Online Supplemental Material

Mendelian Randomization Method

Standard errors and inference for the instrumental variable (IV) estimator

After meta-analysis, we used the IV estimators to quantify the strength of the causal association of dairy intake and BMI.(1) The IV estimator which is identical to that derived by the widely used two-stage least squares method,(2) was calculated as the β of the regression coefficients *MCM6* variantrs4988235-BMI and *MCM6* variantrs4988235-Dairy:

(Eq. 1)
$$\beta_{IV} = \frac{\beta_{SNP\_BMI}}{\beta_{SNP\_Dairy}}$$

The standard error was calculated via the delta method as

(Eq. 2)
$$se_{IV} = abs(\beta_{IV}) \sqrt{\left(\frac{se_{SNP\_Dairy}}{\beta_{SNP\_Dairy}}\right)^2 + \left(\frac{se_{SNP\_BMI}}{\beta_{SNP\_BMI}}\right)^2}$$

Based on these estimates, we appeal to standard-normal asymptotics, with the resulting Wald test statistic and 95% confidence intervals given as

$t_{IV} = \frac{\beta_{IV}}{se_{IV}}$ $\qquad CI_{IV} = \beta_{IV} \pm 1.96\, se_{IV}$

The *P*-value for the H_0: $\beta_{IV} = 0$ was derived from the standard normal distribution. For the association with BMI, the 95% CI estimates were back-transformed through the antilogit and exponentiation, respectively.

For comparing the IV estimate β_{IV} and the conventional estimateβ_{Dairy_BMI}, we consider the difference

(Eq. 3)
$$\beta_{Diff} = \beta_{IV} - \beta_{Dairy\_BMI}$$

The corresponding standard error is

(Eq. 4)
$$se_{Diff} = \sqrt{(se_{IV})^2 + \left(se_{Dairy\_BMI}\right)^2}$$

We use again standard normal asymptotics for the difference, viz.

$t_{Diff} = \frac{\beta_{Diff}}{se_{Diff}}$ $\qquad CI_{Diff} = \beta_{Diff} \pm 1.96\, se_{Diff}$

The *P*-value for the H_0: $\beta_{Diff} = 0$ was derived from the standard normal distribution, and the confidence intervalswere back-transformed as above.

Supplemental Table 1 Baseline characteristics of participating studies

Full Name	Abbreviation	Sample size	Study design	Baseline year	Follow-up time (year)
Atherosclerosis Risk in Communities Study (African Ancestry)	ARIC-AA	1,889	Cohort	1987	5.8
Atherosclerosis Risk in Communities Study (European Ancestry)	ARIC-EA	8,233	Cohort	1987	6
Boston Puerto Rican Health Study	BPRHS	845	Cohort	2003	2.4
Copenhagen City Heart Study	CCHS	8,702	Cohort	1991-1994	20
Copenhagen General Population Study	CGPS	74,128	Cohort	2003-2011	5.7
Cardiovascular Health Study	CHS	1,943	Cohort	1989-90	8.9
Data from an Epidemiological Study on Insulin Resistance Syndrome	DESIR	3,468	Cohort	1994-1996	9
Diet, Cancer and Health Cohort	DCH	1,297	Nested cohort	1993-1997	5
Diet, Obesity and Genes Study (control)	Diogenes-C	1,002	Nested case-cohort	1993-1997	5
Diet, Obesity and Genes Study (weight gain)	Diogenes-W	813	Nested case-cohort	1993-1997	5
Family Heart Study	FamHS	2,167	Family Based Cohort	1992	7.9
Danish General Suburban Population Study	GESUS	14,751	Cohort	2010-2013	2.1
Gene-Lifestyle Interactions and Complex Traits Involved in Elevated Disease Risk Study	GLACIER	3,129	Cohort	1991-2001	9.9
Genetics of Lipid-lowering drugs and diet network	GOLDN	818	Cohort		0
Health Professional Follow-up Study	HPFS	7,599	Cohort	1990	10
InCHIANTI	INCH	647	Cohort	1998	8.7
Inter99	Inter99	6,161	Cohort	1999	5
Malmö Diet and Cancer study	MDCS	3,199	Cohort	1991-1996	16.7
Multi-Ethnic Study of Atherosclerosis	MESA	2,423	Cohort	1990	10
Nurses' Health Study	NHS	12,039	Cohort	1990	10
Prevención con Dieta Mediterránea-Valencia Study	PREDIMED-Valencia	940	Cohort	2003	2
Western Australian Pregnancy Cohort (Raine) Study	RAINE	730	Cohort	2010	2.1
Rotterdam Study	RS	3,215	Cohort	1990	6.5
Women's Genome Health Study	WGHS	23,294	Cohort	1992	2

