International Journal of Epidemiology, 2016, 1938–1952
doi: 10.1093/ije/dyw325
Advance Access Publication Date: 29 December 2016
Original article

Mendelian randomisation and instrumental variable analysis

Coffee intake, cardiovascular disease and all-cause mortality: observational and Mendelian randomization analyses in 95 000–223 000 individuals

Ask Tybjærg Nordestgaard[1,2] and Børge Grønne Nordestgaard[1,2,3]*

[1]Department of Clinical Biochemistry and the Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark, [2]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark and [3]Copenhagen City Heart Study, Frederiksberg Hospital, Copenhagen University Hospital, Frederiksberg, Denmark

*Corresponding author. Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev Ringvej 75, DK-2730 Herlev, Denmark. E-mail: Boerge.Nordestgaard@regionh.dk

Accepted 17 October 2016

Abstract

Background: Coffee has been associated with modestly lower risk of cardiovascular disease and all-cause mortality in meta-analyses; however, it is unclear whether these are causal associations. We tested first whether coffee intake is associated with cardiovascular disease and all-cause mortality observationally; second, whether genetic variations previously associated with caffeine intake are associated with coffee intake; and third, whether the genetic variations are associated with cardiovascular disease and all-cause mortality.

Methods: First, we used multivariable adjusted Cox proportional hazard regression models evaluated with restricted cubic splines to examine observational associations in 95 366 White Danes. Second, we estimated mean coffee intake according to five genetic variations near the *AHR* (rs4410790; rs6968865) and *CYP1A1/2* genes (rs2470893; rs2472297; rs2472299). Third, we used sex- and age adjusted Cox proportional hazard regression models to examine genetic associations with cardiovascular disease and all-cause mortality in 112 509 Danes. Finally, we used sex and age-adjusted logistic regression models to examine genetic associations with ischaemic heart disease including the Cardiogram and C4D consortia in a total of up to 223 414 individuals. We applied similar analyses to *ApoE* genotypes associated with plasma cholesterol levels, as a positive control.

Results: In observational analyses, we observed U-shaped associations between coffee intake and cardiovascular disease and all-cause mortality; lowest risks were observed in individuals with medium coffee intake. Caffeine intake allele score (rs4410790 + rs2470893) was associated with a 42% higher coffee intake. Hazard ratios per caffeine intake allele were 1.02 (95% confidence interval: 1.00–1.03) for ischaemic heart disease,

1.02 (0.99–1.02) for ischaemic stroke, 1.02 (1.00–1.03) for ischaemic vascular disease, 1.02 (0.99–1.06) for cardiovascular mortality and 1.01 (0.99–1.03) for all-cause mortality. Including international consortia, odds ratios per caffeine intake allele for ischaemic heart disease were 1.00 (0.98–1.02) for rs4410790, 1.01 (0.99–1.03) for rs6968865, 1.02 (1.00–1.04) for rs2470893, 1.02 (1.00–1.04) for rs2472297 and 1.03 (0.99–1.06) for rs2472299. Conversely, 5% lower cholesterol level caused by *ApoE* genotype had a corresponding odds ratio for ischaemic heart disease of 0.93 (0.89–0.97).

Conclusions: Observationally, coffee intake was associated with U-shaped lower risk of cardiovascular disease and all-cause mortality; however, genetically caffeine intake was not associated with risk of cardiovascular disease or all-cause mortality.

Key words: Genetics, lifestyle, death, stroke, ischaemic heart disease, nutrition

Key Messages

- In 95 000–223 000 individuals, coffee intake was associated with U-shaped low risk of cardiovascular disease and all-cause mortality in observational analyses, but not in genetic analyses.
- Two studies from Copenhagen as well as Cardiogram and C4D international consortia were included in the genetic analyses.
- Thus, genetic analyses do not support the hypothesis that coffee intake influences risk of cardiovascular disease or all-cause mortality.

Introduction

Coffee is one of the most widely consumed beverages in the world. Multiple observational studies including large meta-analyses have reported associations between high coffee intake and low risk of cardiovascular disease (CVD) and of all-cause mortality.[1–8] However, as observational studies are often influenced by reverse causation and confounding, it is unclear whether these represent causal associations.

We tested the hypothesis that high coffee intake is associated with low risk of CVD and all-cause mortality using a Mendelian randomization study design. In this design, genetic variations are used as instruments for the exposure variable, and associations between the genetic variations and the outcome variables are thus used to examine associations between the exposure variable and the outcome variables. Since genetic variations are inherited randomly at conception, reverse causation and most confounding are avoided, exactly as in a randomized, double-blind controlled trial.[9] The design is based on several assumptions, including: (i) that the genetic variations are robustly associated with the exposure variable; (ii) that the genetic variations are not associated with possible confounding variables; and (iii) that the genetic variations are only related to the outcomes through the association with the exposure variable.[10] Specifically, we used five genetic variations near the *AHR* and *CYP1A1/2* genes involved in caffeine metabolism; these genetic variations have previously been associated with high caffeine intake in genome-wide association studies,[11–14] and since coffee intake is the major source of caffeine consumption in Denmark, we have previously used them to examine potential effects of coffee intake on type 2 diabetes, obesity, metabolic syndrome and components thereof.[15]

Using 95 366 to 112 509 individuals from three Danish cohorts, we tested first whether coffee intake is associated with CVD and all-cause mortality observationally; second, whether the five genetic variations near *AHR* and *CYP1A1/2* genes are associated with coffee intake; and third, whether the five genetic variations are associated with CVD and all-cause mortality indicating causal relationships; for ischaemic heart disease (IHD), we additionally included the Cardiogram and C4D consortia in a total of up to 223 414 individuals. To test our methods, we examined associations between IHD and *ApoE* genotype groups using similar analytical strategies, since *ApoE* genotypes are associated with plasma cholesterol levels and thus would be expected to be causally associated with risk of IHD.[16]

Methods

Participants

The Copenhagen General Population Study (CGPS) started in 2003, and the Copenhagen City Heart Study (CCHS)

started in 1976–78 with four follow-up examinations, were essentially conducted by the same investigators using identical methods.[17–19] At examination, all participants filled out a questionnaire which was reviewed by an investigator in attendance, and had a physical examination performed and blood samples taken for biochemical analyses and DNA extraction. We included 95 366 individuals from the CGPS examined in 2003–13, all with information on coffee intake and genetic variations, and 9743 individuals from the 1991–94 and 2001–03 examinations of the CCHS, all with information on genetic variations. The Copenhagen Ischaemic Heart Disease Study(CIHDS) included patients referred for coronary angiography to Copenhagen University Hospitals from 1989–2012, of whom we included 7400 individuals, all with information on genetic variations.[19,20] All individuals were Whites of Danish descent according to the Danish Central Person Registry. The studies were approved by Herlev and Gentofte Hospital and by Danish ethical committees, and were conducted according to the Declaration of Helsinki. Written informed consent was obtained from participants. Additionally and specifically for IHD (= coronary heart disease) we downloaded genetic risk estimates on up to 80 517 and 30 433 individuals from the Cardiogram and C4D consortia, respectively, as contributed by CARDIoGRAMplusC4D investigators at [http://www.cardiogramplusc4d.org/].[21,22]

In observational analyses to examine the association between coffee intake and risk of CVD and all-cause mortality, we included only the CGPS since information on coffee intake was not available for CCHS, CIHDS, Cardiogram or C4D (Supplementary Figure 1, available as Supplementary data at *IJE* online). Similarly, only the CGPS was used in the genetic analyses examining associations between the five genetic variations and coffee intake. In genetic analyses examining the association between the five genetic variations and CVD and all-cause mortality, we included the CGPS, CCHS and CIHDS but not Cardiogram and C4D, since the latter only included information on risk on IHD. For these analyses, we matched 14 800 controls from among individuals without IHD in the CGPS on sex and age 2:1 with CIHDS cases. Finally, to examine the association between the five genetic variations and IHD with maximum statistical power, we also included Cardiogram and C4D. To test associations between IHD and *ApoE* genotype groups, as a positive control we included the CGPS, CCHS, and CIHDS (not available in Cardiogram and C4D).

Coffee intake

Using the question 'What is your average weekly consumption of coffee(in cups)?', coffee intake at study entry was converted into cups/day.

Cardiovascular disease

We examined CVD with three endpoints: IHD, ischaemic stroke(IS) and the combination of the two as ischaemic vascular disease(IVD). Information on diagnoses of IHD (World Health Organization International Classification of Disease: ICD8 410–414; ICD10 I20–I25) and IS (ICD8 433–34; ICD10 I63) were obtained from the national Danish Patient Registry covering all hospital contacts in Denmark, validated as previously described,[18,23] and from the national Danish Causes of Death Registry (see below). We received register information on all individuals using the unique Danish Personal Identification Number from 1977 to end of follow-up November 2014, that is without losing track of even a single individual.

