International Journal of Epidemiology, 2015, 551–565
doi: 10.1093/ije/dyv083
Advance Access Publication Date: 22 May 2015
Original Article

Mendelian Randomization Causal Analysis

Coffee intake and risk of obesity, metabolic syndrome and type 2 diabetes: a Mendelian randomization study

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

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

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

Accepted 23 April 2015

Abstract

Background: Coffee is one of the most widely consumed beverages. We tested the hypothesis that genetically high coffee intake is associated with low risk of obesity, metabolic syndrome and type 2 diabetes, and with related components thereof.

Methods: We included 93 179 individuals from two large general population cohorts in a Mendelian randomization study. We tested first whether high coffee intake is associated with low risk of obesity, metabolic syndrome and type 2 diabetes, and with related components thereof, in observational analyses; second, whether five genetic variants near the *CYP1A1*, *CYP1A2* and *AHR* genes are associated with coffee intake; and third, whether the genetic variants are associated with obesity, metabolic syndrome and type 2 diabetes, and with related components thereof. Finally, we tested the genetic association with type 2 diabetes in a meta-analysis including up to 78 021 additional individuals from the DIAGRAM consortium.

Results: Observationally, high coffee intake was associated with low risk of obesity, metabolic syndrome and type 2 diabetes. Further, high coffee intake was associated with high body mass index, waist circumference, weight, height, systolic/diastolic blood pressure, triglycerides and total cholesterol and with low high-density lipoprotein cholesterol, but not with glucose levels. In genetic analyses, 9–10 vs 0–3 coffee-intake alleles were associated with 29% higher coffee intake. However, genetically derived high coffee intake was not associated convincingly with obesity, metabolic syndrome, type 2 diabetes, body mass index, waist circumference, weight, height, systolic/diastolic blood pressure, triglycerides, total cholesterol, high-density lipoprotein cholesterol or glucose levels. Per-allele meta-analysed odds ratios for type 2 diabetes were 1.01 (0.98–1.04) for *AHR* rs4410790, 0.98 (0.95–1.01) for *AHR* rs6968865, 1.01 (0.99–1.03) for *CYP1A1/2* rs2470893, 1.01 (0.98–1.03) for *CYP1A1/2* rs2472297 and 0.98 (0.95–1.01) for *CYP1A1* rs2472299.

Conclusions: High coffee intake was associated observationally with low risk of obesity, metabolic syndrome and type 2 diabetes, and was associated observationally with related components thereof, but with no genetic evidence to support corresponding causal relationships.

Key words: Coffee, metabolic syndrome, type 2 diabetes, BMI, blood pressure, plasma cholesterol, plasma glucose

Key Messages

- Several large meta-analyses have consistently found an inverse association between coffee intake and risk of diabetes.
- It is also possible that coffee intake is associated with the related outcomes of obesity and metabolic syndrome and related components thereof, as suggested by previous studies.
- We found that high coffee intake was associated observationally with low risk of obesity, metabolic syndrome and type 2 diabetes, and was associated observationally with related components thereof.
- However, we found no genetic evidence to support corresponding causal relationships. Indeed, genetically high coffee intake was not associated with obesity, metabolic syndrome, type 2 diabetes or related outcomes.
- Unmeasured confounding factors are the most likely explanations for the observational findings.

Introduction

Coffee, one of the most widely consumed beverages in the world, contains several biologically active compounds with potential effects on health, including caffeine, antioxidants and different minerals.[1,2] Multiple studies have examined the association between coffee intake and risk of type 2 diabetes: the majority of studies including several large meta-analyses have consistently found an inverse association between coffee intake and risk of diabetes.[1–9] Because type 2 diabetes coexists with obesity and metabolic syndrome, it is also possible that coffee intake is associated with these outcomes and related components, as suggested by previous studies.[9–13] However, because all previous studies have been observational, the question of causality remains unclear.