Young Finns Study			YFS	1,370	Cohort		1980	4

Continued

Study Name	Male (n(%)	Ancestry	Country	Age (year)	BMI (kg/m^2)	Smoker (n (%))	rs4988235 (n (%))		
							TT	TC	CC
ARIC-AA	684(36.2)	AA	USA	53.2±5.7	29.7±5.9	517(27.4)	1393(73.7)	465(24.6)	31(1.6)
ARIC-EA	3,853(46.8)	Eu	USA	54.3±5.6	27.0±4.7	1,824(22.2)	739(9.0)	3190(38.8)	4304(52.3)
BPRHS	243(28)	Puerto Rican	USA	57±8	32±7	187(22.0)	520(61.5)	292(34.5)	41(4.8)
CCHS	3,904(44.9)	Eu	Denmark	60	25	4,269(49.1)	548(6.3)	3033(34.9)	5121(58.9)
CGPS	33,082(44.6)	Eu	Denmark	57	25.6	14,596(19.7)	4338(5.9)	26530(35.8)	43260(58.4)
CHS	730(37.7)	Eu	USA	71.1±4.3	26.4±4.2	172(8.9)	66(3.4)	839(46.8)	1031(53.3)
DESIR	1,698(49.0)	Eu	France	47.2±9.9	24.6±3.6	652(18.8)	747(21.5)	1696(48.9)	1025(29.6)
DDCH	582(44.9)	Eu	Denmark	55.9±4.4	25.2±3.5	415(32.0)	61(4.7)	415(32.0)	821(63.3)
Diogenes-C	510(50.9)	Eu	Denmark	53.6±2.60	25.5±3.6	287(28.6)	62(6.2)	363(36.2)	577(57.6)
Diogenes-W	399(49.1)	Eu	Denmark	53.4±2.6	26.9±4.0	263(32.4)	42(5.2)	283(34.8)	488(60.0)
FamHS	968(44.7)	Eu	USA	50.65	28.88	265(12.2)	256(11.8)	906(41.8)	1005(46.4)
GESUS	6,747(45.7)	Eu	Denmark	56	26.1	2,665(18.1)	884(6.0)	5339(36.2)	8528(57.8)
GLACIER	1,184(37.8)	Eu	Sweden	45.2±6.7	25.1±3.7	682(21.8)	216(6.9)	1095(35.0)	1818(58.1)
GOLDN	405(50)	Eu	USA	49±16	28±5	67(8.0)	82(10.0)	323(40.0)	413(50.0)
HPFS	7,599(100)	Eu	USA	57.7±11.8	25.9±3.3	3,577(53.0)	1128(16.8)	2702(39.3)	3057(43.8)
InCHIANTI	343(54.3)	Eu	Italy	63.4±14.8	27.1±4.0	127(20.1)	431(68.2)	187(29.6)	14(2.2)
Inter99	3,169(48.6)	Eu	Denmark	46.2±7.9	26.3±4.6	2,546(41.6)	388(6.3)	2222(36.3)	3507(57.3)
MDC	1279(4)	99.6% Eu	Sweden	56.3±5.7	25.4±3.7	729(22.8)	185(5.8)	1073(33.5)	1941(60.7)
MESA	1,134(46.8)	Eu*	USA	60.7±9.6	28.2±5.2	279(11.5)	488(20.1)	1196(49.4)	739(30.5)
NHS	0(0)	Eu	USA	57.3±9.6	26.2±5.2	6,250(55.6)	1541(14.1)	4658(41.0)	5149(44.9)
PREDIMED-V	338(36)	Eu	Spain	67±7	30.1±4.2	113(12.0)	153(16.3)	430(45.7)	357(38.0)
Raine	344(47.3)	Mixed**	Australia	19.9±0.3	24.3±5.1	93(12.7)	110(15.1)	283(38.8)	337(46.2)
Rotterdam	1,314(40.9)	Eu	Netherlands	65.8±6.8	26.3±3.5	702(21.8)	303(9.4)	1256(39.1)	1656(51.5)
WGHS	0(0)	Eu	USA	54.2±7.1	25.9±4.9	10,816(49.0)	2304(10.4)	8460(38.4)	11301(51.2)
YFS	593(43.3)	Eu	Finland	38.09	25.83	229(16.8)	206(15.0)	643(46.9)	521(38.0)

Data are numbers (%)
* Only Whites included
**Majority CEU/CEU-admixed,Asian(Chinese,Vietnamese),Indian,Polynesian,Aboriginal
We included 184,802participants from 25 studies.
AA: African Ancestry; Eu: European Ancestry; PREDIMED-V: PREDIMED-Valencia

Supplemental Table 2 Descriptions of studies

Study	Description of study
ARIC	ARIC is a multi-center prospective investigation of atherosclerotic disease in a predominantly bi- racial population conducted in four U.S. communities, involving both cohort and community surveillance components 2. Study participants aged 45-64 years at baseline were recruited from 4 communities: Forsyth County, North Carolina; Jackson, Mississippi; suburban areas of Minneapolis, Minnesota; and Washington County, Maryland. A total of 15,792 individuals participated in the baseline examination in 1987-1989, with follow-up examinations in approximate 3-year intervals, during 1990-1992, 1993-1995, and 1996-1998. Weight and height were measured. All study participants provided written informed consent.
BPRHS	The Boston Puerto Rican Health Study is an ongoing longitudinal cohort study designed to examine the role of psychosocial stress on presence and development of allostatic load and health outcomes in Puerto Ricans, and potential modification by nutritional status, genetic variation, and social support. Individuals who were self-identified Puerto Ricans, aged 45-75 years and residing in the Boston, MA metro area, were recruited through door-to-door enumeration and community approaches. Data including demographics, medical history, physical function, cognition and dietary data were collected through a comprehensive set of questionnaires and tests. Blood, urine and salivary samples were extracted for biomarker and genetic analysis. Measurements were repeated at a two-year follow-up and a five-year follow up.
CCHS	The Copenhagen City Heart Study (CCHS) is a prospective study of the general urban population in Copenhagen, Denmark, and was initiated in 1976-1978. Participants aged 20+ years were randomly selected among individuals living near Rigshospitalet, Copenhagen University Hospital based on residence according to the national Danish Civil Registration System. Participants from the first examination were re-invited to participate in follow-up examinations in 1981-1993, 1991-1994, and 2001-2003. In each of the follow-up examinations, a selection of first-time participants was invited to supplement the existing study population. Participants in the study completed a general questionnaire and a health examination including a non-fasting blood sample. In this study, approximately 8,700 white participants of Danish descent from the 1991-1994 examination were included, as additional blood samples were drawn at this point, allowing us to obtain genotypes for the LCT-13910 C/T genetic variant. The overall response rate for the 1991-1994 examination was 61%.
CGPS	The Copenhagen General Population Study (CGPS) is an ongoing prospective cohort study initiated in 2003 and includes individuals residing in the greater urban area of Copenhagen. Residence was determined by data from the national Danish Civil Registration System. All individuals aged 40+ years were invited along with a random selection of 25% of individuals aged 20-39 years, and the overall response rate was 45%. Participants in the study completed a general questionnaire and a health examination including a non-fasting blood sample. In this study we included approximately 74,000 white participants of Danish descent who had been genotyped for the LCT-13910 C/T genetic variant.
CHS	The Cardiovascular Health Study (CHS) is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥65 years conducted across four field centers [PMID: 1669507]. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional predominantly African-American cohort of 687 persons was enrolled for a total sample of 5,888. The analyses included 1964 White study participants without cardiovascular disease at baseline, with available genotype data and with follow-up data 9 years after baseline. CHS was approved by institutional review committees at each site, the subjects gave informed consent, and those included in the present analysis consented to the use of their genetic information for the study of cardiovascular disease.