Cardiovascular and all-cause mortality

Information on death from CVD and from all causes was obtained from the national Danish Civil Registration System and the national Danish Cause of Death Registry. The national Danish Civil Registration System records date of death for people living in Denmark. The national Danish Cause of Death Registry ranks main causes of death and contributing causes of death reported by the attending physician in general practice or at a hospital, or by a physician in a forensic or pathology department.[24,25] We classified death from CVD as if one of the top three ranked causes of death was CVD (ICD8 390–458; ICD10 I00–99) and none had a diagnose of cancer (ICD8 140–209, ICD10 C00-C97), since cardiovascular complications could be caused by cancer. We followed each participant using the Personal Identification Number from study entrance to November 2014. The national Danish Causes of Death Registry lags behind the national Danish Civil Registration System by 23 months. Therefore, CVD mortality was only followed until December 2012, and participants with entry after December 2012 were censored in the CGPS ($n = 5003$) together with participants with no cause of death registered in the CGPS ($n = 1789$), CCHS ($n = 388$) and the CIHDS ($n = 563$) in analyses of CVD mortality. For mortality endpoints we likewise did not lose track of even a single individual.

Covariates

Income, alcohol intake, tea intake, cola intake, physical activity, use of antihypertensive medication, use of lipid-lowering medication and hormone replacement therapy were self-reported. Smoking history was captured as smoking status (never, former, current smokers), time since smoking cessation (former smokers only) and cumulated

smoking in pack-years (former and current smokers only); one pack-year was 20 cigarettes or equivalent smoked daily for 1 year. Physical inactivity was leisure time activity less than 4 h weekly and predominantly sedentary work. Plasma triglycerides, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol and glucose were measured using standard hospital assays. Systolic blood pressure, weight and height were measured at examination. Body mass index (BMI) was measured weight (kilograms) divided by measured height squared(metres squared). Diabetes was registry-based diagnosis of type 1 or 2 diabetes, self-reported diabetes, self-reported use of anti-diabetic medication and/or a non-fasting baseline plasma glucose level above 11 mmol/l. Less than 0.6% of covariates were missing. Missing values for systolic blood pressure, triglycerides, LDL cholesterol, HDL cholesterol and BMI were imputed using linear regression analysis with age and sex as predictors, and missing values for income and alcohol intake were grouped separately, when adjusted for in observational analyses; however, if analyses only included individuals with all covariates available, results were similar to those reported.

Genetic analyses

Participants were genotyped for two variations near the *AHR* gene (rs4410790 and rs6968865) and for three variations near the *CYP1A1/2* genes (rs2472297, rs2470893 and rs2472299), using *TaqMan*-based assays(Applied Biosystems Inc., Foster City, CA), including positive controls for each genotype verified by sequencing.[15] Due to reruns, call rates were $\geq$ 99.7%. Genotype distributions did not differ from Hardy-Weinberg equilibrium (χ^2 test, $P \geq 0.1$). As described previously, 76 069 participants were also genotyped for the *ApoE* polymorphisms (rs7412 and rs429358) which define six common *ApoE* genotypes associated with plasma cholesterol levels.[16,26]

Statistical analyses

We used Stata/SE 13.1. In observational analysis to examine the associations between coffee intake and IHD, IS, IVD, CVD mortality and all-cause mortality, we examined individuals from the CGPS using Cox proportional hazards regression models with age as underlying time scale (referred to as age-adjusted) and delayed entry at examination (left truncation), multivariable adjusted for CVD risk factors, including: sex; income; systolic blood pressure; plasma triglycerides; LDL cholesterol; HDL cholesterol; BMI; diabetes; smoking status; cumulated smoking in pack-years; time since smoking cessation in former smokers; alcohol intake; physical inactivity; use of antihypertensive medication; use of lipid-lowering medication and hormone replacement therapy in postmenopausal women. Totals of 5494 individuals with IHD prior to or at study entry, 1231 with IS and 6476 with IVD were excluded in these analyses (249 had both IHD and IS). Associations were examined using cubic restricted splines and, as categorical analyses of coffee intake, groups in steps of one cup/day and two cups/day. For cubic restricted splines, the number of knots between 3 and 7 were chosen to balance best fit and overfitting. To test for possible reverse causation, categorical analyses were also examined using delayed entry at 2 years after examination.

In genetic analyses to examine the effect of the five genetic variations on coffee intake, we calculated mean coffee intake according to genotype in the CGPS. We also grouped each individual according to number of caffeine alleles in an allele score.[15] Since the two *AHR* variations (rs6968865 and rs4410790) and the three *CYP1A1/2* (rs2470893, rs2472297 and rs2472299) are in linkage disequilibrium(Supplementary Figure 2, available as Supplementary data at *IJE* online), only rs4410790 and rs2472297 with the strongest association with coffee intake were included in the allele score. Test for trend was by Cuzick's extension of the Wilcoxon rank sum test. F- and R^2-values were estimated using linear regression models. We applied similar analytical strategies for effect of *ApoE* genotypes on plasma total cholesterol levels. *ApoE* genotypes were grouped as $\varepsilon 44 + \varepsilon 43$, $\varepsilon 33 + \varepsilon 42$ and $\varepsilon 32$ to ensure sufficient number of individuals in each group; $\varepsilon 22$ individuals were excluded as $\varepsilon 22$ may cause remnant hyperlipidaemia.

In genetic analyses to examine the association between genetically high coffee intake and risk of IHD, IS, IVD, CVD mortality and all-cause-mortality, we analysed the GCPS, CCHS and CIHDS. We used Cox proportional hazards regression models with age as underlying time scale adjusted for sex to estimate hazard ratios per caffeine intake allele. For IHD, IS and IHD, follow-up time was from 1977, and for CVD mortality and all-cause mortality from date of examination. Next, we performed random effect meta-analyses of the estimates.

Further, to examine the association between genetically high coffee intake and risk of IHD with maximum statistical power, we also included Cardiogram and C4D. We used logistic regression models adjusted for sex and age to estimate odds ratios per caffeine intake allele for the genetic variations separately in the CGPS, CCHS and the CIHDS, and performed random effect meta-analyses including downloaded risk estimates from Cardiogram and C4D. As a positive control, we applied similar analytical strategies for genetic associations between IHD and *ApoE* genotype groups in the CGPS, CCHS and CIHDS.

Finally, from the Cardiogram and CD4 consortia we examined six additional genetic variations associated with caffeine intake, as identified in a recent GWAS meta-analysis,[14] for risk of IHD. A funnel plot of IHD risk estimates as a function of allele effect size on coffee intake for all five genetic variations available in the Copenhagen General Population study and all 11 genetic variations available in Cardiogram (C4D excluded due to inconsistent data) were used to visually inspect for potential pleiotropic effects.[27]

Results

Baseline characteristics of the 95 366 individuals from the CGPS varied according to coffee intake (Table 1) and at most minimally according to caffeine intake allele score (Table 2); allele score increase associated directly with LDL cholesterol, BMI and alcohol intake, and inversely with tea intake (the latter likely as heavy coffee drinkers simply will not find time to drink tea). Allele score increase associated inversely with age, hormone replacement therapy in post-menopausal women and use of antihypertensive medication in coffee abstainers (Supplementary Table 1, available as Supplementary data at *IJE* online). In coffee-drinking individuals, allele score increase associated directly with LDL cholesterol and hormone replacement therapy in post-menopausal women, and inversely with tea intake (Supplementary Table 2, available as Supplementary data at *IJE* online).

Coffee intake, CVD and all-cause mortality: observational analyses

Among 95 366 CGPS individuals we observed 3822 IHD events during follow-up, 1708 IS events, 4973 IVD events, 971 CVD mortality events and 5422 all-cause mortality events. Median follow-up time was 5.8 years for IHD, 5.9 for IS, 5.8 for IVD, 6.2 for CVD mortality and 6.0 for all-cause mortality. Using cubic restricted splines, we observed U-shaped associations with extent of coffee intake for IHD, IS, IVD and all-cause mortality with lowest risks for individuals with medium coffee intake compared with individuals with no coffee intake (Figure 1). In categorical analyses, using coffee intake groups in steps of one cup/day and two cups/day, we observed similar U-shaped relationships (Figures 2 and 3); however, cardiovascular mortality appeared to be lower with higher coffee intake without a U-shaped relationship (Figure 3, available as Supplementary data at *IJE* online). When excluding events up to 2 years after study entry, results were also similar (Figure 2). Cubic spline models stratified for smoking status are shown in Supplementary Figure 3; as statistical power is lower in these analyses compared with the overall analyses, these results should be read and interpreted cautiously.

