were recommended to be handled by multiple imputation. We will use multiple imputation by chained equations method to impute confounders with more than 5% of observations missing in each study. Each missing value will be imputed five times based on their predictive distribution on the observed data. Next for each of the five-imputed data sets the analysis will be repeated, and the five sets of results will be aggregated using Rubin rules. If less than 5% of values were missing, those subjects with missing information were excluded from the analysis.

1.4 Unit of analysis

The following advantages will derive from basing this meta-analysis on cohort-level estimates for the pooled analyses. This approach should enable appropriate pooling of the hazard ratio (HR) estimates across different studies and the reliable quantification of between-study heterogeneity. Hence, to facilitate such analyses (and to enable a uniform approach to analyses across the participating consortia/studies), the Coordinating Center will provide SAS and R software scripts for direct use (or adaptation as needed). For example, given the strong collinearity between the interaction term and both dietary fats variables and PRS, we will use centering methods using the combined mean value of all cohorts as recommended by Aschard H, *et al*. Genet Epidemiol 2016. In addition, to implement the joint meta-meta-analysis method, it is essential to obtain the covariance matrix of the HR estimates within each cohort to present relative risk for the interaction term and main genetic effects.

1.5 Avoidance of "double-counting" data from participating studies

As some studies to be included in the present analysis have provided data on dietary fat and T2D risk to two of the participating consortia, it is important to ensure that data from such studies are not "double- counted" in the present meta-analysis. This will be achieved in consultation with coordinators of the participating consortia. In such situations, the information from the consortium to which the study has provided the maximum relevant data will provide results for that cohort.

1.6 Subgroup analysis

Interactions between dietary fat quality and genetic predisposition on T2D risk depend not only on the amount and quality of dietary fat intake in different regions, but also on other factors that can impact on interaction estimates such age or BMI. Because all of this, there is no reason to expect a priori that the interaction between dietary fat quality and genetic predisposition on T2D risk will be similar in regions that have different underlying fat intake, age or BMI. Therefore, the analysis will be stratified by major geographical regions (i.e., Europe/North America), mean age of each cohort (≥ 55 years; < 55 years) and mean BMI of each cohort (≥ 26 kg/m2; < 26 kg/m^2)

2. Statistical analyses plan

1. Each cohort lists whether information on the T2D risk-increasing variants discovered to date is available either via direct genotyping or via imputation with a sufficient imputation quality. We acknowledge that not all studies have full information on these T2D risk-increasing variants. However, a table providing the specific variants available by each study will clarify the exclusions made in the main analyses.

2. Each cohort will provide basic descriptive statistics (mean (SD), min, max) of main exposures of interest (total dietary fat and subtypes) and indicate whether this information is only available at baseline or not.

3. The following cohort-level characteristics at baseline will be provided:

1. Total number of participants
2. Total number of T2D cases
3. Mean (SD) of PRS
4. Mean (SD) of age
5. Mean (SD) of BMI
6. Number (%) of males.
7. Number (%) of never/former/current tobacco smokers.
8. Mean (SD) of physical activity index.
9. Number (%) of participants with hypertension

10. Number (%) of participants with dyslipidemia
11. Number (%) of participants with parental history of T2D
12. Mean (SD) of dietary factors (energy intake (kcal/d), protein intake (% of energy intake), dietary fibre (g/d), alcohol intake (g/d).
13. Median (IQR) of follow-up time.
14. Country and region (ie, Europe, North America).
15. Method of assessment of diet.
16. Method of genotyping imputation panels, and quality metrics
17. Ascertainment methods for T2D adjudication (i.e., fasting/non-fasting glucose determinations, treatment with either insulin or a hypoglycaemic drugs at follow-up examinations, or by reviewing medical record).

3. Primary analyses based on Cox proportional-hazards models to estimate the relative risk of T2D will be conducted modelling PRS and total or subtypes of fat intake as continuous variables.

4. Interaction models will incorporate an interaction term between dietary fat variables and PRS. Given the strong collinearity between the interaction term and both dietary fats variables and PRS, we will use centering methods using the combined mean value of all cohorts.

5. Stratified analyses by quartiles of PRS and adjusted for the same confounders as before will be conducted

6. Study specific statistics will be meta-analysed. Between-study heterogeneity will be quantified by the I^2 statistic. In addition, we will exclude one cohort at a time to identify single-cohort drive effects.

7. Sensitivity analyses will be conducted including only cohorts with repeated measurements of diet and cohorts classified as a low risk of bias.

8. Meta-regression analyses will be conducted to examine the impact of moderator variables on interaction estimates including geographical region (Europe; North America), mean age of each cohort (≥55 years; <55 years) and mean BMI of each cohort (≥ 26 kg/m2; <2 6 kg/m2). Additional stratified meta-analyses will be conducted when the effect of moderator variables on meta-analysed interaction estimates is significant.

9. The joint meta-analysis method will be used as a secondary analysis to combine study specific interaction statistics.

10. If required, studies will be contacted for further information or analyses.

3. Data checking

After receiving study specific statistics, data quality control procedures will be conducted to identify extreme outliers for main exposures and covariates, missing information, identification of disparate estimates, and revision of the scripts used by each analyst.

4. List of cohorts that had agreed to collaborate in this effort

Atherosclerosis Risk in Communities study (ARIC), USA. The Bogalusa Heart Study (BHS), USA. The Cardiovascular Health Study (CHS), USA. The Danish Diet, Cancer and Health study (DCH), Denmark. The European Prospective Investigation of Cancer-InterAct (EPIC-InterAct), Europe (10-countries). The Framingham Heart Study (FHS), USA. The National FINRISK study (FINRISK), Finland. The Finish Health 2000 Study (Health 2000), Finland. The Health Professionals' Follow-up Study (HPFS), USA. The Inter99 study (Inter99), Denmark. The Malmö Diet and Cancer-Cardiovascular Cohort study (MDC-CC), Sweden. The Multi-Ethnic Study of Atherosclerosis (MESA), USA. The Nurses' Health Study (NHS), USA. The Rotterdam Study I (RS-I) the Netherlands. The Women's Genome Health Study (WGHS), USA.