DESIR	DESIR (Data from an Epidemiological Study on the Insulin Resistance syndrome) began in 1994 and included 5212 men and women, aged from 30 to 64 years. Participants were recruited from volunteers insured by the French Social Security system, which offers periodic health examinations free of charge. DESIR is a 9-year follow-up study with clinical and biological examinations every 3 years. A medical interview provided information about use of medication, and personal and familial history of diseases. Weight and height were measured by trained personnel with subjects lightly clothed. Blood pressure was measured by a doctor with a sphygmomanometer on the right arm, with subjects lying at rest for at least 5 minutes. Two measures of blood pressure were taken, and means were used for the analysis. At each visit, participants had to complete self-questionnaires about lifestyle (including diet). A 23-item questionnaire was completed by each participant to determine the frequency and level of consumption of different foods. This questionnaire has been validated by comparison with the dietary history method consisting of 30min interviews by trained dietitians. Two items concerned dairy products: cheese / milk and other dairy products (except cheese). Our analysis included only Caucasian subjects with lactase genotype, dietary data and examined 9 years after inclusion (n=3478). The study was approved by the ethics committee of the Kremlin Bicêtre Hospital and by the CNIL (Commission Nationale de l'Informatique et des Libertes).
DCH	The analyses include a subsample of 1,297 randomly selected individuals from the population-based Diet, Cancer and Health cohort of 57,053 individuals in the age of 50-64 years and born in Denmark. At baseline, anthrometric measurements were taken and blood pressure was measured. Information on usual diet and lifestyle was obtained using self-administered questionnaires, and biological samples were collected. At 5-year follow-up, the participants completed a repeat questionnaire and self-measured their BMI and waist circumference. The 1,297 participants were free of diabetes, cardiovascular disease, and cancer at baseline and at the 5-year follow-up.
Diogenes	The analyses include a subsample of 1,815 individuals from the the population-based Diet, Cancer and Health cohort of 57,053 Danish individuals in the age of 50-64 years and born in Denmark. At baseline, anthrometric measurements were taken and blood pressure was measured. Information on usual diet and lifestyle was obtained using self-administered questionnaires, and biological samples were collected. At 5-year follow-up, the participants completed a repeat questionnaire and self-measured their BMI and waist circumference. Of the 1,815 individuals included in the present analyses, 813 were weight-gainers and 1,002 were randomly selected control individuals. Weight gainers were defined as those individuals who had experienced the greatest degree of unexplained weight gain and were identified by using the residuals from a regression model of annual weight change on baseline values of age, weight, height, and smoking status (current/former/nonsmokers) and follow-up time. The participants were younger than 60 years at baseline and younger than 65 years at follow-up, with stable smoking habits, without cancer, cardiovascular disease, or diabetes at baseline and at the 5 year follow-up, and with a weight change not more than 5 kg/year.
FamHS	The Family HS began in 1992 with the ascertainment of 1,200 families (50% randomly sampled, and 50% high risk for CHD). The families (~6,000 individuals,) were sampled on the basis of information on probands from four population-based parent studies: the Framingham Heart Study, the Utah Family Tree Study, and two ARIC centers (Minneapolis, and Forsyth County, NC). Approximately eight years later, study participants belonging to the largest pedigrees were invited for a second clinical exam. A total of 2,767 participants of European descent in 510 extended families were examined. A total of 2,167 adults with available DNA and who provided valid dietary information were eligible for the current study.
GESUS	The Danish General Suburban Population Study (GESUS) was initiated in 2010 and concluded in 2013. The GESUS is a study of the suburban general population in Naestved Municipality located approximately 70km south of Copenhagen. All individuals aged 30+ and a random selection of 25% of the younger population aged 20-30 years were invited. Participants in the study completed a general questionnaire and a health examination including a non-fasting blood sample. In this study, we included approximately 14,000 white participants of Danish descent, with known LCT-13910 C/T genotypes. The response rate for this subset of the study population was 50%, and the GESUS had an overall response rate of 43% when the study was concluded in 2013.

GLACIER	The GLACIER Study (N~19,000) is a population-based cohort study of initially non-diseased adults living in the county of Västerbotten in Northern Sweden, nested within the Northern Sweden Health and Disease Study (Kurbasic A et al., Curr Nutr Rep, 2014). Clinical characteristics and lifestyle data were obtained as part of a population-wide health screening initiative called the Västerbotten's Health Survey (also called the Västerbotten Intervention Program), where habitants are invited to attend an extensive health examination the year of their 40th, 50th, and 60th birthday. The total number of GLACIER participant with genotype and phenotype data available for the current analysis was 3,129. Ethical approval for the GLACIER Study was obtained from the Regional Ethical Review Board in Umeå.
GOLDN	The GOLDN Study belongs to the PROgram for GENetic Interaction (PROGENI) Network, which includes family studies examining gene-environment interactions via controlled interventions. GOLDN participants were recruited from three-generational pedigrees previously identified in the Minneapolis, MN, and Salt Lake City, UT, field centers of the National Heart, Lung, and Blood Institute Family Heart Study. Clinical, dietary and biochemical measurements were collected at baseline and following an intervention with fenofibrate. A postprandial study fat-challenge was also conducted pre - and post- fenofibrate. Individuals were >= 18 years with fasting triglycerides < 1500 mg/dl.
HPFS	The HPFS was initiated in 1986, and was composed of 51,529 male dentists, pharmacists, veterinarians, optometrists, osteopathic physicians, and podiatrists, aged 40-75 y at baseline. The male participants returned a baseline questionnaire about detailed medical history, lifestyle, and usual diet. Questionnaires were collected at baseline and biennially thereafter, to update information on lifestyle factors and the occurrence of chronic diseases. In the current analysis, we used 1990 as baseline in the HPFS, when the earliest complete dietary data were collected. Our analysis included 8,000 men whose genotype data were available. All of the participants were Caucasians and were free of diabetes, cardiovascular disease, and cancer at baseline. The study protocol was approved by the institutional review boards of Brigham and Women's Hospital and Harvard School of Public Health.
InCHIANTI	The InCHIANTI study is a population-based epidemiological study aimed at evaluating the factors that influence mobility in the older population living in the Chianti region in Tuscany, Italy. Overnight fasted blood samples were for genomic DNA extraction and genotyping using Illumina Infinium HumanHap 550K SNP arrays were used for genotyping. Dairy consumption was assessment using a questionnaire designed for the EPIC study. The analysis was restricted to those with data on dairy intake, BMI, blood pressure, and genotyping. The study protocol was approved by the Italian National Institute of Research and Care of Aging Institutional Review and Medstar Research Institute (Baltimore, MD).
Inter99	The Danish population-based Inter99 study (ClinicalTrials.gov ID-no: NCT00289237) is a non-pharmacological intervention study for ischemic heart disease, initiated in 1999 at the Research Centre for Prevention and Health, Glostrup, Denmark. A random sample of 13,016 individuals living in Copenhagen County from seven different age groups (30-60 years, grouped with five year intervals) was drawn from the Civil Registration System and 6,784 of these attended the health examination
MDCS	The MDCS was a population based study conducted 1991-1996. In total, 28098 individuals (45-74 years) completed all baseline examinations. The MDC cardiovascular sub-cohort (n=6103) were invited to a follow-up study 2007-2012. Individuals from the follow-up study with data available on genotypes were included in this study (n=3199)
MESA	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 men and women aged 45-84 at baseline. Thirty-eight percent of the recruited participants were white, who were included in the current analyses. Participants were recruited from six field centers in 2000-2002 across the United States and had been followed-up five times at the time of data analysis with an average time period of follow-up of 20 months between each visit. The tenets of the Declaration of Helsinki were followed and institutional review board approval was granted at all MESA sites. Written informed consent was obtained from each participant.