Genetic variations and coffee intake

For each of the genotypes we observed stepwise higher mean coffee intake from non-carriers through heterozygotes to homozygotes (P for trend $\leq 4 \times 10^{-19}$) (Figure 4). For caffeine intake allele score we observed stepwise higher mean coffee intake up to 42% from 0 to 4 alleles (P for trend across five allele score groups $= 3 \times 10^{-178}$). When stratified into current, former and never smokers, corresponding relative higher coffee intakes were 26%, 53% and 41%, respectively, reflecting the mean higher coffee intake among current smokers with similar absolute difference in coffee intake from 0 to 4 alleles among all strata (approximately 1 cup/day). F-value for statistical strength of the genetic instruments was 827 and R^2 as a measure of explained variation in coffee intake was 0.009. For grouped *ApoE* genotypes we observed a 10% lower mean plasma total cholesterol for $\varepsilon 32$ versus $\varepsilon 44 + \varepsilon 43$ (P for trend across three genotype groups $= 8 \times 10^{-300}$); F-value was 1394 and R^2 was 0.02. The per allele higher coffee intake was 9%, whereas the per ApoE genotype group lower plasma total cholesterol was 5%.

Coffee intake, CVD and all-cause mortality: genetic analyses

Among 112 509 individuals from the CGPS, CCHS and CIHDS, we observed 18 997 IHD events, 4589 IS events, 21 695 IVD events, 3671 CVD mortality events and 12 656 all-cause mortality events. Median follow-up time was 37.6 years for IHD, 37.8 for IS, 37.6 for IVD, 6.2 for CVD mortality and 6.0 for all-cause mortality. Per caffeine intake allele and 9% higher coffee intake meta-analysed hazard ratios were 1.02 (95% confidence interval: 1.00–1.03) for IHD, 1.02 (0.99–1.05) for IS, 1.02 (1.00–1.03) for IVD, 1.02 (0.99–1.06) for CVD mortality and 1.01 (0.99–1.03) for all-cause mortality (Figure 5).

When including Cardiogram and C4D together with the Copenhagen studies, up to 223 414 individuals and 55 689 IHD events/cases were examined. In meta-analyses, odds ratios per one caffeine intake allele were 1.00 (0.98–1.02) for *AHR* rs4410790, 1.01 (0.99–1.03) for *AHR* rs6968865, 1.02 (1.00–1.04) for *CYP1A1/2* rs2470893, 1.02 (1.00–1.04) for *CYP1A1/2* rs2472297 and 1.03 (0.99–1.06) for *CYP1A1/2* rs2472299 (Figure 6). In contrast, a 5% lower cholesterol level caused by *ApoE* genotype had a corresponding odds ratio for IHD of 0.93 (0.89–0.97) in the Copenhagen studies. For each genetic variation we had 80% power at two-sided $P < 0.05$ to

Table 1. Baseline characteristics of the 95 366 individuals from the Copenhagen General Population Study according to coffee intake

Coffee intake (cups/day)	0	–1	1–2	2–3	3–4	4–5	> 5	*P* for trend
No. of individuals	10 156	16 466	18 570	19 773	10 148	10 745	9509	
Age (years)	50 (42–62)	63 (51–72)	61 (50–70)	58 (48–67)	58 (49–67)	56 (48–64)	56 (49–64)	$4*10^{-26}$
Men (%)	34	38	38	46	48	57	63	$<1*10^{-300}$
Income below 400,000 DKK (%)	34	43	39	34	35	29	34	$3*10^{-60}$
Systolic blood pressure (mmHg)	134 (121–149)	140 (127–156)	140 (127–156)	140 (127–155)	140 (127–155)	140 (127–154)	140 (128–154)	$4*10^{-12}$
Plasma triglycerides (mmol/l)	1.31 (0.90–1.96)	1.40 (0.96–2.05)	1.38 (0.96–2.03)	1.39 (0.96–2.04)	1.39 (0.97–2.03)	1.40 (0.98–2.06)	1.48 (1.03–2.20)	$9*10^{-40}$
LDL cholesterol (mmol/l)	3.00 (2.40–3.60)	3.17 (2.50–3.80)	3.20 (2.59–3.80)	3.20 (2.60–3.85)	3.25 (2.64–3.90)	3.30 (2.70–3.90)	3.30 (2.70–4.00)	$6*10^{-152}$
HDL cholesterol (mmol/l)	1.51 (1.20–1.86)	1.61 (1.28–1.98)	1.61 (1.28–1.99)	1.58 (1.26–1.95)	1.58 (1.26–1.96)	1.53 (1.22–1.89)	1.47 (1.17–1.83)	$1*10^{-50}$
Body mass index (kg/m^2)	25.2 (22.7–28.4)	25.4 (23.0–28.3)	25.5 (23.0–28.4)	25.6 (23.3–28.4)	25.7 (23.3–28.4)	25.9 (23.6–28.6)	26.1 (23.9–28.8)	$1*10^{-81}$
Diabetes mellitius (%)	5	7	6	5	5	5	6	0.001
Ischaemic vascular disease (%)	5	9	8	7	6	6	6	$7*10^{-5}$
Current smokers (%)	12	14	14	15	18	23	36	$1*10^{-151}$
Cumulated smoking (pack-years)[a]	10 (4–23)	14 (4–30)	13 (4–27)	14 (5–27)	15 (6–29)	18 (7–7)	25 (13–39)	$<1*10^{-300}$
Time since smoking cessation (years)[b]	18 (8–29)	19 (9–31)	18 (8–30)	17 (8–27)	17 (7–28)	15 (6–26)	13 (5–24)	$6*10^{-100}$
Alcohol intake (g/week)	48 (12–120)	96 (36–168)	96 (48–180)	108 (48–192)	108 (48–192)	120 (60–204)	108 (48–204)	$1*10^{-253}$
Tea intake (cups/week)	10 (2–25)	3 (0–12)	5 (0–14)	2 (0–10)	1 (0–7)	0 (0–5)	0 (0–2)	$6*10^{-100}$
Cola and cola light drinkers (%)	22	10	11	13	13	15	15	0.06
Physical inactivity (%)	48	60	58	53	55	49	50	$3*10^{-44}$
Antihypertensive medication (%)	14	25	23	20	20	17	16	$6*10^{-32}$
Lipid-lowering medication (%)	7	14	13	12	12	10	11	0.06
Hormone replacement therapy (%)[c]	17	17	17	15	15	16	13	$8*10^{-9}$

Values are median (interquartile range) for continuous variables. To convert triglyceride values to mg/dl, multiply values in mmol/l by 88. To convert cholesterol values to mg/dl, multiply values in mmol/l by 38.6.
DKK, Danish Kroner as income per year; LDL, low-density lipoprotein; HDL, high-density lipoprotein.
[a]Current and former smokers only.
[b]Former smokers only.
[c]Post-menopausal women only.

Table 2. Baseline characteristics of the 95 366 individuals from the Copenhagen General Population Study according to allele score

Allele score	0	1	2	3	4	*P* for trend
No. of individuals	5844	26 691	39 302	20 263	3266	
Age (years)	58 (48–67)	58 (48–67)	58 (48–67)	58 (48–67)	58 (48–67)	0.56
Men (%)	45	45	45	45	46	0.98
Income below 400,000 DKK (%)	35	36	37	36	34	0.23
Systolic blood pressure (mmHg)	140 (126–155)	140 (126–155)	140 (126–155)	140 (126–155)	140 (126–153)	0.37
Plasma triglycerides (mmol/l)	1.38 (0.96–2.02)	1.39 (0.96–2.06)	1.39 (0.96–2.05)	1.40 (0.97–2.05)	1.39 (0.96–2.01)	0.28
LDL cholesterol (mmol/l)	3.20 (2.56–3.80)	3.20 (2.60–3.81)	3.20 (2.60–3.80)	3.20 (2.60–3.90)	3.24 (2.60–3.88)	$6*10^{-5}$
HDL cholesterol (mmol/l)	1.57 (1.25–1.95)	1.57 (1.25–1.94)	1.57 (1.25–1.93)	1.55 (1.24–1.94)	1.56 (1.25–1.93)	0.07
Body mass index (kg/m^2)	25.5 (23.2–28.4)	25.6 (23.2–28.4)	25.6 (23.2–28.5)	25.6 (23.3–28.5)	25.6 (23.3–28.5)	0.02
Diabetes mellitius (%)	5	7	7	7	6	0.18
Ischaemic vascular disease (%)	5	5	6	5	5	0.61
Current smokers (%)	18	18	18	18	19	0.22
Cumulated smoking (pack-years)[a]	15 (5–30)	15 (5–30)	15 (5–30)	15 (5–30)	15 (5–30)	0.42
Time since smoking cessation (years)[b]	17 (7–29)	17 (7–28)	17 (7–28)	17 (7–28)	17 (8–30)	0.43
Alcohol intake (g/week)	96 (48–180)	96 (48–180)	96 (48–180)	96 (48–180)	108 (48–192)	0.001
Tea intake (cups/week)	3 (0–10)	3 (0–10)	2 (0–10)	2 (0–10)	2 (0–10)	$4*10^{-17}$
Cola and cola light drinkers (%)	14	13	13	13	14	0.75
Physical inactivity (%)	54	54	54	55	54	0.33
Antihypertensive medication (%)	20	20	20	20	20	0.38
Lipid-lowering medication (%)	11	12	12	12	12	0.02
Hormone replacement therapy (%)[c]	15	16	16	17	17	0.12

Values are median (interquartile range) for continuous variables. To convert triglyceride values to mg/lL, multiply values in mmol/l by 88. To convert cholesterol values to mg/dl, multiply values in mmol/l by 38.6.