A Mendelian randomization study takes advantage of the fact that genetic variance is inherited randomly at gamete formation and, by studying known genetic determinants of the exposure variable of interest, reverse causation and most confounding may be circumvented.[14] To be used for this purpose, to study potential causal effect of coffee intake, five genetic variants located near the *CYP1A1*, *CYP1A2* and *AHR* genes have recently been associated with high coffee/caffeine intake in genome-wide association studies.[15–17]

In this study we tested the hypothesis that genetically high coffee intake is associated with low risk of obesity, metabolic syndrome and type 2 diabetes, and with related components thereof. Using 93 179 individuals from two large general population cohorts, we tested first whether coffee intake is associated with obesity, metabolic syndrome, type 2 diabetes and related components thereof, in observational analyses; second, whether the five genetic variants near the *CYP1A1*, *CYP1A2* and *AHR* genes involved in caffeine metabolism are associated with coffee intake; and third, whether the genetic variants are associated with obesity, metabolic syndrome, type 2 diabetes and related components thereof. Finally, we tested the genetic association with type 2 diabetes in a meta-analysis including up to 78 021 additional individuals from the DIAGRAM consortium.

Methods

Participants

We studied age-stratified, randomly selected individuals ($n = 93\,179$) from two similar general population cohorts, essentially conducted by the same investigators and using identical methods: the Copenhagen General Population Study (CGPS) and the Copenhagen City Heart Study (CCHS).[18–20] Participants were Whites of Danish descent, according to the Danish Central Person Registry. The two studies were approved by Herlev Hospital and Danish ethical committees, and conducted according to the Declaration of Helsinki. Written informed consent was

obtained from all participants. At each examination, all participants filled out a questionnaire, reviewed by an investigator present, and had a physical examination performed and blood samples taken for biochemical analyses and DNA extraction. There was no overlap of participants between the two studies. We also included up to 78 021 individuals from the DIAGRAM consortium.[21]

In observational analyses, to examine the association between coffee intake and obesity, metabolic syndrome, type 2 diabetes and related components, we included 83 436 individuals from the 2003–12 examinations of the CGPS with available information on coffee intake. Only participants from the CGPS were included in these analyses, since coffee intake was not available for the CCHS or the DIAGRAM consortium. For the same reason, only CGPS individuals were used in the genetic analyses examining the association of the five genetic variants near *CYP1A1, CYP1A2* and *AHR* genes with coffee intake. However, to examine the association between the five coffee-intake alleles and obesity, metabolic syndrome, type 2 diabetes and related outcomes, we pooled the 83 436 participants from the CGPS with 9743 participants from the 1991–94 and 2001–03 examinations of the CCHS and analysed them as a single study. Finally, to examine the association between genetically derived coffee intake and type 2 diabetes with maximum statistical power, we performed meta-analyses of up to 171 200 individuals from the pooled CGPS and CCHS with the DIAGRAM consortium.

Coffee intake

Using the question 'What is your average weekly consumption of coffee (in cups)?', baseline coffee intake was converted into cups/day. In 2010, the Danes had the fourth largest coffee intake in the world, and only 0.3 % of the total import of coffee into Denmark consisted of decaffeinated coffee.[22]

Anthropometric variables, blood pressure, lipids and glucose

Waist circumference, weight, height and systolic and diastolic blood pressure were measured at examination. Body mass index (BMI) was measured as weight (kilograms) divided by measured height squared (metres squared). Plasma levels of triglycerides, high-density lipoprotein cholesterol, total cholesterol and glucose were measured using standard hospital assays. For these 10 outcomes, less than 0.3% of data were missing among CGPS participants. Obesity was defined as body mass index over 30 kg/m^2.