(list in alphabetical order, updated September 2017).

Appendix 2: PRISMA-IPD checklist

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Title			
Title	1	Identify the report as a systematic review and meta-analysis of individual participant data.	1, lines 1-2
Abstract			
Structured summary	2	Provide a structured summary including as applicable:	8,9 lines 162-193
		Background: state research question and main objectives, with information on participants, interventions, comparators and outcomes.	
		Methods: report eligibility criteria; data sources including dates of last bibliographic search or elicitation, noting that IPD were sought; methods of assessing risk of bias.	
		Results: provide number and type of studies and participants identified and number (%) obtained; summary effect estimates for main outcomes (benefits and harms) with confidence intervals and measures of statistical heterogeneity. Describe the direction and size of summary effects in terms meaningful to those who would put findings into practice.	
		Discussion: state main strengths and limitations of the evidence, general interpretation of the results and any important implications.	
		Other: report primary funding source, registration number and registry name for the systematic review and IPD meta-analysis.	
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known.	10-11, lines 208-230
Objectives	4	Provide an explicit statement of the questions being addressed with reference, as applicable, to participants, interventions, comparisons, outcomes and study design (PICOS). Include any hypotheses that relate to particular types of participant-level subgroups.	11, lines 231-234
Methods			
Protocol and registration	5	Indicate if a protocol exists and where it can be accessed. If available, provide registration information including registration number and registry name. Provide publication details, if applicable.	12, lines 255-256 appendix 1
Eligibility criteria	6	Specify inclusion and exclusion criteria including those relating to participants, interventions, comparisons, outcomes, study design and characteristics (e.g. years when conducted, required minimum follow-up). Note whether these were applied at the study or individual level i.e. whether eligible participants were included (and ineligible participants excluded) from a study that included a wider population than specified by the review inclusion criteria. The rationale for criteria should be stated.	12-13, lines 258-279 appendix 1 appendix 3
Identifying studies - information sources	7	Describe all methods of identifying published and unpublished studies including, as applicable: which bibliographic databases were searched with dates of coverage; details of any hand searching including of conference proceedings; use of study registers and agency or company databases; contact with the original research team and experts in the field; open adverts and surveys. Give the date of last search or elicitation.	12, lines 264-271 appendix 3

Identifying studies - search	8	Present the full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	12, lines 259-264 appendix 3
Study selection processes	9	State the process for determining which studies were eligible for inclusion.	11, lines 264-268 appendix 3
Data collection processes	10	Describe how IPD were requested, collected and managed, including any processes for querying and confirming data with investigators. If IPD were not sought from any eligible study, the reason for this should be stated (for each such study). If applicable, describe how any studies for which IPD were not available were dealt with. This should include whether, how and what aggregate data were sought or extracted from study reports and publications (such as extracting data independently in duplicate) and any processes for obtaining and confirming these data with investigators.	12-13, lines 271-273 15, lines 330-333 appendix 1
Data items	11	Describe how the information and variables to be collected were chosen. List and define all study level and participant level data that were sought, including baseline and follow-up information. If applicable, describe methods of standardising or translating variables within the IPD datasets to ensure common scales or measurements across studies.	13-15, lines 282-324 appendix 1
IPD integrity	A1	Describe what aspects of IPD were subject to data checking (such as sequence generation, data consistency and completeness, baseline imbalance) and how this was done.	15, lines 326-333 appendix 1
Risk of bias assessment in individual studies.	12	Describe methods used to assess risk of bias in the individual studies and whether this was applied separately for each outcome. If applicable, describe how findings of IPD checking were used to inform the assessment. Report if and how risk of bias assessment was used in any data synthesis.	15, lines 326-328 appendix 1
Specification of outcomes and effect measures	13	State all treatment comparisons of interests. State all outcomes addressed and define them in detail. State whether they were pre-specified for the review and, if applicable, whether they were primary/main or secondary/additional outcomes. Give the principal measures of effect (such as risk ratio, hazard ratio, difference in means) used for each outcome.	13, lines 282-287 appendix 1
Synthesis methods	14	Describe the meta-analysis methods used to synthesise IPD. Specify any statistical methods and models used. Issues should include (but are not restricted to): • Use of a one-stage or two-stage approach. • How effect estimates were generated separately within each study and combined across studies (where applicable). • Specification of one-stage models (where applicable) including how clustering of patients within studies was accounted for. • Use of fixed or random effects models and any other model assumptions, such as proportional hazards. • How (summary) survival curves were generated (where applicable). • Methods for quantifying statistical heterogeneity (such as I^2 and τ^2).	15-18, lines 335-394 appendix 1