NHS	The NHS began in 1976, when 121,700 female registered nurses aged 30-55 y residing in 11 states were recruited to complete a baseline questionnaire about their lifestyle and medical history. Questionnaires were collected at baseline and biennially thereafter, to update information on lifestyle factors and the occurrence of chronic diseases. In the current analysis, we used 1990 as baseline in the NHS, when the earliest complete dietary data were collected. Our analysis included 13,000 women whose genotype data were available. All of the participants were Caucasians and were free of diabetes, cardiovascular disease, and cancer at baseline. The study protocol was approved by the institutional review boards of Brigham and Women's Hospital and Harvard School of Public Health.
PREDIMED-Valencia	The PREDIMED study was initiated in 2003 y was composed of 1094 participants. LCT genotypes were available for 940. PREDIMED is a randomized intervention trial with Mediterranean diet versus a control diet. FFQ questionnaires were completed at baseline and yearly. BMI and blood pressure were directly measured yearly. We used data corresponding to the 2-y follow-up period.
RAINE	The Western Australian Pregnancy Cohort (Raine) Study is a prospective pregnancy cohort where 2900 were recruited from King Edward Memorial Hospital between 1989 and 1991. Data were collected throughout pregnancy and the children have been followed-up at ages 1, 2, 3, 5, 8, 10, 14, 17, 18, 20 and 22. Ethnics approval for this study was obtained from King Edward Memorial Hospital and Princess Margaret Hospital. Participants were consented to being involved in this study prior to each follow-up. In the current analysis, we used the 20 year follow-up as baseline in the Raine Study. Analyses were performed at the 22 year follow-up on xx number of participants who had genotyping data, food frequency data from the 20 year follow-up amd outcomes of interests for this study.
RS	The Rotterdam Study is a population-based prospective cohort study ongoing since 1990 in the city of Rotterdam in The Netherlands. The study was designed to investigate the prevalence and incidence of and risk factors for chronic diseases in the elderly. All inhabitants of Ommord, a district of Rotterdam, the Netherlands, aged 55 years and older were invited. At enrollment, participants were interviewed at home (2h) and were examined in detail at our research facility (5h). These measurements were repeated every 3 to 4 years.
WGHS	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.
YFS	The Cardiovascular Risk in Young Finns (YFS) is a population-based 27 year follow up-study (http://med.utu.fi/cardio/youngfinnsstudy/). The first cross-sectional survey was conducted in 1980, when 3,596 Caucasian subjects aged 3-18 years participated. In adulthood, the latest 27-year follow-up study was conducted in 2007 (ages 30-45 years) with 2,204 participants. The study cohort for the present analysis comprised subjects who had participated in the study in 2007 and had validated dietary data from FFQ. The dietary intake of nutrients was assessed using a modified 131-item food frequency questionnaire developed by the Finnish National Institute for Health and Welfare. The study was approved by the local Ethical Committees and was performed according to Helsinki declaration.

Supplemental Table 3 Software for analysis and ethical approvals

Study	Software for analysis	Ethic issue*
	SAS/R/SPSS/STATA	
ARIC-AA	STATA	Yes
ARIC-EA	STATA	Yes
BPRHS	SAS	Yes
CCHS	STATA	Yes
CGPS	STATA	Yes
CHS	STATA	Yes
DESIR	SYSTAT	Yes
DCH	SAS	Yes
Diogenes	SAS	Yes
FamHS	SAS	Yes
GESUS	STATA	Yes
GLACIER	SAS	Yes
GOLDN	SAS	Yes
HPFS	SAS	Yes
INCH	SAS	Yes
Inter99	R	Yes
MDCS	SPSS	Yes
MESA	SAS	Yes
NHS	SAS	Yes
PREDIMED-Valencia	SPSS	Yes
RAINE	R	Yes
RS	SPSS	Yes
WGHS	R	Yes
YFS	R	Yes

*All participants provided written, informed consent, and ethical approval was granted by local ethics committees for participating studies.