Metabolic syndrome

Metabolic syndrome was defined as three or more of the following five metabolic abnormalities: (i) waist circumference at least 94 cm in men and at least 80 cm in women; (ii) systolic blood pressure at least 130 mmHg and/or diastolic blood pressure at least 85 mmHg and/or antihypertensive treatment; (iii) non-fasting plasma triglycerides at least 1.7 mmol/l; (iv) high-density lipoprotein cholesterol level less than 1.0 mmol/l in men and less than 1.3 mmol/l in women; (v) registry-based diagnosis of diabetes (see below) and/or self-reported diabetes and/or use of anti-diabetic medication and/or non-fasting baseline glucose levels above 11 mmol/l.

Diabetes

Information on diagnosis of type 2 diabetes was obtained from the national Danish Patient Registry covering all hospital contacts in Denmark, and from the national Danish Causes of Death Registry containing death certificate information from general practitioners and doctors at hospitals. Registry-based diagnoses were obtained by following each participant using the unique Personal Identification Number from 1977 to end of follow-up in April 2013, recording diagnoses of type 2 diabetes (World Health Organization International Classification of Diseases: ICD-8 250; ICD-10 E11, E13-E14); no individuals were lost to follow-up, and the few individuals who emigrated were censored at date of emigration

Covariates

Smoking status, physical inactivity and antihypertensive and lipid-lowering medication were self-reported. Physical inactivity was defined as leisure time activity less than 4 h weekly and predominantly sedentary work. Covariates were complete for all CGPS participants.

Genetic analyses

Participants were genotyped for three variants near the *CYP1A1* and *CYP1A2* genes (rs2472297, rs2470893 and rs2472299) and two variants near the *AHR* gene (rs4410790 and rs6968865), using *TaqMan*-based assays (Applied Biosystems, Foster City, CA), including positive controls for each genotype verified by sequencing. Since we performed re-runs, call rates were $\geq$99.3%. All genotype distributions were in Hardy-Weinberg equilibrium ($P \geq 0.1$). Participants were also genotyped for three non-coding polymorphisms in *FTO*, *MC4R* and *TMEM18* (rs9939609, rs17782313 and rs6548238) using *TaqMan*-based arrays. These polymorphisms with the largest known association with body mass index were included in genetic analyses to examine potential

reverse causation from BMI to coffee intake, and as a positive control for risk of type 2 diabetes.[23–26]

Statistical analyses

Data were analysed using Stata/SE 13.1 (StataCorp LP, College Station, TX). In the observational analyses, we used linear regression models unadjusted and multivariable adjusted for sex, age, smoking status, physical inactivity and use of antihypertensive and lipid-lowering medication to estimate associations of coffee intake with body mass index, waist circumference, weight, height, systolic and diastolic blood pressure, plasma triglycerides, high-density lipoprotein cholesterol, plasma total cholesterol and plasma glucose; the covariates for adjustment were chosen because they could influence coffee intake and/or risk of obesity, metabolic syndrome, type 2 diabetes and related components. In cross-sectional analyses, we divided the participants into groups according to coffee intake and used logistic regression models to estimate odds ratios for obesity and metabolic syndrome, multivariable adjusted for the same covariates as in the linear regression analyses. In prospective analyses, we divided the participants into groups according to coffee intake and used Cox proportional hazards regression models with age as underlying time scale, and delayed entry at examination (left truncation) to estimate hazard ratios for type 2 diabetes multivariable adjusted for the same covariates as above; individuals with a diagnosis of diabetes defined as registry-based diagnosis of diabetes type 1 or 2, self-reported diabetes, use of anti-diabetic medication, and/or a non-fasting baseline glucose level above 11 mmol/l, prior to or at study entry, were excluded ($n = 3559$) in the prospective analyses. For test of trend, median coffee intake of each coffee intake subgroup was entered into the regression models as a continuous variable. To test for possible interactions, two-factor interaction terms were inserted into the model.