DKK, Danish Kroner as income per year; LDL, low-density lipoprotein; HDL, high-density lipoprotein.

[a]Current and former smokers only.

[b]Former smokers only.

[c]Post-menopausal women only.

detect a causal odds ratio per allele of 0.97 for the approximately 8% higher genetic coffee intake per allele, compared with an observed hazard ratio of 0.86 for an approximately 350% higher observational coffee intake (0 cups/day versus 4–5 cups/day).

Sensitivity analyses

When individuals with a coffee intake of 0–1 cups/day where used as reference group in observational analyses, results were similar to those presented in Figure 2. In genetic analyses, when adjusting for smoking status in the CGPS since this might influence CYP1A1/2 activity,[11,13,28,29] results were similar to those presented in Figure 5. Genetic results were also similar when adjusting for LDL cholesterol, BMI, alcohol intake and tea intake in all individuals, when adjusting for age, hormone replacement therapy in post-menopausal women and use of antihypertensive medication in coffee-abstaining individuals, and when adjusting for LDL cholesterol, tea intake and hormone replacement therapy in post-menopausal women in coffee drinkers (data not shown). Further, when stratifying individuals from the GCPS into coffee abstainers, coffee drinkers, coffee drinkers excluding tea and cola drinkers, genetic results were similar (Figure 7).

Since the completion of our study an additional six genetic variations have been shown to associate with caffeine intake. We have included associations of these genetic variations with IHD from Cardiogram and C4D consortia in Supplementary Figure 4, available as Supplementary data at *IJE* online. Further, we have included odds ratios for IHD per effect allele as a function of coffee intake increment per effect allele for all available genetic variations in the CGPS and the Cardiogram consortium (C4D excluded due to inconsistent data) to examine for potential pleiotropic effects in Figure 8; on visual inspection, no clear evidence for pleiotropic effects was evident.

Discussion

As examined in 95 000–223 000 individuals, observationally coffee intake was associated with U-shaped low risk of cardiovascular disease and all-cause mortality; however, genetically coffee intake was not associated with risk of cardiovascular disease or all-cause mortality. The latter are novel findings.

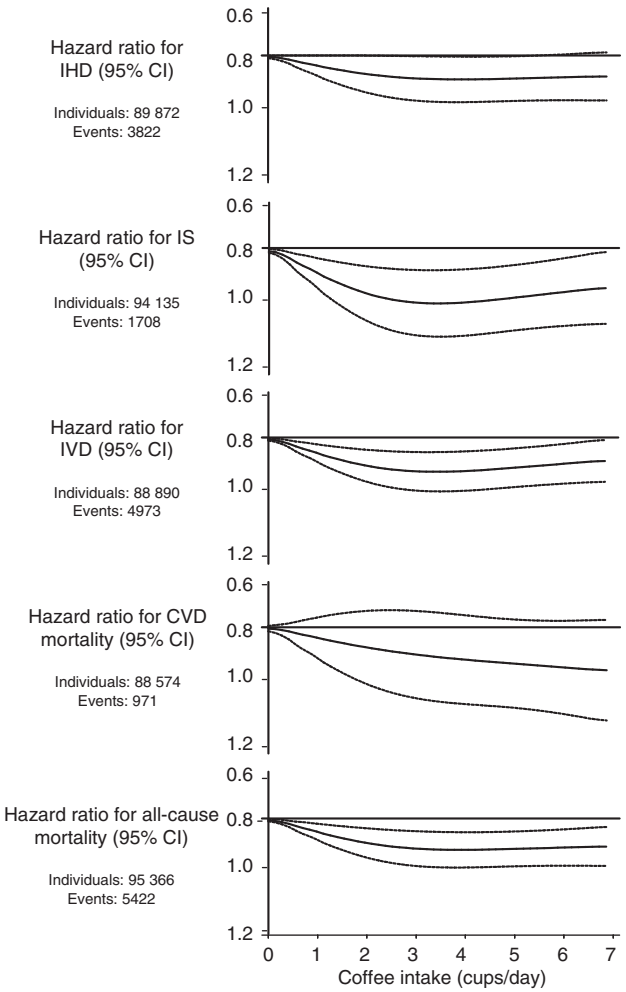

Figure 1. Hazard ratios for cardiovascular disease and all-cause mortality according to coffee intake as cubic restricted spline models. Analyses included 95 366 individuals from the Copenhagen General Population Study: 5494 individuals with IHD, 1231 with IS and 6476 with IVD prior to or at study entry were excluded in the analyses. Also, 6792 individuals with study entry after December 2012 or with no registered cause of death were excluded from the analyses of CVD mortality. Hazard ratios were adjusted for age, sex, income, systolic blood pressure, plasma triglycerides, LDL cholesterol, HDL cholesterol, BMI, diabetes, smoking status, cumulated smoking in pack-years, time since smoking cessation in former smokers, alcohol intake, physical inactivity, use of antihypertensive medication, use of lipid-lowering medication and hormone replacement therapy in post-menopausal women. IHD, ischaemic heart disease; IS, ischaemic stroke; IVD, ischaemic vascular disease; CVD, cardiovascular disease; HR, hazard ratio; CI, confidence interval.

U-shaped associations between coffee intake and CVD and all-cause mortality have previously been described in observational studies.[1,2,5,7] Coffee is a complex mixture, and several biological mechanisms have been examined to explain both beneficial and harmful effects of coffee intake on risk of CVD and all-cause mortality.[1,6,30,34] It has been proposed that potential beneficial effects are dominant with moderate coffee intake, whereas harmful effects increase with high coffee intake, which produces the observed U-shaped association with CVD. As also suggested elsewhere,[35] associations could be heavily confounded by smoking as U-shaped associations are less prominent in never smokers compared with former and current smokers (Supplementary Figure 3, available as Supplementary data at *IJE* online). In line with this, we believe that our genetic results also suggest confounding as possible explanations for the previously observed 'beneficial' associations with moderate coffee intake, and thus that coffee intake is unlikely to influence risk of CVD or all-cause mortality. Since observational results were similar when excluding events up to 2 years after study entry, reverse causation is less likely to explain any differences between observational and genetic results.

Strengths of our study include the large sample size of individual participant data and the fact that all participants in the Copenhagen studies were Whites of Danish descent, which ensures a homogeneous population and excludes genetic admixture. Conversely, this might limit the generalizability of our results, although it is reassuring that the findings for Cardiogram and C4D, including individuals from other countries,[21,22] were similar to those for the Copenhagen studies. Another strength of the Copenhagen studies is that we did not lose track of even a single individual. Further, it is notable that mean coffee intake in individuals with respectively 0 and 4 caffeine intake alleles were 2.2 and 3.1 cups/day, thus a 42% higher coffee intake, and that the statistical F-value for allele score was very high at 827, suggesting large statistical power to exclude associations of coffee intake with the endpoints studies. In support, genetic variations with comparable effect sizes on intermediate variables have been used in Mendelian randomization studies to document causal associations of BMI and smoking on endpoints.[36,39] Finally, it is also reassuring that in our studies low cholesterol level caused by *ApoE* genotype associated causally with the expected low risk of IHD.

Limitations include pleiotropic effects of the genetic variations, since *AHR* and *CYP1A1/2* genes are involved in the metabolism and regulation of several components besides caffeine, although our use of five different genetic variations close to three different genes with similar results argues against this as a major problem.[10] Also, Figure 8 using a funnel plot of 11 different caffeine alleles provided no clear evidence for pleotropic effects. In our study, higher caffeine intake allele score was associated positively with LDL cholesterol, BMI and alcohol intake, and inversely with tea intake, but when adjusting for these potential confounding variables, genetic results were similar to those in the overall analyses. That said, it is interesting that genetic analysis suggests that higher coffee intake causes a slightly higher levels of LDL cholesterol, although the