		• How studies providing IPD and not providing IPD were analysed together (where applicable). • How missing data within the IPD were dealt with (where applicable).	
Exploration of variation in effects	A2	If applicable, describe any methods used to explore variation in effects by study or participant level characteristics (such as estimation of interactions between effect and covariates). State all participant-level characteristics that were analysed as potential effect modifiers, and whether these were pre-specified.	15-18, lines 335-394 appendix 1
Risk of bias across studies	15	Specify any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to not obtaining IPD for particular studies, outcomes or other variables.	17, lines 372-382 appendix 1
Additional analyses	16	Describe methods of any additional analyses, including sensitivity analyses. State which of these were pre-specified.	16-17, lines 362-364 17-18, lines 373-390
Results			
Study selection and IPD obtained	17	Give numbers of studies screened, assessed for eligibility, and included in the systematic review with reasons for exclusions at each stage. Indicate the number of studies and participants for which IPD were sought and for which IPD were obtained. For those studies where IPD were not available, give the numbers of studies and participants for which aggregate data were available. Report reasons for non-availability of IPD. Include a flow diagram.	19, lines 408-410 appendix 3 supplementary fig 1
Study characteristics	18	For each study, present information on key study and participant characteristics (such as description of interventions, numbers of participants, demographic data, unavailability of outcomes, funding source, and if applicable duration of follow-up). Provide (main) citations for each study. Where applicable, also report similar study characteristics for any studies not providing IPD.	table 1, table 2, supplemental tables 2,3,5,6,7 appendix 4
IPD integrity	A3	Report any important issues identified in checking IPD or state that there were none.	19, lines 413-414
Risk of bias within studies	19	Present data on risk of bias assessments. If applicable, describe whether data checking led to the up-weighting or down-weighting of these assessments. Consider how any potential bias impacts on the robustness of meta-analysis conclusions.	19, lines 410-413 20, lines 447-451 supplementary table 6
Results of individual studies	20	For each comparison and for each main outcome (benefit or harm), for each individual study report the number of eligible participants for which data were obtained and show simple summary data for each intervention group (including, where applicable, the number of events), effect estimates and confidence intervals. These may be tabulated or included on a forest plot.	fig 1, fig2, supplemental figs 2-7 table 3
Results of syntheses	21	Present summary effects for each meta-analysis undertaken, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified, and report the numbers of studies and participants and, where applicable, the number of events on which it is based.	19-22, lines 425-480
		When exploring variation in effects due to patient or study characteristics, present summary interaction estimates for each	

		characteristic examined, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified. State whether any interaction is consistent across trials.	
		Provide a description of the direction and size of effect in terms meaningful to those who would put findings into practice.	
Risk of bias across studies	22	Present results of any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to the availability and representativeness of available studies, outcomes or other variables.	20, lines 446-450 21, lines 464-465
Additional analyses	23	Give results of any additional analyses (e.g. sensitivity analyses). If applicable, this should also include any analyses that incorporate aggregate data for studies that do not have IPD. If applicable, summarise the main meta-analysis results following the inclusion or exclusion of studies for which IPD were not available.	20-21, lines 451-452 21, lines 466-468
Discussion			
Summary of evidence	24	Summarise the main findings, including the strength of evidence for each main outcome.	23, lines 496-501
Strengths and limitations	25	Discuss any important strengths and limitations of the evidence including the benefits of access to IPD and any limitations arising from IPD that were not available.	25-27, lines 563-597
Conclusions	26	Provide a general interpretation of the findings in the context of other evidence.	27, lines 599-606
Implications	A4	Consider relevance to key groups (such as policy makers, service providers and service users). Consider implications for future research.	27, lines 602-606
Funding			
Funding	27	Describe sources of funding and other support (such as supply of IPD), and the role in the systematic review of those providing such support.	28, line 627 appendix 5

Appendix 3: Systematic literature search procedures

Data sources

Studies published between January 1970 and February 2017 were identified, without any language restriction, through electronic searches using MEDLINE, EMBASE, and SCOPUS and discussion with investigators. The computer-based search strategy is detailed below. Upon identification of studies and eligible independent cohorts, 19 cohorts were asked to participate in a standardized individual participant data analysis of dietary fat quality, genetic risk, and T2D incidence by March 2017. Four eligible cohorts were unable to contribute due to infrastructure constrains (n=3) and methodological reasons (n=1).

Study selection

Prospective cohort studies were included if they had available genome-wide genetic information and they had reported on the association of diet or dietary fat quality with T2D incidence. Prospective cohort studies or multi-cohort consortia were eligible for inclusion if they satisfied all of the following criteria: 1) had genome-wide genetic data and information about dietary fat quality and T2D incidence; 2) included >500 European ancestry participants not selected on the basis of having any previous chronic disease; 3) recruited adults (ages ≥20 years and ≤ 80 years at baseline), and 4) had accrued 5 years or more of median follow-up. A literature search flow chart is provided below.

Data on the following characteristics were extracted independently by two investigators (JM, HSD) according to a pre-specified protocol: full study name, year of publication, prospective cohort study, study location, and number of participants. Discrepancies were resolved by discussion and by adjudication of a third reviewer (CES).

Search strategy

Publication database: MEDLINE

#1	"diet" OR "dietary fat" OR "fat quality" OR "fat intake"
#2	"genetics" OR "genotype" OR "gene"
#3	"diabetes"
#4	"Cohort Studies" [Mesh] OR "cohort" OR "prospective" OR "risk ratio" OR "relative risk" OR "hazard ratio" OR "risk ratios" OR "relative risks" OR "hazard ratios"
Search	#1 AND #2 AND #3 AND #4
Restrictions	Clinical Trial
# identified studies	490