In the genetic analyses, to examine the effect of the five genetic variants on coffee intake we first calculated mean coffee intake according to genotype. Test of trend was by Cuzicks extension of the Wilcoxon rank sum test. To obtain maximum statistical power in genetic analyses, we examined all possible genotype combinations and grouped each participant according to number of coffee-intake allele score. Allele scores of 0–3 and 9–10 were grouped together to ensure sufficient numbers of participants in groups with extreme allele scores. F- and R^2-values were calculated using the *regress* command in STATA. Similar analytical strategies were applied for allele score of *FTO*, *MC4R* and *TMEM18* genotypes known to be associated with body mass index.[23–26] Genetic variants near the *AHR* gene (rs6968865rs4410790) are in strong linkage disequilibrium (Supplementary Figure 1, available as Supplementary data at *IJE* online); however, because we still observed a higher mean coffee intake between each allele score group, both variants were included in further analyses. Using data from the HapMap database in HaploView 4.2 (Broad Institute, Cambridge, MA), we calculated linkage disequilibrium between the five genetic variants and other nearby genetic variants associated with risk of diabetes (Supplementary Figures 2 and 3, available as Supplementary data at *IJE* online).[27,28]

To test whether genetically high coffee intake is associated with low risk of obesity, metabolic syndrome, type 2 diabetes and related outcomes thereof, we first used linear regression models adjusted for age and sex to estimate associations of coffee-intake allele score with body mass index, waist circumference, weight, height, systolic and diastolic blood pressure, plasma triglycerides, high-density lipoprotein cholesterol, plasma total cholesterol and plasma glucose levels. Next, we used logistic regression models adjusted for age and sex to estimate odds ratios for obesity and metabolic syndrome according to number of coffee-intake alleles in the CGPS and the CCHS pooled together. We used Cox proportional hazard regression models adjusted for age and sex to estimate hazard ratios for type 2 diabetes according to number of coffee-intake alleles in the same participants. The observed odds/hazard ratios for a one-cup of coffee/day higher intake in observational analyses was used to calculate the theoretically predicted risk of obesity, metabolic syndrome and type 2 diabetes, as done previously.[20] Similar analytical strategies were applied for *FTO*, *MC4R* and *TMEM18* genotypes, which we included as a body mass index allele score as a positive control for type 2 diabetes.

Furthermore, to examine the association between type 2 diabetes and genetically high coffee intake with maximum statistical power, we used logistic regression models: we estimated odds ratios per coffee-intake allele adjusted for age and sex for each genotype separately in the CGPS and CCHS pooled together. We then performed fixed and random effect models meta-analyses of the estimates from the combined CGPS and CCHS with the DIAGRAM consortium (at http://diagram-consortium.org/). Finally, the odds ratios for type 2 diabetes per each coffee-intake allele increase from 0 to 10 alleles were calculated for CGPS+CCHS, for DIAGRAM and in meta-analysis for CGPS +CCHS + DIAGRAM.

Using an online sample size and power calculator for Mendelian randomization with a binary outcome (http://spark.rstudio.com/sb452/power/) we had 80% power at two-sided $P < 0.05$ to detect a causal odds ratio for an SD increase in coffee intake using allele score with an R^2 of 0.007 of 0.73 for obesity (15 245 cases), 0.74 for metabolic syndrome (29 107 cases) and 0.57 for type 2 diabetes (3963 cases) in the CGPS + CCHS with 93 179 participants; the corresponding value for type 2 diabetes for

Table 1. Baseline characteristics of the 83 436 participants from the Copenhagen General Population Study according to coffee intake

Coffee intake (cups/day)	0	0.1–1	1.1–2	2.1–3	3.1–4	4.1–5	>5	*P* for trend
Participants	8948	14 327	16 034	17 128	8830	9479	8690	
Age (years)	49 (42–61)	62 (51–71)	60 (49–69)	58 (48–66)	58 (49–67)	56 (48–64)	56 (49–64)	$1*10^{-4}$
Men (%)	34	38	38	46	48	56	62	$<1*10^{-300}$
Current smokers (%)	13	15	15	16	20	24	37	$<1*10^{-300}$
Physical inactivity (%)	48	60	58	53	55	49	50	$7*10^{-30}$
Antihypertensive medication (%)	14	25	23	19	20	17	16	$2*10^{-19}$
Lipid-lowering medication (%)	7	14	13	11	12	10	11	0.71

Values are median (interquartile range) for continuous variables and percent for categorical variables. P values for trend were by Cuzick's extension of the Wilcoxon rank-sum test.