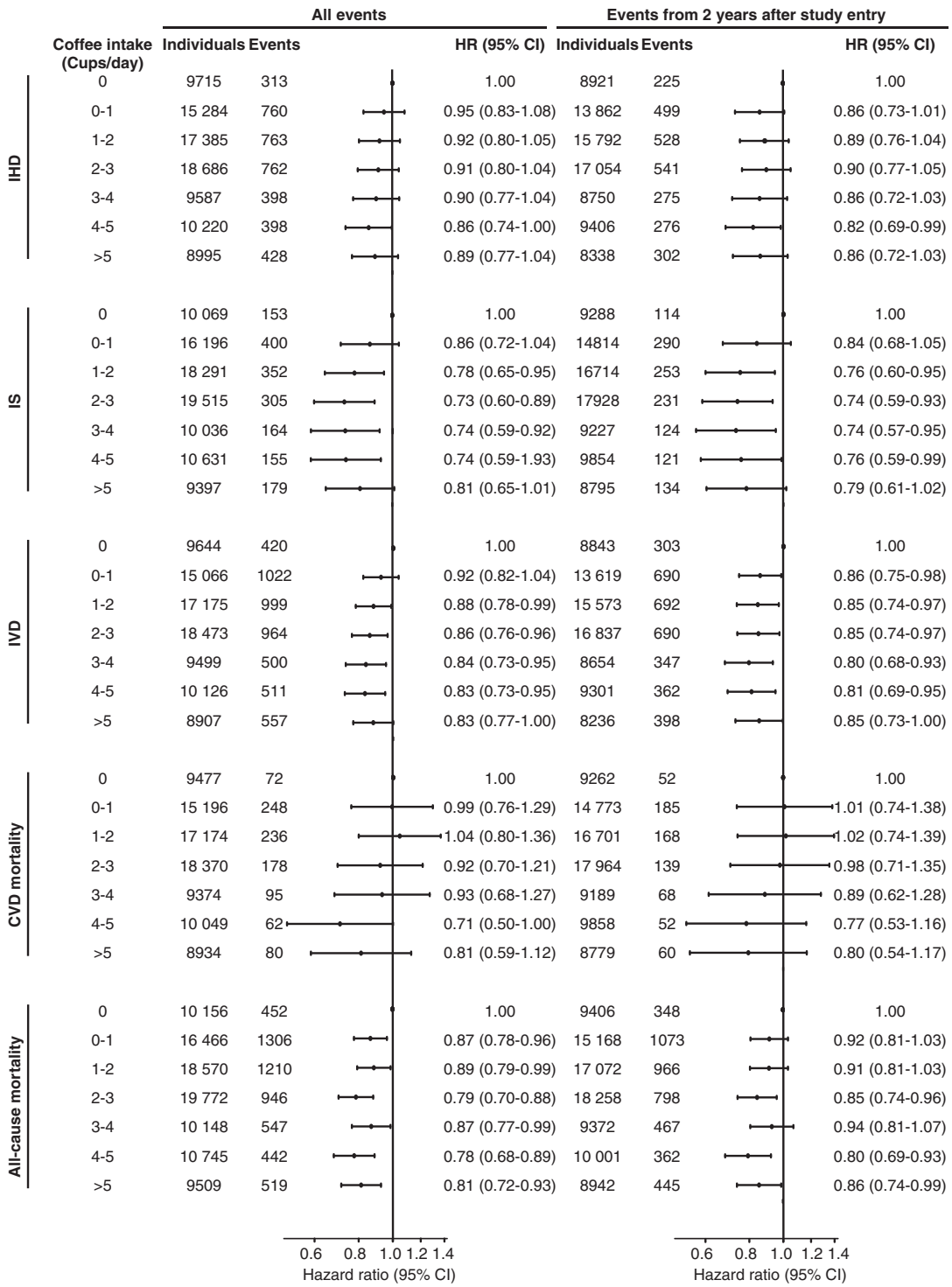

Figure 2. Hazard ratios for cardiovascular disease and all-cause mortality according to coffee intake in steps of one cup/day, including all events and after excluding events occurring within 2 years of examination. Analyses included 95 366 individuals from the Copenhagen General Population Study: 5494 individuals with IHD, 1231 with IS and 6476 with IVD prior to or at study entry were excluded in the analyses including all events; corresponding numbers after excluding events occurring within 2 years of examination were 13 243, 8746 and 14 303, respectively, that is after additionally excluding events occurring prior to or at 2 years after study entry or exclusion of individuals with less than 2 years' follow-up time. Also, 6792 (8840) individuals with study entry after December 2012 or with no registered cause of death (or CVD mortality prior to or at 2 years after study entry or less than 2 years' follow-up time) were excluded from the analyses of CVD mortality, and 7237 individuals with mortality prior to or at 2 years after study entry or less than 2 years' follow-up time were excluded in lag analyses of all-cause mortality. Hazard ratios were adjusted for age, sex, income, systolic blood pressure, plasma triglycerides, LDL cholesterol, HDL cholesterol, BMI, diabetes, smoking status, cumulated smoking in pack-years, time since smoking cessation in former smokers, alcohol intake, physical inactivity, use of antihypertensive medication, use of lipid-lowering medication and hormone replacement therapy in post-menopausal women. IHD, ischaemic heart disease; IS, ischaemic stroke; IVD, ischaemic vascular disease; CVD, cardiovascular disease; HR, hazard ratio; CI, confidence interval.

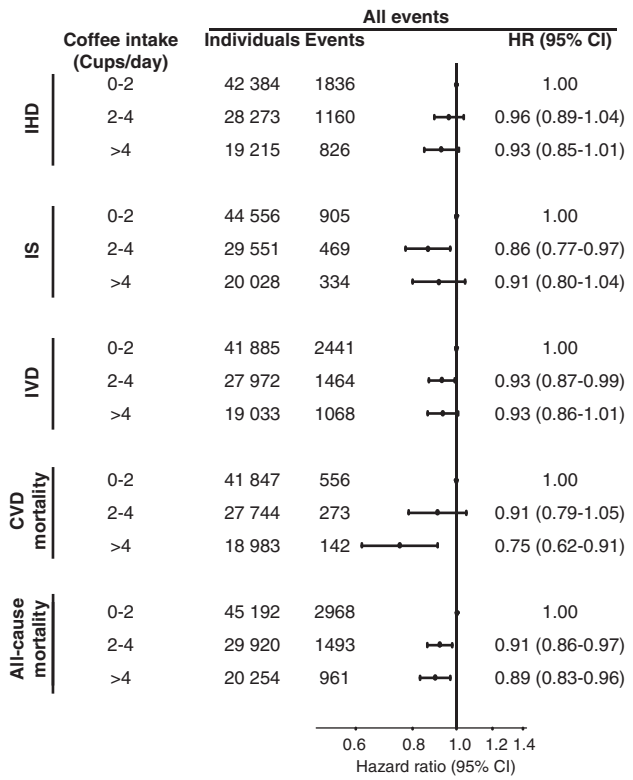

Figure 3. Hazard ratios for cardiovascular disease and all-cause mortality according to coffee intake in steps of two cups/day. Analyses included 95 366 individuals from the Copenhagen General Population Study: 5494 individuals with IHD, 1231 with IS and 6476 with IVD prior to or at study entry were excluded in the analyses. Also, 6792 individuals with study entry after December 2012 or with no registered cause of death were excluded from the analyses of CVD mortality. Hazard ratios were adjusted for age, sex, income, systolic blood pressure, plasma triglycerides, LDL cholesterol, HDL cholesterol, BMI, diabetes, smoking status, cumulated smoking in pack-years, time since smoking cessation in former smokers, alcohol intake, physical inactivity, use of antihypertensive medication, use of lipid-lowering medication and hormone replacement therapy in post-menopausal women. IHD, ischaemic heart disease; IS, ischaemic stroke; IVD, ischaemic vascular disease; CVD, cardiovascular disease; HR, hazard ratio; CI, confidence interval.

extent of LDL cholesterol increase would be considered to be clinically unimportant.

Since the five genetic variants are associated with caffeine consumption and not only coffee consumption, genetic results could be confounded by other caffeine-containing beverages such as tea and cola. Reassuringly though, coffee was the major source of dietary caffeine in terms of quantity with mean coffee, tea and cola intakes in the CGPS of 2.6 cups/day, 1.1 cups/day and 58 ml/day, respectively. Further, approximate caffeine contents for coffee, tea and cola are 85, 25 and 15 mg/dl, respectively, in Denmark,[40] and the present genetic results were similar when excluding tea and cola drinkers. In other countries with other patterns of beverage consumption, tea and soft drinks might be a larger source of caffeine intake, and thus results from Cardiogram and C4D consortia are more likely to be influenced from other caffeine sources. It would be informative to perform similar Mendelian randomization analyses in, for example, the UK to examine whether high tea intake is associated with CVD or all-cause mortality and a possible confounding variable. Unfortunately, we do not have the data to distinguish between caffeinated and decaffeinated coffee in our cohorts. However, in Denmark decaffeinated coffee comprises only 0.5% of total coffee import, and thus is not a common beverage.

Another concern is linkage disequilibrium; however, allele score included only two genetic variations with no linkage in between. In our study we did not ask participants about type, strength or size of the cups of coffee, and it is therefore possible that individuals with high coffee intake drink bigger coffee servings of stronger coffee, compared with those with low coffee intake. Observationally, such a scenario would likely lead to overestimated associations. In genetic analyses, such inconsistency would most likely affect mean coffee intake by genotype, but would not influence the association between caffeine intake allele score and risk of CVD or all-cause mortality. Additionally, if U-shaped associations between coffee intake and CVD and all-cause mortality indeed are true, then a Mendelian randomization approach might not detect an eventual association, since this approach is based on the assumption of linearity between all coffee intake categories and thus will not be capturing non-linear differences between very low and very high coffee intakes. Finally, it is a limitation that we did not include all genetic variations previously associated with coffee intake in all of our genetic analyses;[14] however, these additional genetic variations would only capture a small proportion of additional variance compared with the five genetic variants already included in our main analyses.