Supplemental Table 4 Dairy consumption assessment

Study	Dairy consumption				
	Unit	**Mean (SD)**	**Measurement**	**Dairy products included**	**Time of assessment**
ARIC-AA	Servings/day	1.25 (1.37)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline
ARIC-EA	Servings/day	1.80 (1.39)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline
BPRHS	servings/day	2.3(1.5)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline
CCHS	NA	NA	NA	NA	NA
CGPS	Servings/day	1.69(1.21)	general questionnaire	Milk (whole, semi-skimmed, skimmed), cheese	Baseline
CHS	Servings/day	1.37 (0.70)	Picture-sort FFQ	Skim milk/buttermilk, low fat milk/beverages with low fat milk, whole milk/beverages with whole milk, ice cream, yogurt, cottage cheese, other cheese/cheese spread	Baseline
DESIR	Servings/day	1.36 (0.67)	FFQ (very short questionnaire)	milk, yogurt, cottage/fresh soft cheese, custard type desserts	Baseline
DCH	Servings/day	1.64 (1.21)	FFQ	Skimmed milk, Semi-skimmed milk, Whole fat milk, Buttermilk, Fermented dairy (low-fat), Fermented dairy (whole-fat), Ice cream (low fat), Ice cream (whole fat), Cream	Baseline
Diogenes	Servings/day	1.59 (1.40)	FFQ	Skimmed milk, Semi-skimmed milk, Whole fat milk, Buttermilk, Fermented dairy (low-fat), Fermented dairy (whole-fat), Ice cream (low fat), Ice cream (whole fat), Cream	Baseline
FamHS	Servings/day	2.04 (1.46)	66-item interviewer administered modified Willett FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, and other cheese	Baseline
GESUS	Servings/day	2.38(1.51)	general questionnaire	Milk(whole, semi-skimmed, skimmed, butter milk, lactose free), cheese, fermented milk	Baseline
GLACIER	Servings/day	3.32 (1.55)	FFQ	cream, sour cream, hard cheese 28% fat, hard cheese 10-17% fat, sour milk and yoghurt (0.5 and 3% fat), ice cream, milk (0.5, 1.5 and 3% fat)	Baseline
GOLDN	Servings/day	1.97 (1.5)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline
HPFS	Servings/day	1.37 (0.70)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline

INCH	Servings/day	1.09 (0.8)	FFQ	whole milk, low fat milk, icecream, yogurt, latte, butter, cheese, soft cheeseer	Baseline
Inter99	Servings/day	5 (4.6)	FFQ	Skim/low fat milk, whole milk, chocolate milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline
MDCS	Servings/day	4.8 (2.2)	Modified diet history	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese and cream	Baseline
MESA	Servings/day	2.08 (1.78)	FFQ	Yogurt, dairy-based desserts, ice cream, cottage cheese / ricotta, other cheesem skim/ low-fat milk, whole milk, cream in coffee	Baseline
NHS	Servings/day	1.37 (0.70)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline
PREDIMED-Valencia	Servings/day	1.85 (1.10)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline
RAINE	Servings/day	1.90 (1.01)	FFQ	Full cream milk, Reduced fat milk, Skim milk, hard cheese, firm cheese, ricotta, low fat cheese, flavoured milk, ice cream, yoghurt, cream cheese, soft cheese	Baseline
RS	Servings/day	3.99 (2.37)	FFQ	Whole milk, skimmed milk, yogurt, hard cheese, soft cheese, whip cream, coffee cream, ice cream, custard	Baseline
WGHS	Servings/day	1.98 (1.36)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage cheese, cream cheese, other cheese, cream, sour cream, and sherbert	Baseline
YFS	Servings/day	4.30 (2.52)	FFQ	Skim/low fat milk, whole milk, ice cream, yogurt, cottage/ricotta cheese, cream cheese, other cheese, and cream	Baseline

FFQ, food frequency questionnaire

Supplemental Table 5 Measurement of body mass index

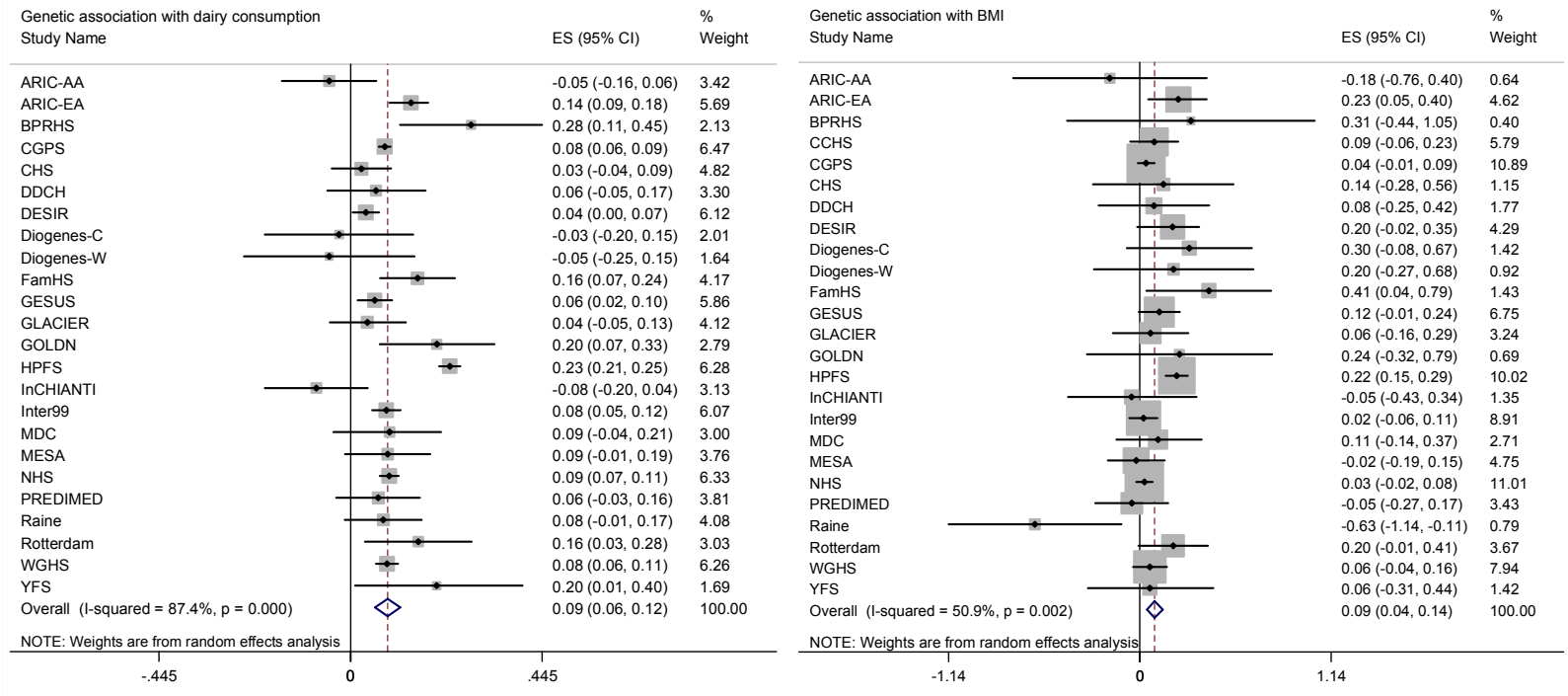
Study	Outcome BMI	
	Clinical measurement/ self-reported	**Time of measurement**
ARIC-AA	Clinical measurement	End of follow-up
ARIC-EA	Clinical measurement	End of follow-up
BPRHS	clinical measurement	Baseline
CCHS	Clinical measurement	Baseline
CGPS	Clinical measurement	Baseline
CHS	Clinical measurement	End of follow-up
DESIR	Clinical measurement	End of follow-up
DCH	Self-reported	End of follow-up
Diogenes	Self-reported	End of follow-up
FamHS	Clinical measurement	End of follow-up
GESUS	Clinical measurement	Baseline
GLACIER	Clinical measurement	End of follow-up
GOLDN	clinical measurement	Baseline
HPFS	Self-reported	End of follow-up
InCHIANTI	Clinical measurement	Baseline
Inter99	Clinical measurement	End of follow-up
MDCS	Clinical measurement	End of follow-up
MESA	Self-reported	End of follow-up
NHS	Self-reported	End of follow-up
PREDIMED-Valencia	Clinical measurement	End of follow-up
RAINE	Clinical measurement	End of follow-up
RS	Clinical measurement	End of follow-up
WGHS	Self-reported	End of follow-up
YFS	Clinical measurement	End of follow-up