Table 2. Baseline characteristics of the 83 436 participants from the Copenhagen General Population Study according to number of coffee-intake-increasing alleles

Allele score	0–3	4	5	6	7	8	9–10	*P* for trend
Participants	16 870	12 008	15 363	15 746	8514	10 864	4071	
Age (years)	58 (48–67)	58 (48–67)	58 (48–67)	58 (48–67)	58 (48–67)	58 (48–67)	58 (48–66)	0.24
Men (%)	45	46	45	45	44	44	45	0.39
Current smokers (%)	19	19	19	19	18	19	19	0.35
Physical inactivity (%)	54	54	54	54	55	55	54	0.40
Antihypertensive medication (%)	20	20	21	19	20	20	20	0.07
Lipid-lowering medication (%)	11	12	12	11	12	11	12	0.37

Values are median (interquartile range) for continuous variables and percent for categorical variables. P values for trend were by Cuzick's extension of the Wilcoxon rank-sum test.

CGPS + CCHS + DIAGRAM was 0.79 with 171 200 participants and 26 632 cases. Also, for each allele score group vs 0–3 alleles, we calculated for each binary outcome the lowest odds ratio we could exclude at 80% power and two-sided $P < 0.05$ using NCSS-PASS.

Results

Baseline characteristics of the 83 436 White individuals of Danish descent from the CGPS differed according to coffee intake (Table 1), but not according to coffee-intake allele score overall (Table 2). When stratifying participants into those with and without coffee intake, among individuals without coffee intake age and use of antihypertensive medication decreased with increasing coffee-intake allele score (Supplementary Tables 1 and 2), likely explained by collider stratification bias.

Coffee intake, anthropometric variables, blood pressure, lipids and glucose: observational analyses

In observational linear regression analyses, 83 436 individuals from the CGPS were included. High coffee intake was associated with high body mass index, waist circumference, weight, height, systolic/diastolic blood pressure, triglycerides, and total cholesterol, with low high-density lipoprotein cholesterol, but not with glucose levels (Figure 1); thus for a one cup/day higher coffee intake, observed changes were 0.075 kg/m^2 (95% confidence interval, 0.062–0.088), 0.496 cm (0.456–0.536), 0.705 kg (0.660–0.751), 0.526 cm (0.498–0.555), 0.160 mmHg (0.095–0.224), 0.217 mmHg (0.182–0.252), 0.013 mmol/l (0.010–0.017), 0.031 mmol/l (0.028–0.035), −0.011mmol/l ((−0.012)–(−0.009)) and −0.001 mmol/l ((−0.005)–0.003), respectively.

Coffee intake, obesity, metabolic syndrome and type 2 diabetes: observational analyses

In observational cross-sectional and prospective analyses among 83 436 individuals, 13 806 and 26 046 prevalent cases of obesity and metabolic syndrome, respectively, and 789 incident cases of type 2 diabetes from the CGPS, were identified. In cross-sectional analyses, there was a lower risk of obesity (only up to four cups/day) and of metabolic syndrome according to higher coffee intake (trend for both, $P \leq 0.001$) (Figure 2). Compared with individuals with no coffee intake, odds ratios for obesity were 0.82 (95% confidence interval, 0.77–0.89) with 0.1–1 cup/day, 0.86 (0.81–0.93) with 1.1–2 cups/day, 0.86 (0.80–0.92)