In coffee-abstaining individuals from the CGPS, allele score increase was associated inversely with age, use of antihypertensive medication and hormone replacement therapy in post-menopausal women (Supplementary Table 1, available as Supplementary data at *IJE* online). These associations are likely explained by collider stratification bias, since age is strongly associated directly with coffee intake among coffee consumers only, use of antihypertensive medication and hormone replacement therapy. In coffee drinkers, direct associations with LDL cholesterol and hormone replacement therapy were observed. Interestingly, direct associations with LDL cholesterol were more pronounced in coffee drinkers than in all individuals (compare Table 2 and Supplementary Table 2), suggesting that coffee intake has an LDL cholesterol- and perhaps a minor BMI-increasing effect. Associations with hormone replacement therapy could be explained by possible pleiotropic effects of

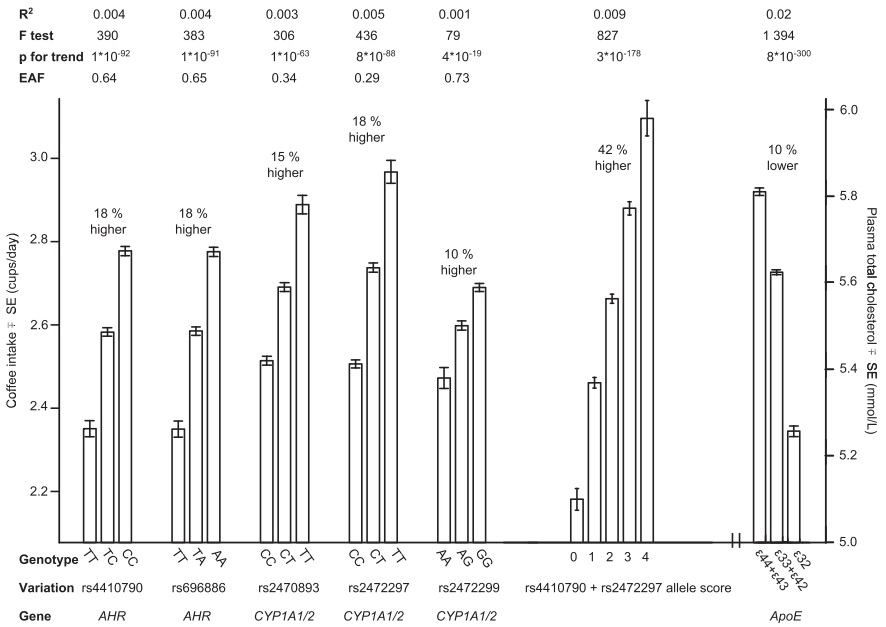

Figure 4. Coffee intake according to caffeine intake genotypes and caffeine intake allele score, and plasma total cholesterol according to *ApoE* genotype groups. Analyses included 95 366 individuals from the Copenhagen General Population Study. Values represent mean ∓ standard error. *P*-values are for trend using Cuzick's extension of the Wilcoxon rank sum test. F test is for statistical strength of the genetic instruments. R^2 is a measure of explained variation in coffee intake by genotype or allele score. Plasma total cholesterol levels by *ApoE* genotype groups were added as a positive control. SE, standard error; EAF, exposure allele frequency.

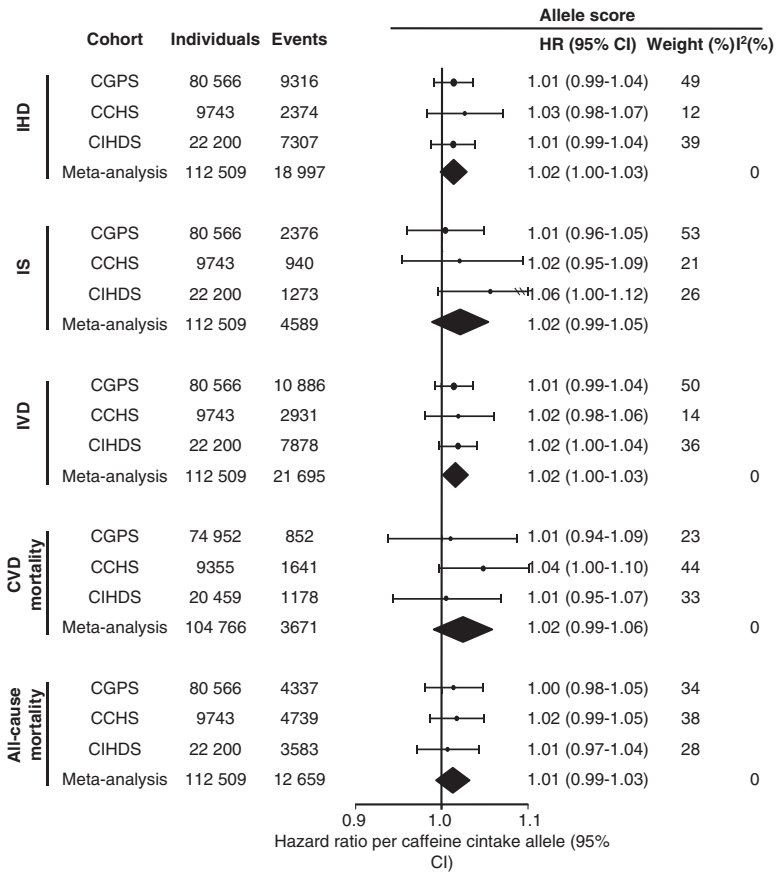

Figure 5. Hazard ratios per caffeine intake allele for cardiovascular disease and all-cause mortality, including rs441079 and rs2472297. Analyses included 112 509 individuals from the Copenhagen General Population Study (CGPS), Copenhagen City Heart Study(CCHS) and Copenhagen Ischaemic Heart Disease study(CIHDS) separately and combined as random effect meta-analysed hazard ratios; 7743 individuals with study entry after December 2012 or with no registered cause of death were excluded from analyses of CVD mortality. Hazard ratios were adjusted for sex and for age used as underlying intensity. I^2 represents percentage heterogeneity between studies. IHD, ischaemic heart disease; IS, ischaemic stroke; IVD, ischaemic vascular disease; CVD, cardiovascular disease; HR, hazard ratio; CI, confidence interval.