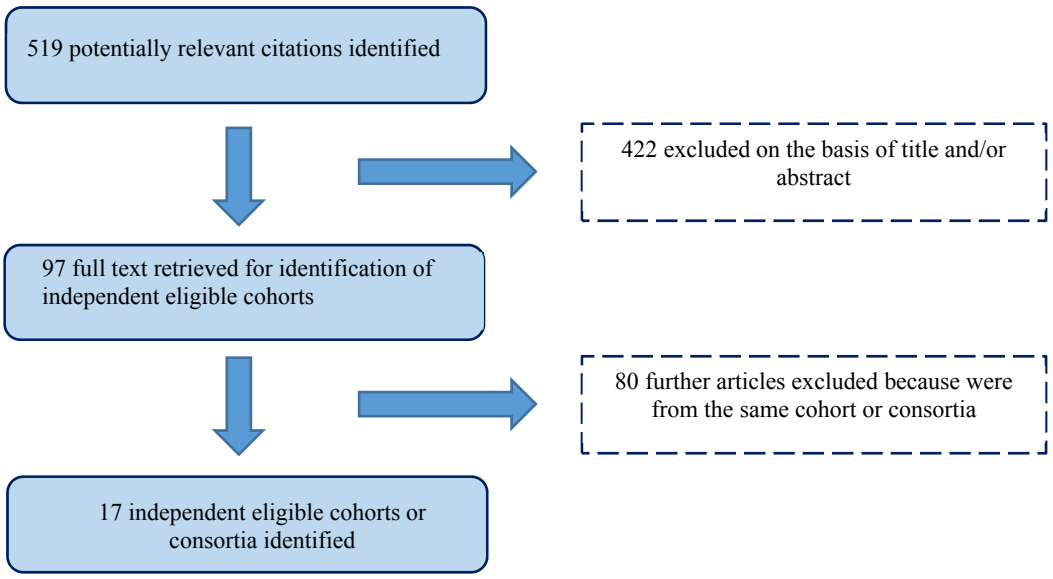
Publication database: EMBASE

#1	(diet OR dietary fat OR fat quality OR fat intake).af
#2	(genetics OR genotype OR gene).af
#3	(diabetes).af
#4	(cohort OR prospective OR risk ratio OR relative risk OR hazard ratio OR risk ratios OR relative risks OR hazard ratios).af
Search	#1 AND #2 AND #3 AND #4
Restrictions	Clinical Trial
# identified studies	424

Publication database: SCOPUS

#1	TITLE-ABS-KEY (diet OR dietary fat OR fat quality OR fat intake)
#2	TITLE-ABS-KEY (genetics OR genotype OR gene)
#3	TITLE-ABS-KEY (diabetes)
#4	TITLE-ABS-KEY (cohort OR prospective OR risk ratio OR relative risk OR hazard ratio OR risk ratios OR relative risks OR hazard ratios)
Search	#1 AND #2 AND #3 AND #4
Restrictions	Clinical Trial
# identified studies	407

Literature search flow chart



Appendix 4: Description of the participating cohorts and outcome ascertainment

Participants for the current IPD meta-analysis were drawn from 15 prospective cohort studies, including the Atherosclerosis RIsk in Communities study (ARIC), the Bogalusa Heart Study (Bogalusa), the Cardiovascular Health Study (CHS), the Danish Diet, Cancer and Health study (DCH), the European Prospective Investigation of Cancer-InterAct (EPIC-InterAct), the National FINRISK study (FINRISK), the Framingham Heart Study (FHS), the Health Professionals' Follow-up Study (HPFS), the Health 2000 Survey (Health 2000), The Inter99 study (Inter99), the Malmö Diet and Cancer-Cardiovascular Cohort study (MDC-CC), the Multi-Ethnic Study of Atherosclerosis (MESA), the Nurses' Health Study (NHS), the Rotterdam Study I (RS-1) and the Women's Genome Health Study (WGHS). Two European studies were case-cohort studies; EPIC-InterAct, which included sample from a total of 340,234 participants in the EPIC-Europe cohort study and DCH that selected a sample from a total of 57,053 free cancer participants living in Denmark. For this analysis, overlapping participants between EPIC-InterAct, DCH and MDC-CC were excluded from DCH and MDC-CC respectively (detailed below).

The **Atherosclerosis Risk in Communities** (ARIC) study is a population-based cohort study designed to study new and established risk factors for atherosclerosis and community trends in coronary heart disease. In 1987-89, baseline data was collected on 15,792 adults, aged 45–64 y, living in four U.S. communities (Forsyth County, NC; Jackson, MS; northwest Minneapolis suburbs, MN; Washington County, MD). Follow-up examinations including demographic, lifestyle and clinical determinations were conducted in approximate 3-year intervals. Up to 6,690, European ancestry adults with available DNA, valid dietary information, and consent to share genetic data were eligible for the current analysis.
The ARIC study ascertained T2D incidence in follow-up examinations (visit 3, 1993-5) and defined using published criteria as a fasting glucose level ≥7mmol/l (126 mg/dl), a random plasma glucose level ≥11.1mmol/l (200 mg/dl), or a history of or treatment for diabetes.

Study design reference: The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol. 1989; 129:687-702.

The **Bogalusa Heart Study** (BHS), which began in 1973, is a long-term epidemiologic study designed to examine the natural history of cardiovascular risk factors from childhood to adulthood in a well-defined biracial population (65 percent European descendent and 35 percent African American) community in Bogalusa, Louisiana. Between 1973 and 2010, nine cross-sectional surveys of children ages 4 to 18 years and ten cross-sectional surveys of adults, aged 19 to 52 years, who had been previously examined as children were conducted in Bogalusa. This panel design of repeated cross-sectional examinations has resulted in serial observations every 2 to 3 years from childhood to adulthood. The longitudinal cohort of this study consisted of 790 adult subjects who had confirmed European ancestry and had valid dietary information and consent to share genetic data.
Based on the American Diabetes Association criteria, ascertainment of T2D was defined as having fasting plasma glucose ≥7mmol/l or a history of or treatment for diabetes.
Study design reference: Berenson GS, et al. Cardiovascular disease risk factor variables at the preschool age. The Bogalusa heart study. Circulation. 1978; 57:603-12.