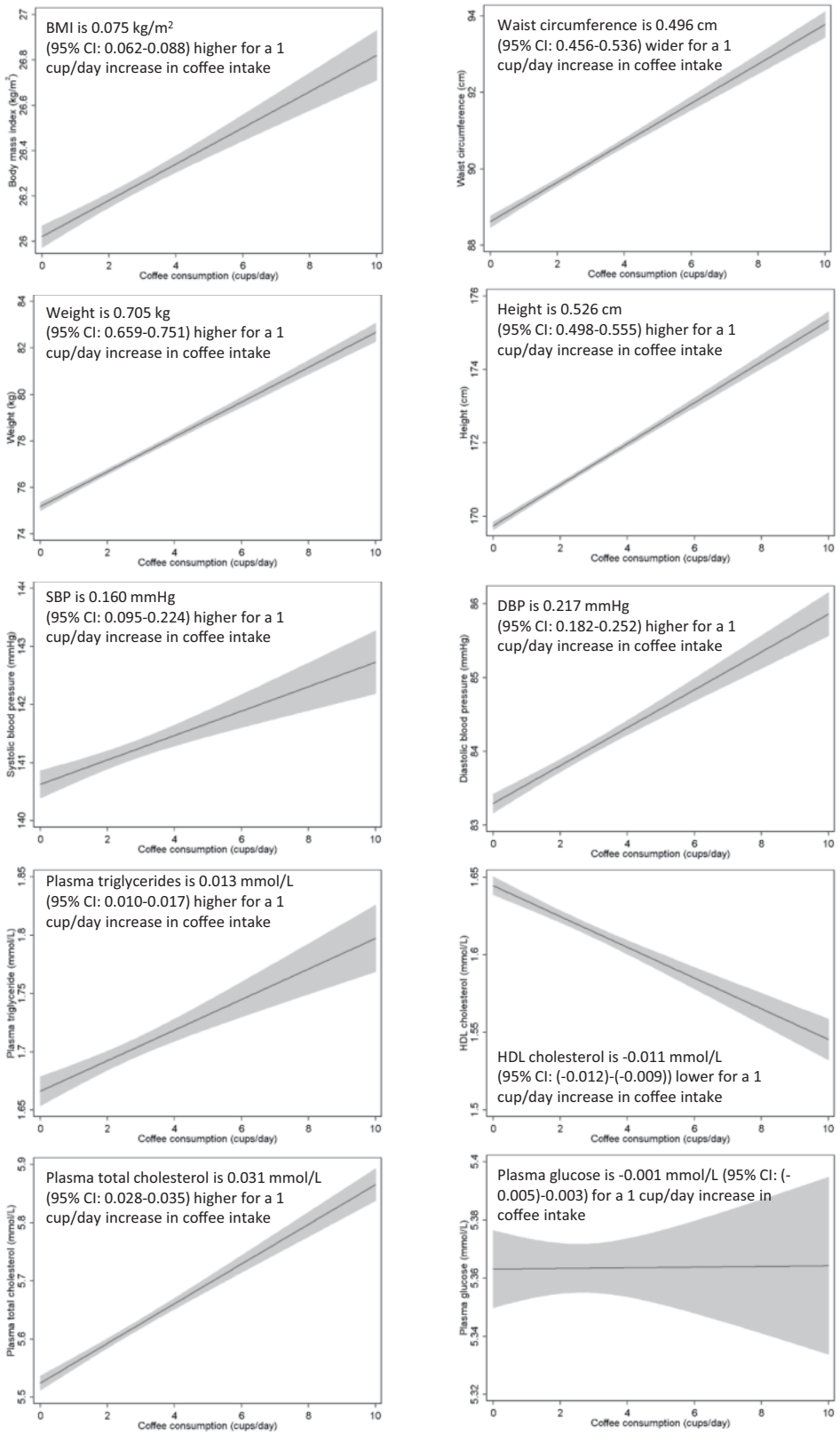

Figure 1. Linear regressions for body mass index, waist circumference, weight, height, systolic and diastolic blood pressure, plasma triglycerides, high-density lipoprotein cholesterol, plasma total cholesterol and plasma glucose according to coffee intake. Analyses were unadjusted and included 83 436 individuals from the Copenhagen General Population Study, with slight variation in number of individuals according to availability of data. CI, confidence interval; BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density lipoprotein.