Supplemental Table 6 SNP and proxy SNP selection and genotyping information

Study	Genotype					
	SNP	Genotyping method	Sample call rate	Proxy SNP/R^2	MAF	HWE *P*-value
ARIC-AA	rs1446585	Imputed	NA	R^2=1 in Africans	12.9	
ARIC-EA	rs4988235	Imputed	NA	NA	30.2	
BPRHS	rs4988235	TaqMan	>97%	NA	21.9	0.5
CCHS	rs4988236	TaqMan	>99.9%	NA	6.3	0.0006
CGPS	rs4988237	TaqMan	>99.9%	NA	5.85	0.0014
CHS	rs4988235	Whole genome scan using Illumina 370CNV BeadChip systemSNP, and imputation to HapMap build 36 using BIMBAM v0.99.	≥ 97%	NA	70	NA
DESIR	rs4988235	Kaspar	97%	NA	46	0.36
DCH	rs4988235	Imputed from 1000Genomes based on Illumina HumanCoreExome BeadChip genotypes	>95%	NA	20.70%	0.36
Diogenes	rs4988235	Illumina Metabochip	>95%	NA	23.53%	0.64
FamHS	rs4988235	Imputed to 1000Genomes from Illumina GWAS arrays	100%	NA	32.72	0.019
GESUS	rs4988238	Competitive Allele-Specific PCR-based assay (KASP)	>99.9%	NA	5.99	0.21
GLACIER	rs4988235	Metabochip array (Illumina Inc. San Diego, CA, USA)	100%	NA	25.4	0.00025
GOLDN	rs4988235	TaqMan	>97%	NA	29.8	0.06
HPFS	rs4988235	TaqMan	97%	NA	35.9	0.37
INCHIANTI	rs4988235	Illumina 550K	97%	NA	17	0.23
Inter99	rs4988235	The Cardio-MetaboChip on an Illumina HiScan system	99.3%	NA	24.50%	0.16
MDCS	rs309137	Illumina HumanOmniExpress BeadChip v. 1	99.8%	0.77	22.6	0.02
MESA	rs4988235	Affy 6.0	≥ 95%	NA	0.45	0.24
NHS	rs4988235	TaqMan	97%	NA	33.9	0.35
PREDIMED-Valencia	rs4988235	TaqMan	97%	NA	Minor allele is T	0.221
RAINE	rs4988235	Illumina 660W Quad Array; Human	97%	NA	660W: 30.3;	660W: 1.8 E-05;

		Omni Express BeadChip			Omni:47.1	Omni: 1
RS	rs4988235	Illumina 550	97.5%	NA	28.9	0.13
WGHS	rs4988235	Imputed	NA	NA	29.7	0.873
YFS	rs4988235	custom Illumina BeadChip Human670K	95%	NA	39.57	0.19268

Supplemental Figure 1 Genetic association and estimated causality between dairy intake and BMI in additive model



Linear regression was used to test the association of *MCM6* variantrs4988235 in additive model with dairy intake or BMI after adjustment of age, sex, ethnicity, and region in each study.

Random-effects meta-analysis was used to pool the genetic association with dairy intake under additive modelin 176,100participants from 24 studies due to the heterogeneity between studies (I^2 =87.4%, P<0.001). Random-effects meta-analysis was used to pool the genetic association with BMI association under additive model in 184,802participants from 25 studies due to the heterogeneity between studies (I^2 =50.9%, P=0.002).

We used the IV estimators to quantify the strength of the causal association of dairy intake and BMI in each study. TheIV estimator which is identical to that derived by the widely used two-stage least squares method, will be calculated as the β of the regression coefficients *MCM6* rs4988235-BMI and *MCM6* rs4988235-Dairy.