The **Cardiovascular Health Study** (CHS) is a population-based prospective cohort study of cardiovascular disease in adults older than 65 years, and includes 5,888 participants ≥ 65 years of age identified from four U.S. communities using Medicare eligibility lists (Forsyth County, NC; Sacramento County, CA; Washington County, MD; Pittsburgh, PA). The original cohort included 5,201 participants recruited in 1989–1990 and 687 additional subjects were recruited in 1992–1993 to enhance the racial/ethnic diversity of the cohort. The CHS genome-wide association study (GWAS), which had the primary aim of studying incident cardiovascular events, included 3,980 CHS participants who were free of clinical cardiovascular disease at study baseline, consented to genetic testing, and had DNA available for genotyping. For this analysis, we included 2,813 Caucasian individuals free of diabetes and clinical cardiovascular disease at baseline with available lifestyle, genetic and dietary information.
In the CHS, participants were classified as having new-onset T2D based on the initiation of insulin or oral hypoglycemic therapy or having fasting glucose level ≥7.0mmol/L.
Study design reference: The Cardiovascular Health Study: design and rationale. Ann Epidemiol. 1991; 1:263-76.

The **Danish Diet, Cancer and Health study (**DCH) study cohort was established between 1993 and 1997 with the primary objective to prospectively investigate the etiologic role of diet and lifestyle in the development of cancer in

57,053 participants (27,178 men and 29,876 women). The study participants were aged between 50 and 64 years at baseline, lived in the urban areas of Copenhagen and Aarhus, and did not have a cancer diagnosis registered in the Danish Cancer Registry at baseline. Participants were, for this study, followed from 1993-1997 until the end of December 2011. The DCH Study is part of the European Prospective Investigation into Nutrition and Cancer (EPIC). The case-cohort sample of 8,788 participants included in the present analyses does not overlap with the case-cohort sample genotyped in the EPIC-InterAct project.

The DCH study participants were followed up in the National Diabetes Register for incident diagnoses of diabetes. A person is included in the National Diabetes Register if one of the following criteria is met: registration in National Prescription Registry with a diagnosis of diabetes; registration of chiropody (as diabetic patient) in the National Health Insurance Service Register; five blood glucose measurements within one year or two blood glucose measurements per year in five consecutive years in the National Health Insurance Service Register; or registration in The Drug Prescription Register either with purchase of oral glucose lowering drugs within 6 months or prescribed insulin.
Study design reference: Tjonneland A, et al. Study design, exposure variables, and socioeconomic determinants of participation in diet, cancer and health: A population-based prospective cohort study of 57,053 men and women in Denmark. Scand. J. Public Health. 2007; 35:432–41.

The large prospective **EPIC-InterAct** type 2 diabetes case–cohort study was nested within the EPIC study, one of the largest cohort studies in the world. EPIC was initiated in the late 1980s and involves collaboration between 23 research institutions across Europe in ten countries (Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden and the UK). InterAct consortium partners ascertained and verified 12,403 T2D cases and selected a random subcohort of 16,835 individuals with baseline plasma samples occurring among 340,234 persons with 3.99 million person-years of follow-up (1991–2007) in eight out of ten countries of the EPIC study. After exclusions (n = 5,287 (19%) individuals without genetics data, and a further 592 (2.1%) individuals with extreme values of energy intake), the sample for this analysis included 21,900 individuals: 12,749 subcohort members and 9,742 individuals with incident T2D (591 of these cases were also in the subcohort). The median follow-up was 10.9 years. Lifestyle and dietary information was collected at baseline.
Incident T2D was ascertained until 31 December 2007 by reviewing multiple sources of evidence, including self-report, linkage to primary-care registers and secondary-care registers, medication use, hospital admissions, and mortality data. No diabetes cases were ascertained solely by self-report, and further evidence was sought for cases with information on incident T2D from fewer than two independent sources.
Study design reference: InterAct Consortium. Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. Diabetologia. 2011; 54:2272-8.

The **Framingham Heart Study** (FHS) is a community-based longitudinal study designed to examine CVD risk in the offspring of the Original Cohort participants of the Framingham Heart Study and their spouses. In 1971, 5,124 individuals were enrolled; since then, the Offspring Cohort has been examined every 3–4 y. between 1998 and 2001, during the 7th examination cycle, 3,539 adults, with a mean age of 61y, underwent a standardized medical history and physical examination. Beginning in 2002, 4,095 Gen III participants, who had at least one parent in the offspring cohort, were enrolled in the Framingham Heart Study. At the first cycle of the Gen III study, 4,095 individuals with a mean age of 40 y, underwent the standard clinic examination. For the present study both cohorts were combined for the analysis. A total of 6,710 adults with available DNA, valid dietary information, and consent to share genetic data were eligible for the current study.
In FHS, incident diabetes was defined as a fasting plasma glucose ≥7.0mmol/L or or treatment with either insulin or a hypoglycemic agent at the follow-up examinations. Chart review was conducted to identify participants with type 1 diabetes mellitus; those individuals were excluded from the analyses.
Study design references: Dawber TR, et al. An approach to longitudinal studies in a community: the Framingham Study. Ann N Y Acad Sci. 1963; 107:539-56.
Feinleib M, et al. The Framingham Offspring Study. Design and preliminary data. Prev Med. 1975; 4:518-25.
Splansky GL, et al. The Third-Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. Am J Epidemiol. 2007; 165:1328-35.