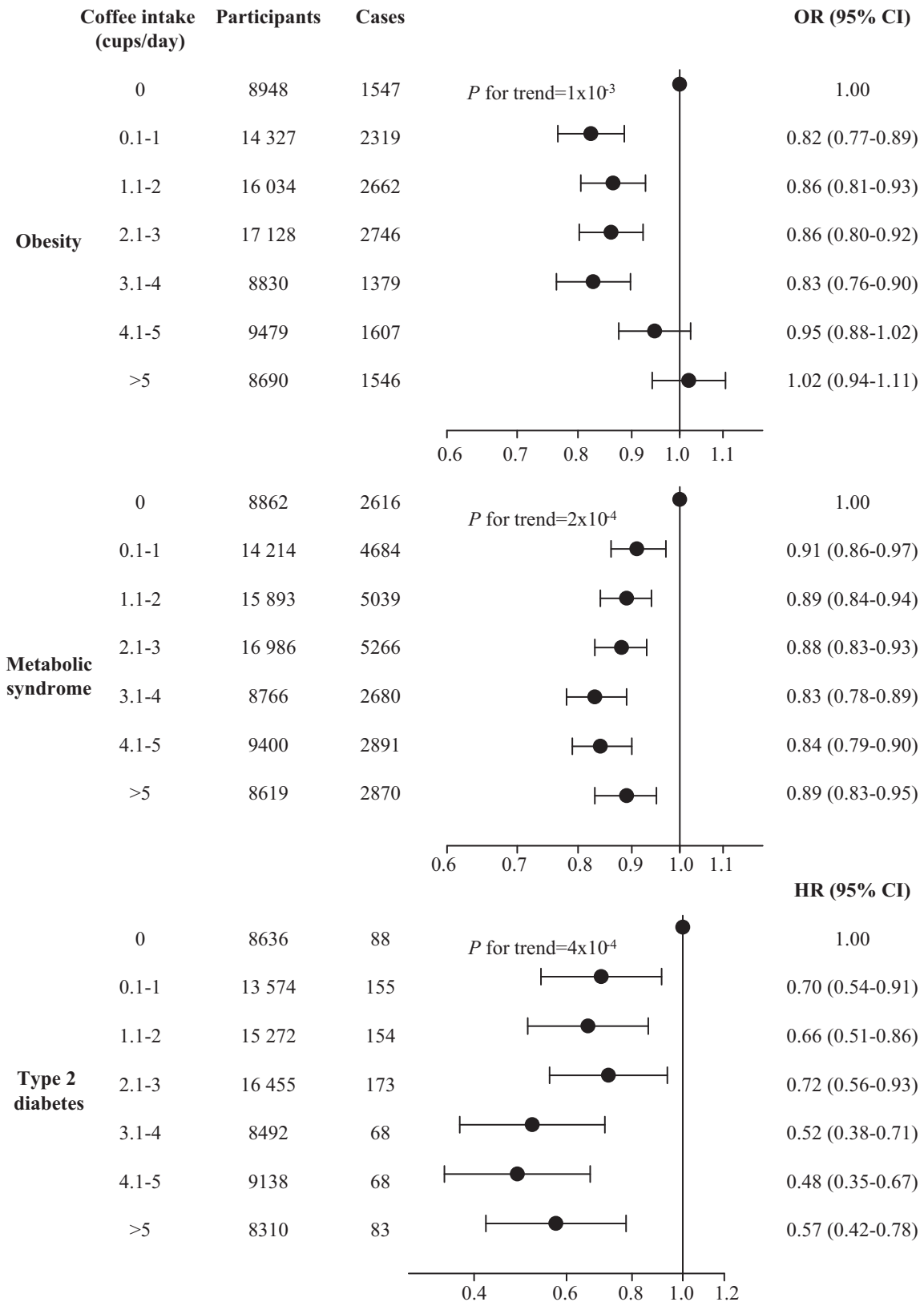

Figure 2. Odds ratios for obesity and metabolic syndrome and hazard ratios for incident type 2 diabetes according to coffee intake. Analyses included 83 436 individuals from the Copenhagen General Population Study; 3559 individuals with type 1 or 2 diabetes at baseline were excluded in prospective analyses for type 2 diabetes. Odds and hazard ratios were multivariable adjusted for age, sex, smoking status, physical inactivity and use of antihypertensive and lipid-lowering medication. OR, odds ratio; HR, hazard ratio; CI, confidence interval.

with 2.1–3 cups/day, 0.83 (0.76–0.90) with 3.1–4 cups/day, 0.95 (0.88–1.02) with 4.1–5 cups/day and 1.02 (0.94–1.11) with >5 cups/day. Corresponding odds ratios for metabolic syndrome were 0.91 (0.86–0.97), 0.89 (0.84–0.94), 0.88 (0.83–0.93), 0.83 (0.78–0.89), 0.84 (0.79–0.90) and 0.89 (0.83–0.95), respectively. In prospective analyses, there was a lower risk of type 2 diabetes according to higher coffee intake (trend, $P = 0.0004$). Compared with individuals with no coffee intake, hazard ratios were 0.70 (0.54–0.91) with 0.1–1 cup/day, 0.66 (0.51–0.86) with 1.1–2 cups/day, 0.72 (0.56–0.93) with 2.1–3 cups/day, 0.52 (0.38–0.71) with 3.1–4 cups/day, 0.48 (0.35–0.67) with 4.1–5 cups/day and 0.57 (0.42–0.78) with >5 cups/day.

The odds ratios for obesity and metabolic syndrome for an increase of one cup of coffee/day in observational analyses were 1.02 (95% confidence interval, 1.01–1.02) and 0.99 (0.98–0.99). Corresponding hazard ratio for type 2 diabetes was 0.94 (0.91–0.97). When dividing participants