PREDIMED: PREDIMED-Valencia

Supplemental Table7 Causal association between Dairyand BMI in recessive model (TT vs CT+CC)

Subgroup	Dairy (outcome)			BMI (outcome)			Estimated Causality	
	Studies	SNP (IV)		Studies	SNP (IV)			
	No.	β(95%CI)	*P*-value	No.	β(95%CI)	*P*-value	β(95%CI)	*P*-value
All	24	0.06(0.03,0.10)	<0.001	25	0.10(0.04,0.16)	<0.001	1.59(0.28,2.91)	<0.001
Age, y								
≥50	19	0.08(0.04,0.12)	<0.001	19	0.12(0.05,0.19)	0.001	1.57(0.36,2.78)	<0.001
<50	5	0.01(-0.08,0.11)	0.783	6	0.03(-0.10,0.16)	0.642	-0.19(-2.58,2.20)	0.875
BMI, kg/m^2								
≥25	19	0.09(0.04,0.13)	0.001	19	0.11(0.04,0.18)	0.002	1.28(0.25,2.31)	<0.001
<25	5	-0.01(-0.07,0.05)	0.788	6	0.08(-0.05,0.21)	0.227	-0.94(-3.18,1.31)	0.413
Follow, y								
≥5	13	0.06(0.01,0.12)	0.019	14	0.13(0.05,0.21)	0.002	2.00(-0.06,4.06)	0.061
<5	11	0.06(0.03,0.10)	<0.001	11	0.04(-0.04,0.12)	0.298	0.58(-0.77,1.92)	0.4
Sample Size								
≥1000	20	0.07(0.03,0.10)	0.001	21	0.10(0.04,0.16)	0.002	1.55(0.22,2.88)	<0.001
<1000	4	0.01(-0.24,0.26)	0.939	4	0.20(-0.23,0.63)	0.363	0.54(-2.55,3.62)	0.734
Ethnic group								
Eu	19	0.07(0.03,0.11)	<0.001	20	0.11(0.05,0.16)	<0.001	1.50(0.32,2.68)	<0.001
Non-Eu	5	0.01(-0.08,0.10)	0.789	5	-0.05(-0.79,0.68)	0.885	-3.57(-11.96,4.83)	0.405
Study Design								
Cohort	19	0.06(0.01,0.10)	<0.001	20	0.11(0.04,0.18)	<0.001	1.83(0.04,3.63)	0.041
Cross-sectional	5	0.06(0.05,0.08)	<0.001	5	0.04(-0.03,0.10)	0.125	0.67(-0.42,1.76)	0.212

Linear regression was used to test the association of *MCM6* variant rs4988235with dairy intakeor BMI after adjustment of age, sex, ethnicity, region, total energy and principal component for population stratification, as appropriate.

We pooled β coefficients across studies using random-effect meta-analysis due to the heterogeneity between studies (I^2>50%, *P*<0.001).

We used the IV estimators toquantify the strength of the causal association of dairy intake and BMI.

The IV estimator which is identical to that derived by the widely used two-stage least squares method, will be calculated as the β of the regression coefficients *MCM6* rs4988235-BMI and *MCM6* rs4988235-Dairy.

Supplemental Table 8 Acknowledgments and Funding Sources

Study	Acknowledgments and Funding Sources
ARIC	The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by 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. 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.
BPRHS	This study was supported by the National Institutes of Health grants P01 AG023394 and P50 HL105185. C.E. Smith is supported by K08 HL112845.
CCHS	The Copenhagen City Heart Study was funded by the Danish Heart Foundation.
CGPS	The Copenhagen General Population Study was funded by The Danish Council for Independent Research; Medical Sciences (FSS); Herlev Hospital, Copenhagen University Hospital; Copenhagen County Foundation; and Chief Physician Johan Boserup and Lise Boserup's Fund, Denmark.
CHS	Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. This CHS research 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). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. 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. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
DESIR	The DESIR study has been supported by INSERM contracts with CNAMTS, Lilly, Novartis Pharma and Sanofi-Aventis; by INSERM (Réseaux en Santé Publique, Interactions entre les déterminants de la santé, Cohortes Santé TGIR 2008), the Association Diabète Risque Vasculaire, the Fédération Française de Cardiologie, La Fondation de France, ALFEDIAM, CNIEL, ONIVINS, Société Francophone du Diabète, Ardix Medical, Bayer Diagnostics, Becton Dickinson, Cardionics, Merck Santé, Novo Nordisk, Pierre Fabre, Roche, Topcon. The funding sources had no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript. The DESIR Study Group: INSERM CESP U1018: B Balkau, P Ducimetière, E Eschwège; INSERM U367: F. Alhenc-Gelas; CHU d'Angers: A Girault; Bichat Hospital: F Fumeron, M Marre, R Roussel; CHU de Rennes: F Bonnet; CNRS UMR8090, Lille: S Cauchi P Froguel; Centres d'Examens de Santé: Alençon, Angers, Blois, Caen, Chartres, Chateauroux, Cholet, Le Mans, Orléans, Tours; Institute de Recherche Médecine Générale: J Cogneau; General practitioners of the region; Institute inter-Regional pour la Santé: C Born, E Caces, N Copin, JG Moreau, O Lantieri, F Rakotozafy, J Tichet, S Vol.
DCH	This work was carried out as a part of the research program of the UNIK: Food, Fitness & Pharma for Health and Disease (see www.foodfitnesspharma.ku.dk). The UNIK project was supported by the Danish Ministry of Science, Technology and Innovation. Tuomas O. Kilpeläinen was supported by the Danish Council for Independent Research (DFF – 1333-00124 and Sapere Aude program grant DFF – 1331-00730B).