The National **FINRISK Study** consisted of population surveys on the risk factors of chronic non-communicable diseases. Participants included in this study took part in two phases of the National FINRISK 2007 Study conducted

by the National Institute for Health and Welfare in Finland. An independent random sample of 10,000 men and women aged 25–74 years was drawn from the national population register in five large geographical areas at the end of 2006. The sample was stratified by sex, 10-year age category and area. The first phase took place between January and March 2007 including a health examination (including measurements of weight, height, waist circumference, hip circumference and blood pressure as well as collection of blood samples) and health questionnaires on sociodemographic factors, health behaviour (e.g. leisure-time physical activity, smoking habits and alcohol intake) and medical history. All 6,258 participants who took part in the first phase of the survey were invited to a more detailed examination on the dietary, lifestyle and genetic determinants of obesity and the metabolic syndrome from April to June 2007. This phase included measurements of anthropometric parameters (weight, waist circumference, hip circumference and body fat percentage), oral glucose tolerance test and a self-administered FFQ. The response rate for the second phase was 80 %; thus, 5,024 participated in the study. For this study, we included data from 3,822 men and 2570 women (48 % of those invited to participate).
T2D ascertainment was based on the Hospital Discharge Register and Causes of Death Register: E10-E14 (ICD-10) / 250 (ICD-8/9) as main diagnosis symptom, cause and underlying cause of death. E10, E11, E14 (ICD-10) / 250 (ICD-8/9) as the first, the second and the third side diagnoses (symptoms) the first and the second side diagnoses (cause) or as immediate cause of death, the first, the second and the third contributing cause of death. In addition, T2D was ascertainment based on the National Social Insurance Institution (KELA) prescribed medication register (ATC class A10 / A10A / A10B) and KELA specially reimbursed medication register (code 103).
Study design reference: Borodulin K, et al. Forty-year trends in cardiovascular risk factors in Finland. Eur J Public Health. 2015; 3: 539–46.

The study population was derived from the **Health Professionals' Follow-up study** HPFS, a prospective cohort of 51,529 male health care professionals in the United States aged 40–75 years at enrolment in 1986. The present analysis included 5,399 participants free of diabetes at baseline (1986) of validated self-reported European ancestry for whom genotype data based on genome wide association studies were available from nested case-control studies. Since the inception of the study, data on lifestyle and medical history have been ascertained through a self-administered questionnaire, including a semi-quantitative FFQ every 2-4 years.
T2D was ascertained if occurred between return of the questionnaire in 1990 and January 31 in 2012. Men who reported a diagnosis of diabetes in the biannual follow-up questionnaires were sent a supplementary questionnaire to confirm the diagnosis. A supplementary questionnaire regarding symptoms, diagnostic tests, and hypoglycaemic therapy was completed by participants who reported a diagnosis of diabetes. T2D cases were confirmed if at least one of the following was reported on the supplementary questionnaire according to the 1997 American Diabetes Association criteria: 1) one or more classic symptoms (excessive thirst, polyuria, weight loss, hunger) plus fasting plasma glucose concentrations ≥ 7.0mmol/L or random plasma glucose concentrations ≥11.1mmol/L), 2) at least 2 elevated plasma glucose concentrations on different occasions (fasting concentrations ≥ 7.0mmol/L, random plasma glucose concentrations ≥ 11.1mmol/L, and/or concentrations ≥ 11.1mmol/L after ≥2-h oral-glucose-tolerance test) in the absence of symptoms, or 3) treatment with hypoglycaemic medication (insulin or oral hypoglycaemic agent). Only confirmed cases were included in the analysis.
Study design reference: Rimm EB, et al. Prospective study of alcohol consumption and risk of coronary disease in men. Lancet. 1991; 338:464–8.

The **Health 2000 Survey** was a, comprehensive, population-based health examination survey. A nationally representative sample of 8,028 individuals aged 30 years or older were randomly selected from the Finnish population register from 80 health service districts throughout Finland using a two-stage stratified cluster sampling procedure. The longitudinal cohort of this study consisted of 1,946 adults free of diabetes who had attended the health examination proper in 2000-2001 with available genetic, lifestyle and clinical data.
In Health 2000 ascertainment of T2D was based on the same criteria as in the FINRISK study.
Study design reference: Aromaa A, et al. Health and Functional Capacity in Finland. Baseline Results of the Health 2000 Health Examination Survey. Helsinki: Publications of the National Public Health Institute B12; 2004. Available from http://urn.fi/URN:NBN:fife201204193452. Accessed June 25th 2017.

The **Inter99** is a population-based pre-randomized lifestyle intervention study aiming to prevent ischemic heart disease and T2D. The study population comprised 61,301 individuals living in the western part of Copenhagen County. An age- and sex-stratified random sample of 13,016 individuals, with a majority of individuals at the age of 40 to 50 years, was drawn from the study population by the Civil Registration System using computer generated random numbers and pre-randomized into two groups. Baseline data were collected from March 1999 until January

2001. Follow-up was conducted after one, three and five years for a health examination, completion of questionnaires and risk assessment. The total sample size included in this study was 5,607 participants. T2D was defined according to WHO 1999 criteria.

Study design reference: Jørgensen T, et al. A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. Eur J Cardiovasc Prev Rehabil. 2003; 10:377-86.