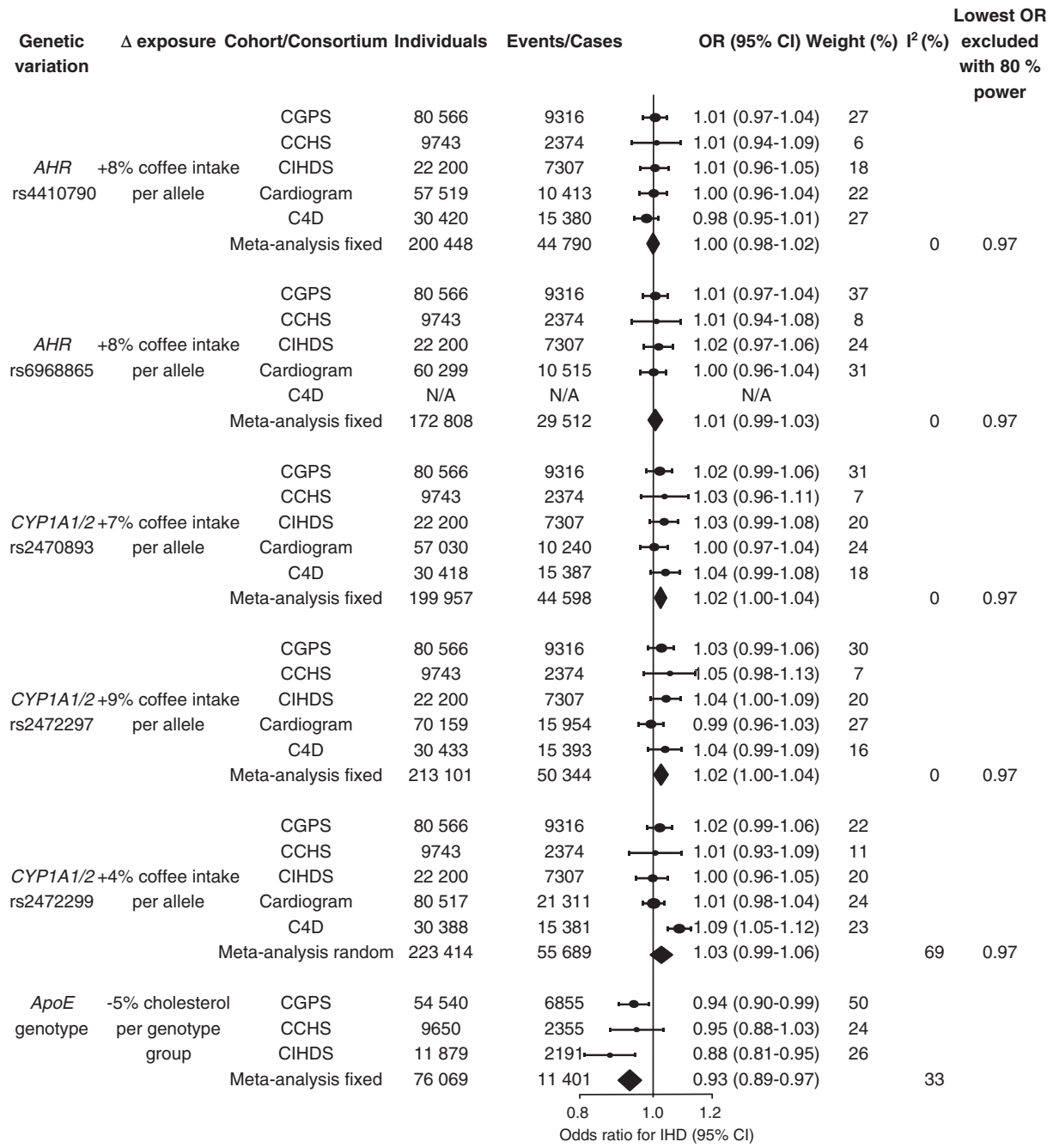

Figure 6. Odds ratios per caffeine intake allele for IHD for each of the five genetic variations and for *ApoE* genotype groups affecting cholesterol levels. Analyses included up to 223 414 individuals from the Copenhagen General Population Study(CGPS), Copenhagen City Heat Study(CCHS), Copenhagen Ischaemic Heart Disease Study(CIHDS), Cardiogram and C4D separately and combined into random effect meta-analysed odds ratios. Numbers of individuals and IHD events/cases vary according to availability of data from Cardiogram and C4D. Odds ratios were adjusted for sex and age in the Copenhagen studies and for sex, age and study-specific covariates in Cardiogram and C4D. I^2 represents percentage heterogeneity between studies. Odds ratios per *ApoE* genotype group were added as a positive control. IHD, ischaemic heart disease; OR, odds ratio. CI, confidence interval.

the genetic variations on estrogen metabolism and signalling.[41] Tea and alcohol intake were associated inversely with allele score increase in all individuals, tea intake was associated inversely with allele score increase in coffee drinkers, and neither was associated with allele score of coffee abstainers. This suggests that tea and alcohol intake are related to coffee-drinking habits, that is coffee drinkers drink less tea and more alcohol in a Danish context.

Although the present study does not provide genetic evidence that coffee intake is causally associated with risk of CVD and all-cause mortality, it is not possible to entirely negate the hypothesis. The genetic associations with coffee

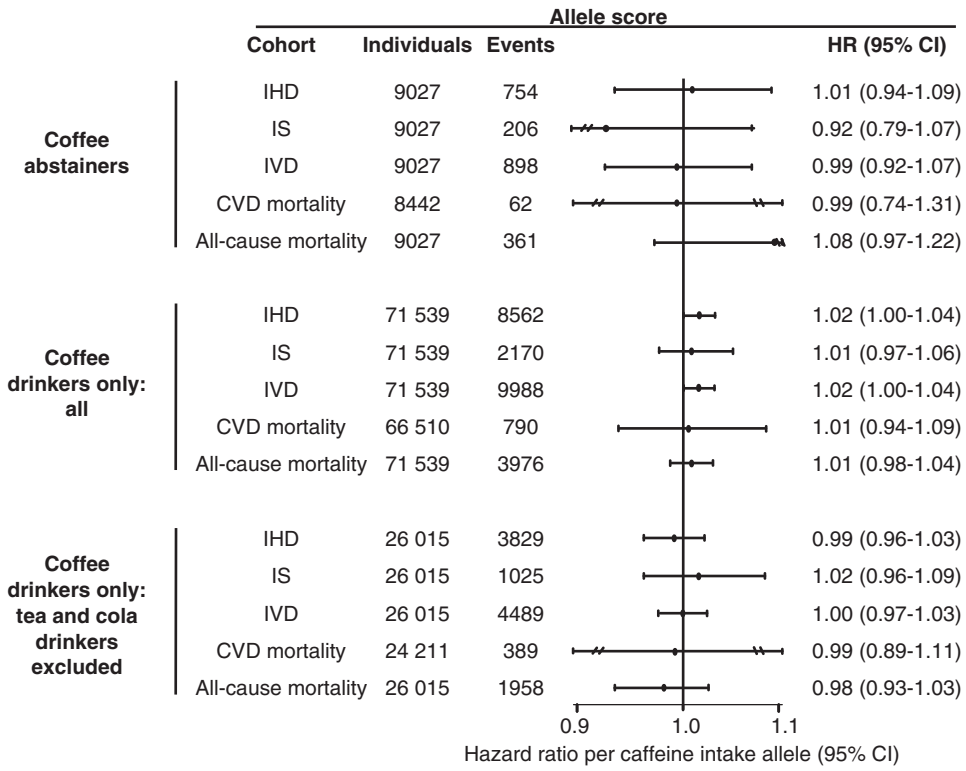

Figure 7. Hazard ratios per caffeine intake allele score of rs441079 and rs2472297 for cardiovascular disease and all-cause mortality in coffee abstainers, coffee drinkers and coffee drinkers excluding tea and cola drinkers. Analyses included 80 566 individuals from the Copenhagen General Population Study(CGPS) including 45 524 tea and/or cola drinkers; 5614 individuals with study entry after December 2012 or with no registered cause of death were excluded from analyses of CVD mortality. Hazard ratios were adjusted for sex and for age used as underlying intensity. IHD, ischaemic heart disease; IS, ischaemic stroke; IVD, ischaemic vascular disease; CVD, cardiovascular disease; HR, hazard ratio; CI, confidence interval.

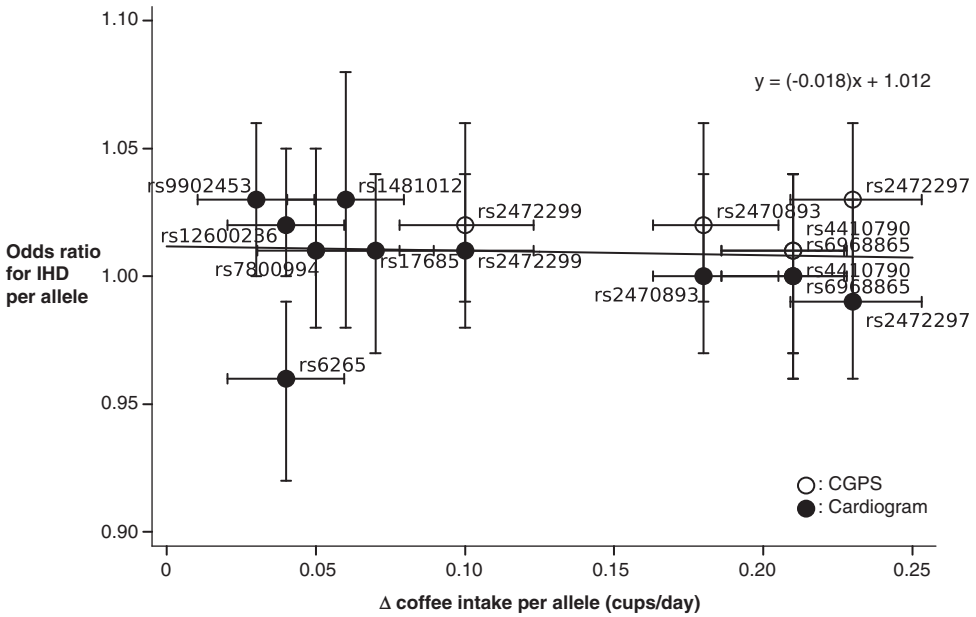

Figure 8. Odds ratios for IHD per caffeine intake effect allele as a function of coffee intake increment per effect allele. Analyses included 95 366 individuals from the Copenhagen General Population Study(CGPS) and 54 802–81 013 individuals from the Cardiogram consortium (C4D excluded due to inconsistent data). For rs1260236, rs1481012, rs7800944, rs17685, rs6265 and rs9902453, coffee intake increment per effect allele is from CornelisMC, Byrne EM, Esko T *et al.* Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol Psychiatry* 2015;**20**:647–56. Linear regressions for CGPS available genetic variations separately and for Cardiogram available genetic variations separately were $y = (-0.004)x + 1.019$ and $y = (-0.077)x + 1.014$, respectively. IHD, ischaemic heart disease.

intake used in the present study are relatively small (we could exclude odds ratio per allele of 0.97 for 8% higher coffee intake with 80% power and two-sided $P < 0.05$) compared with observational differences in coffee intake (we observed odds ratio of 0.86 for approximately 350% higher coffee intake), and thus even more individuals and events are required to thoroughly exclude a causal association from moderate coffee intake to reduced risk of CVD and all-cause mortality using genetic instruments. With a 42% higher coffee intake, from 0 to 4 alleles a 1:1 case-control study size of approximately 225 000 cases and 225 000 controls would be needed to exclude the observed hazard ratio of 0.86 for an approximately 350% higher coffee intake (0 cups/day versus 4–5 cups/day), and thus to exclude a potential effect of coffee intake on CVD and all-cause mortality at two-sided $P < 0.05$ and 80% power.[42]

In conclusion, genetic coffee intake was not associated with risk of CVD or all-cause mortality and does therefore not support the hypothesis that coffee intake influences risk of CVD and all-cause mortality.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

This work was supported by Herlev and Gentofte Hospital, Copenhagen University Hospital and the Danish Heart Foundation. The funding organizations had no role in the design and conduct of the study, the collection, analysis and interpretation of the data or in the writing of the paper.