Diogenes	This work was carried out as a part of the research program of the UNIK: Food, Fitness & Pharma for Health and Disease (see www.foodfitnesspharma.ku.dk). The UNIK project was supported by the Danish Ministry of Science, Technology and Innovation. Tuomas O. Kilpeläinen was supported by the Danish Council for Independent Research (DFF – 1333-00124 and Sapere Aude program grant DFF – 1331-00730B).
FamHS	This study was supported by grants DK089256 and HL117078 from the National Insitutes of Health.
GESUS	The Danish General Suburban Population Study was funded by the Region Zealand Foundation, Naestved Hospital Foundation. Edith and Henrik Henriksens Memorial Scholarship, Johan and Lise Boserup Foundation, TrygFonden, Johannes Fog's Foundation, Region Zealand, Naestved Hospital, The National Board of Health, and the Local Government Denmark Foundation.
GLACIER	The authors thank the participants who dedicated their time to the Västerbotten Intervention Project, the health professionals for data collection and the staff at the Department of Biobank Research for data management. The authors also thank Kerstin Enqvist and Thore Johansson (Biobank Norr, Umeå, Sweden) for DNA preparation and Felicity Payne (Wellcome Trust Sanger Institute, Hinxton, UK) for support with genotyping. The GLACIER Study was funded by project grants from the Swedish Heart-Lung Foundation, the Swedish Diabetes Association, the Påhlsson's Foundation, Region Skåne, the Swedish Research Council, the Umeå Medical Research Foundation, Novo Nordisk, and The Heart Foundation of Northern Sweden (all to PWF). The authors also acknowledge the funding agencies supporting the Northern Sweden Diet Database and the Västerbotten Intervention Project, including the Swedish Research Council.
GOLDN	This study was supported by National Heart, Lung, and Blood Institute (NHLBI) grant no. U01HL072524 (Genetic and Environmental Determinants of Triglycerides), NHLBI R01 HL091357 (Genomewide Association Study of Lipid Response to Fenofibrate and Dietary Fat), NHLBI grant number HL54776 and HL078885; and by contracts 53-K06-5-10 and 58-1950-9-001 from the US Department of Agriculture, Agriculture Research Service. C.E. Smith is supported by K08 HL112845.
HPFS	This study was supported by grants HL126024, HL034594, DK100383, DK091718, HL071981, HL073168, CA87969, CA49449, CA055075, HL34594, HL088521, U01HG004399, DK080140, P30DK46200, U01CA137088, U54CA155626, DK58845, DK098311, U01HG004728, EY015473, CA134958, DK70756 and DK46200 from the National Institutes of Health, with additional support for genotyping from Merck Research Laboratories, North Wales, PA. LQ is a recipient of the American Heart Association Scientist Development Award (0730094N). LRP is supported by the Arthur Ashley Williams Foundation and a Harvard Ophthalmology Scholar Award (Harvard Medical School) from the Harvard Glaucoma Center of Excellence. ATC is a Damon Runyon Cancer Foundation Clinical Investigator. The funding sources had no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.
InCHIANTI	The InCHIANTI study baseline (1998-2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336).
Inter99	None
MDCS	The Malmö Diet & Cancer Study, (Malmö) was initiated and planned in collaboration with the International Agency for Research on Cancer, the Swedish Cancer Society, and Swedish Medical Research Council and the Faculty of Medicine Lund University, Sweden. The study is also funded by Region Skåne, City of Malmö, Påhlsson Foundation and the Swedish Heart and Lung Foundation.
MESA	MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) 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-001079, UL1-TR-000040, and DK063491.
NHS	This study was supported by grants HL126024, HL034594, DK100383, DK091718, HL071981, HL073168, CA87969, CA49449, CA055075, HL34594, HL088521, U01HG004399, DK080140, P30DK46200, U01CA137088, U54CA155626, DK58845, DK098311, U01HG004728, EY015473, CA134958, DK70756 and DK46200 from the National Institutes of Health, with additional support for

genotyping from Merck Research Laboratories, North Wales, PA. LQ is a recipient of the American Heart Association Scientist Development Award (0730094N). LRP is supported by the Arthur Ashley Williams Foundation and a Harvard Ophthalmology Scholar Award (Harvard Medical School) from the Harvard Glaucoma Center of Excellence. ATC is a Damon Runyon Cancer Foundation Clinical Investigator. The funding sources had no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

PREDIMED-Valencia: The PREDIMED-Valencia study was supported by the Spanish Ministry of Health (Instituto de Salud Carlos III) and the Ministerio de Economía y Competitividad (projects G03/140, CIBER 06/03, RD06/0045, PI07-0954, CNIC-06, PI11/02505, SAF2009-12304, AGL2010-22319-C03-03, SAF2016-80532-R), the Ministerio de Educacion, Cultura y Deporte (PRX14/00527 and PRX17/00500), Fondo Europeo de Desarrollo Regional, by the University Jaume I (Project P1-1B2013-54), Fundació La Marató de TV3 (538/U/2016) and by the Generalitat Valenciana (AP111/10, AP-042/11, BEST/2015/087, GVACOMP2011-151, ACOMP/2011/145, ACOMP/2012/190 and ACOMP/2013/159). The PREDIMED-Valencia study also acknowledges the collaborative support provided by the Real Colegio Complutense at Harvard University.

RAINE: This study was supported by the National Health and Medical Research Council of Australia [grant numbers 403981 and 003209] and the Canadian Institutes of Health Research [grant number MOP-82893]. The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the NH&MRC for their long term funding to the study over the last 25 years and also the following institutes for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA) , Curtin University, the Raine Medical Research Foundation, the UWA Faculty of Medicine, Dentistry and Health Sciences, the Telethon Kids Institute, the Women's and Infant's Research Foundation (King Edward Memorial Hospital) and Edith Cowan University). The authors gratefully acknowledge the assistance of the Western Australian DNA Bank (National Health and Medical Research Council of Australia National Enabling Facility). We would also like to acknowledge the Raine Study participants for their ongoing participation in the study, the Raine Study Team for study co-ordination and data collection, the UWA Centre for Science for utilisation of the facility and the Sleep Study Technicians. The 22 year Raine Study follow-up was funded by NHMRC project grants 1027449, 1044840 and 1021855. Funding was also generously provided by Safework Australia. This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and Government of Western Australia.

RS: 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 generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810.

WGHS: The WGHS 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.

YFS	The Young Finns Study has been financially supported by the Academy of Finland: grants 286284 (T.L.), 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001 for T.L.); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research (T.L.); Finnish Cultural Foundation; Tampere Tuberculosis Foundation (T.L.); Emil Aaltonen Foundation (T.L.); and Yrjö Jahnsson Foundation (T.L.). The expert technical assistance in the statistical analyses by Ville Aalto, Irina Lisinen and Mika Helminen is gratefully acknowledged.

References

1. Wald A. The fitting of straight lines if both variables are subject to error. Ann Math Stat 1940;11:284–300.
2. Palmer TM, Sterne JA, Harbord RM, Lawlor DA, Sheehan NA, Meng S, et al. Instrumental variable estimation of causal risk ratios and causal odds ratios in mendelian randomization analyses. Am J Epidemiol 2011;173:1392-403.