The **Malmö Diet and Cancer-Cardiovascular Cohort study** (MDC-CC**)** consists of individuals randomly (50%) invited to be involved in additional baseline examinations between 1991 and 1994. In total 6,103 individuals (46-68 y, 58% females) participated in the additional examinations. For this analysis, we included 3,745 individuals without prevalent diabetes for whom data on genotype and dietary intakes were available and who did not overlap with the case-cohort sample genotyped in the EPIC-InterAct project (i.e. cases ascertained until 31 December 2007 and a random subcohort).
T2D cases were defined as individuals with fasting whole blood glucose ≥ 6.5mmol/L or fasting plasma glucose ≥7.0mmol/L (verified with plasma glucose or oral glucose tolerance test (OGTT) in subsequent examination), ≥ 11.0mmol/L 2-h after OGTT, intake of diabetes medication (A10 drugs), or who have reported having diabetes in a questionnaire were identified as incident diabetes cases. In addition, T2D cases were identified via at least one of seven registries or at examinations during follow-up. The National Diabetes Register and the regional Diabetes 2000 Register required a proven diagnosis by a physician at the hospital based on international diagnosis standards (fasting plasma glucose concentration ≥ 7.0mmol/L, measured twice). Individuals with at least two HbA1c values above 6.0% with the Swedish Mono-S standardization system (corresponding to 6.9% in the US National Glycohemoglobin Standardization Program and 52mmol/mol with the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) units) were categorized as diabetes cases in the Malmö HbA1c Registry. Finally, cases were identified via registries from the National Board of Health and Welfare: the Swedish National Inpatient Registry, the Swedish Hospital-based outpatient care, the Cause-of-death Registry and the Swedish Prescribed Drug Registry. The subjects were followed until date of diabetes diagnosis, death, migration from Sweden or end of follow-up (31 December 2014), whichever occurred first.
Study design reference: Berglund G, et al. The Malmo diet and cancer study. Design and feasibility. J Intern Med. 1993; 233:45–51.

The **Multi-Ethnic Study of Atherosclerosis** (MESA) is a study of the characteristics of subclinical cardiovascular disease (disease detected non-invasively before it has produced clinical signs and symptoms) and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. MESA researchers study a diverse, population-based sample of 6,814 asymptomatic individuals of European- African-Hispanic- and Chinese American ancestry ascertained across six field centres across the United States. Baseline data for the current analyses were taken from the first clinic exam conducted in 2000 – 2002. In addition to yearly phone calls, follow-up clinic exams are conducted approximately every two years, and at the time the current analyses were conducted incident diabetes was available until the fifth clinic exam conducted in 2010-2012. The sample included in this study was composed of those 1,535 European descent individuals who were free of diabetes at baseline with available genetic and dietary information.
In MESA, incident diabetes was defined as a fasting plasma glucose ≥ 7.0mmol/L or treatment with either insulin or a hypoglycaemic agent at the follow-up examinations.
Study design reference*:* Bild DE, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. Am J Epidemiol. 2002; 156:871-81.

The study population was derived from the **Nurses' Health Study** (NHS), a prospective cohort of 121,700 female registered nurses aged 30–55 years at enrolment in 1976. Since the inception of the NHS, data on lifestyle and medical history have been ascertained through a self-administered questionnaire, including a validated semi-quantitative FFQ every 2–4 years. For this analysis, we used 1984 as baseline and a total of 8,510 healthy women of validated self-reported European ancestry with both genotype and dietary data available from genome wide association studies available from nested case–control studies. T2D was ascertained if occurred between return of the questionnaire and June 31 in 2012.
In NHS, T2D was defined using the same exact criteria as HPFS.
Study design reference: Colditz GA, et al. The Nurses' Health Study: 20-year contribution to the understanding of health among women. J Womens Health. 1997; 6:49–62.

The **Rotterdam Study** is a prospective population-based cohort study in Ommoord, a suburb of Rotterdam, designed to investigate the prevalence and incidence of and risk factors for chronic diseases in the elderly. The baseline exam of the first cohort (RS-I) was conducted between 1990 and 1993. A total of 7,983 adults, aged 55 years and over, participated in the study. Data on dietary intake was collected at baseline, and lifestyle and clinical characteristics were collected during four follow-ups visits up to 2012. For the current analysis, 2,509 adults were eligible as they had available data on DNA, dietary intake and outcome information, and consent to share genetic data.
In RS-I, T2D incidence was defined according to WHO guidelines as a fasting glucose level >7mmol/l, non-fasting glucose level >11.1mmol/l, or use of glucose-lowering medication.
Study design reference: Hofman A et al. The Rotterdam Study: 2016 objectives and design update. Eur J Epidemiol. 2015; 30:661-708.

The **Women's Genome Health Study** (WGHS) is a prospective cohort of initially healthy, female North American health care professionals at least 45 years old at baseline representing participants in the Women's Health Study (WHS) who provided a blood sample at baseline and consent for blood-based analyses. The WHS was a 2x2 trial beginning in 1992-1994 of vitamin E and low dose aspirin in prevention of cancer and cardiovascular disease with about 10 years of follow-up. Since the end of the trial, follow-up has continued in observational mode. Additional information related to health and lifestyle were collected by questionnaire throughout the WHS trial and continuing observational follow-up. Dietary data was collected from the WGHS only at baseline. We used a 131-item food frequency questionnaire (FFQ) to obtain information on usual intake of food and beverages at baseline and information on dietary and clinical characteristics were requested every 2 to 4-year cycles.
In WGHS, the diagnosis of T2D was based on revised American Diabetes Association diagnostic criteria. Cases were confirmed if 1 or more of the following conditions were met: (1) presence of more than 1 classic symptom of hyperglycaemia (i.e., polyuria, polydipsia, weight loss with or without polyphagia, and blurred vision) plus either a fasting plasma glucose ≥7.0mmol/L or higher or random plasma glucose ≥11.1mmol/L; (2) in the absence of symptoms, 2 or more elevated plasma glucose concentrations (fasting plasma glucose ≥7.0mmol/L, random plasma glucose ≥11.1mmol/L, or 2-hour plasma glucose ≥11.1mmol/L during oral glucose tolerance testing); or (3) use of insulin or an oral hypoglycaemic agent. The primary care physician's office was contacted for supporting documentation as necessary.
Study design reference: Ridker PM, et al. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy American women. Clin Chem. 2008; 54:249-55.

Appendix 5: Acknowledgement of sources of funding for participating cohort studies

The **Atherosclerosis Risk in Communities Study** is carried out as a collaborative study supported in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C, HHSN268201700001I, HHSN268201700003I, HHSN268201700005I, HHSN268201700004I, HHSN2682017000021. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. Dr. Selvin was supported by NIH/NIDDK grants K24DK106414 and R01DK089174.