Author Contributions

A.T.N. and B.G.N. designed the study and gathered, analysed and interpreted the data. A.T.N. generated laboratory data, performed statistical analyses, drafted the paper and prepared display items, all supervised by B.G.N. Both authors had full access to all data, and both revised and finally approved the manuscript before submission. B.G.N. had the final responsibility to submit for publication.

Conflict of interest: None declared.

References

1. Ding M, Bhupathiraju SN, Satija A, van Dam RM, Hu FB. Long-term coffee consumption and risk of cardiovascular disease: a systematic review and a dose-response meta-analysis of prospective cohort studies. *Circulation* 2014;**129:**643–59.
2. Freedman ND, Park Y, Abnet CC, Hollenbeck AR, Sinha R. Association of coffee drinking with total and cause-specific mortality. *N Engl J Med* 2012;**366:**1891–904.
3. Grioni S, Agnoli C, Sieri S *et al*. Espresso coffee consumption and risk of coronary heart disease in a large Italian cohort. *PLoS One* 2015;**10:**e0126550.
4. Je Y, Giovannucci E. Coffee consumption and total mortality: a meta-analysis of twenty prospective cohort studies. *Br J Nutr* 2014;**111:**1162–73.
5. Malerba S, Turati F, Galeone C *et al*. A meta-analysis of prospective studies of coffee consumption and mortality for all causes, cancers and cardiovascular diseases. *Eur J Epidemiol* 2013;**28:**527–39.
6. Rebello SA, van Dam RM. Coffee consumption and cardiovascular health: getting to the heart of the matter. *Curr Cardiol Rep* 2013;**15:**403.
7. Crippa A, Discacciati A, Larsson SC, Wolk A, Orsini N. Coffee consumption and mortality from all causes, cardiovascular disease, and cancer: a dose-response meta-analysis. *Am J Epidemiol* 2014;**180:**763–75.
8. Larsson SC, Orsini N. Coffee consumption and risk of stroke: a dose-response meta-analysis of prospective studies. *Am J Epidemiol* 2011;**174:**993–1001.
9. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?. *Int J Epidemiol* 2003;**32:**1–22.
10. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;**27:**1133–63.
11. Sulem P, Gudbjartsson DF, Geller F *et al*. Sequence variants at CYP1A1-CYP1A2 and AHR associate with coffee consumption. *Hum Mol Genet* 2011;**20:**2071–77.
12. Cornelis MC, Monda KL, Yu K *et al*. Genome-wide meta-analysis identifies regions on 7p21 (AHR) and 15q24 (CYP1A2) as determinants of habitual caffeine consumption. *PLoS Genet* 2011;**7:**e1002033.
13. Rodenburg EM, Eijgelsheim M, Geleijnse JM *et al*. CYP1A2 and coffee intake and the modifying effect of sex, age, and smoking. *Am J Clin Nutr* 2012;**96:**182–87.
14. Cornelis MC, Byrne EM, Esko T *et al*. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol Psychiatry* 2015;**20:**647–56.
15. Nordestgaard AT, Thomsen M, Nordestgaard BG. Coffee intake and risk of obesity, metabolic syndrome and type 2 diabetes: a Mendelian randomization study. *Int J Epidemiol* 2015;**44:**551–65.
16. Frikke-Schmidt R, Tybjaerg-Hansen A, Steffensen R, Jensen G, Nordestgaard BG. Apolipoprotein E genotype: epsilon32 women are protected while epsilon43 and epsilon44 men are susceptible to ischaemic heart disease: the Copenhagen City Heart Study. *J Am Coll Cardiol* 2000;**35:**1192–99.
17. Frikke-Schmidt R, Nordestgaard BG, Stene MC *et al*. Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischaemic heart disease. *JAMA* 2008;**299:**2524–32.
18. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *JAMA* 2009;**301:**2331–39.

19. Zacho J, Tybjaerg-Hansen A, Jensen JS *et al*. Genetically elevated C-reactive protein and ischaemic vascular disease. *N Engl J Med* 2008;**359:**1897–908.
20. Afzal S, Brondum-Jacobsen P, Bojesen SE, Nordestgaard BG. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. *BMJ* 2014;**349:**g6330.
21. Schunkert H, Konig IR, Kathiresan S *et al*. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 2011;**43:**333–38.
22. Peden JF, Hopewell JC, Saleheen D *et al*. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* 2011;**43:**339–44.
23. Varbo A, Nordestgaard BG, Tybjaerg-Hansen A *et al*. Nonfasting triglycerides, cholesterol, and ischaemic stroke in the general population. *Ann Neurol* 2011;**69:**628–34.
24. Helweg-Larsen K. The Danish Register of Causes of Death. *Scand J Public Health* 2011;**39:**26–29.
25. Nielsen SF, Nordestgaard BG, Bojesen SE. Statin use and reduced cancer-related mortality. *N Engl J Med* 2012;**367:** 1792–802.
26. Rasmussen KL, Tybjaerg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. Plasma levels of apolipoprotein E and risk of dementia in the general population. *Ann Neurol* 2015;**77:**301–11.
27. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;**44:**512–25.
28. Dobrinas M, Cornuz J, Oneda B *et al*. Impact of smoking, smoking cessation, and genetic polymorphisms on CYP1A2 activity and inducibility. *Clin Pharmacol Ther* 2011;**90:**117–25.
29. Guessous I, Eap CB, Bochud M. Blood pressure in relation to coffee and caffeine consumption. *Curr Hypertens Rep* 2014;**16:**468.
30. Greenberg JA, Boozer CN, Geliebter A. Coffee, diabetes, and weight control. *Am J Clin Nutr* 2006;**84:**682–93.
31. Hartley TR, Lovallo WR, Whitsett TL. Cardiovascular effects of caffeine in men and women. *Am J Cardiol* 2004;**93:**1022–26.
32. Lopez-Garcia E, van Dam RM, Qi L, Hu FB. Coffee consumption and markers of inflammation and endothelial dysfunction in healthy and diabetic women. *Am J Clin Nutr* 2006;**84:** 888–93.
33. Svilaas A, Sakhi AK, Andersen LF *et al*. Intakes of antioxidants in coffee, wine, and vegetables are correlated with plasma carotenoids in humans. *J Nutr* 2004;**134:**562–67.
34. Mesas AE, Leon-Munoz LM, Rodriguez-Artalejo F, Lopez-Garcia E. The effect of coffee on blood pressure and cardiovascular disease in hypertensive individuals: a systematic review and meta-analysis. *Am J Clin Nutr* 2011;**94:**1113–26.
35. Ding M, Satija A, Bhupathiraju SN *et al*. Association of coffee consumption with total and cause-specific mortality in 3 large prospective cohorts. *Circulation* 2015;**132:**2305–15.
36. Timpson NJ, Harbord R, Davey Smith G *et al*. Does greater adiposity increase blood pressure and hypertension risk? Mendelian randomization using the FTO/MC4R genotype. *Hypertension* 2009;**54:**84–90.
37. Tyrrell J, Huikari V, Christie JT *et al*. Genetic variation in the 15q25 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNB4) interacts with maternal self-reported smoking status during pregnancy to influence birth weight. *Hum Mol Genet* 2012;**21:**5344–58.
38. Nordestgaard BG, Palmer TM, Benn M *et al*. The effect of elevated body mass index on ischaemic heart disease risk: causal estimates from a Mendelian randomisation approach. *PLoS Med* 2012;**9:**e1001212.
39. Winslow UC, Rode L, Nordestgaard BG. High tobacco consumption lowers body weight: a Mendelian randomization study of the Copenhagen General Population Study. *Int J Epidemiol* 2015;**44:**540–50.
40. Tobberup R, Kjeldsen HC, Baggersgaard R *et al*. 2015. https://www.sundhed.dk/borger/sygdomme-a-aa/sundhedsoplysning/kost/koffein/ (5 July 2016, date last accessed).
41. Labrecque MP, Takhar MK, Hollingshead BD *et al*. Distinct roles for aryl hydrocarbon receptor nuclear translocator and AH receptor in estrogen-mediated signaling in human cancer cell lines. *PLoS One* 2012;**7:**e29545.
42. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol* 2014;**43:**922–29.