The **Bogalusa Heart Study was** supported by funds from the National Heart, Lung and Blood Institute (HLI5103, HL071981, HL034594, and HL126024), National Institute on Aging (AG041200), the National Institute of Diabetes and Digestive and Kidney Diseases (DK115679, DK091718, DK100383, DK078616). The authors would like to thank the staff and participants of the Bogalusa Heart Study for their important contributions.

The **Cardiovascular Health Study** research reported in this article was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, and R01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The **Danish Diet, Cancer and Health study** was carried out as a part of the research program of the UNIK: Food, Fitness & Pharma for Health and Disease which was supported by the Danish Ministry of Science, Technology and Innovation. The Diet, Cancer and Health cohort was supported by the Danish Cancer Society. The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent research center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk). Tuomas O. Kilpeläinen was supported by the Danish Council for Independent Research (DFF – 1333-00124 and Sapere Aude program grant DFF – 1331-00730B).

The **EPIC-InterAct** study received funding from the European Union (Integrated Project LSHM-CT-2006-037197 in the Framework Programme 6 of the European Community). We thank all EPIC participants and staff for their contribution to the study. We thank Nicola Kerrison (MRC Epidemiology Unit, University of Cambridge, Cambridge, UK) for managing the data for the InterAct Project. 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).

The **Framingham Heart Study** was conducted using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). Funding support for the Framingham Food Frequency Questionnaire dataset was provided by ARS Contract #53-3k06-5-10, ARS Agreement #'s 58-1950-9-001 and 58-1950-7-707. Dr. Merino is supported by a postdoctoral fellowship funded by the European Commission; Horizon 2020 program; Marie Skłodowska-Curie actions (H2020-MSCA-IF-2015-703787). Drs Meigs and Dupuis are also supported by National Institute for Diabetes and Digestive and Kidney Diseases (NIDDK) R01 DK078616. Dr. Florez is a Massachusetts General Hospital Research Scholar and is supported by NIDDK K24 DK110550.

The **National FINRISK study** was partly supported by The Finnish Academy grant no. 269517 and the EU FP7 under grant agreements nr. 313010 (BBMRI-LPC), nr. 305280 (MIMOmics), and HZ2020 nr. 633589 (Ageing with Elegans), and the The Yrjö Jahnsson Foundation and The Juho Vainio Foundation. Dr. Kuokkanen was supported by The Päivikki and Sakari Sohlberg Foundation. Dr. Salomaa was supported by the Finnish Foundation for Cardiovascular Research.

The **Health Professionals' Follow-up study** was supported by grants UM CA167552, UM CA87969, UM CA49449, UM CA055075, R01 HL35464, R01 HL071981, R01 HL073168, R01 HL34594, R01 HL088521, U01HG004399, DK080140, DK091718, 5P30DK46200, U54CA155626, DK58845, U01HG004728-02, EY015473, DK70756 and DK46200 from the National Institutes of Health, with additional support for genotyping from Merck Research Laboratories, North Wales, PA USA.

The **Health 2000 Survey** was funded by the National Institute for Health and Welfare (THL), the Finnish Centre for Pension (ETK), the Social Insurance Institution (KEVA), and other institutions listed on the website of the survey (http://www.terveys2000.fi). The authors would like to thank the volunteering participants.

The **Inter99 study** was funded by: Danish Research Councils; The Health Foundation; The Danish Centre for Evaluation and Health Technology Assessment; Copenhagen County; Danish Heart Foundation; Ministry of Health and Prevention; Association of Danish Pharmacies; Augustinus Foundation; Novo Nordisk; Velux Foundation; Becket Foundation and Ib Henriksens Foundation.

The **Malmö Diet and Cancer-Cardiovascular Cohort study** was supported by the Swedish Research Council, the Swedish Heart and Lung Foundation, the Novo Nordic Foundation, the Swedish Diabetes Foundation, and the Påhlsson Foundation, and by equipment grants from the Knut and Alice Wallenberg Foundation, the Region Skåne, Skåne University Hospital, the Linneus Foundation for the Lund University Diabetes Center and the European Research Council.

The Multi-Ethnic Study of Atherosclerosis (MESA) and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420, UL1-TR-001881, and DK063491. Dr. Frazier-Wood was supported by an American Heart Association grant [14BGIA18740011], and also by the United States Department of Agriculture/Agricultural Research Service (USDA/ARS) [CRIS 309-5-001-058]. This work is a publication of the USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, Texas and funded in part with federal funds from the USDA/ARS under Cooperative Agreement No. 58-6250-0-008. The contents of this publication do not necessarily reflect the views or policies of

the USDA, nor does mention of trade names, commercial products, or organizations imply endorsement from the US government.

The **Nurses' Health Study** was supported by grants of UM1 CA186107, UM1 CA176726, R01 CA49449, and R01 CA67262 from the National Institutes of Health. Current study was supported by the National Heart, Lung, and Blood Institute (HL071981, HL034594, HL126024), the National Institute of Diabetes and Digestive and Kidney Diseases (DK58845, DK091718, DK100383), the Boston Obesity Nutrition Research Center (DK46200), and United States – Israel Binational Science Foundation Grant2011036. Dr. Guasch-Ferre is supported by American Diabetes Association grant #1-18PMF-029.

The **Rotterdam Study** is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

The **Women's Genome Health Study** is supported by the National Heart, Lung, and Blood Institute (HL043851 and HL080467) and the National Cancer Institute (CA047988 and UM1CA182913), with collaborative scientific support and funding for genotyping provided by Amgen.