into different subgroups to test for possible interactions, results were similar to those presented in the main analysis (Supplementary Figure 4, available as Supplementary data at *IJE* online): there was no convincing interaction between these covariates and coffee intake on risk of obesity, metabolic syndrome or type 2 diabetes.

Genotypes, coffee intake and body mass index

For each genotype there was a stepwise higher mean coffee intake from noncarriers through heterozygotes to homozygotes (trend, $P \leq 5 \times 10^{-16}$) (Figure 3). Also, when combining genotypes by number of coffee-intake alleles, 9–10 vs 0–3 coffee-intake alleles were associated with 29% higher coffee intake, but with no change in body mass index. For body mass index-associated genotypes, 4–6 vs 0–1 alleles were associated with 5% higher coffee intake and 3% higher body mass index.

Coffee intake, anthropometric variables, blood pressure, lipids and glucose: genetic analyses

In genetic linear regression analyses, 93 179 individuals from the CGPS and the CCHS were included. Genetically derived high coffee intake was not associated convincingly with body mass index, waist circumference, weight, height, systolic/diastolic blood pressure, triglycerides, total cholesterol, high-density lipoprotein cholesterol or with glucose levels (Figure 4): for one-allele higher allele score, estimated changes were 0.011 kg/m^2 (95% confidence interval (−0.002)−0.024), 0.038 cm (0.004–0.073), 0.048 kg (0.006–0.091), 0.011 cm ((−0.009)−0.031), −0.044 mmHg ((−0.108)−0.019), −0.046 mmHg ((−0.081)–(−0.011)), 0.000 mmol/l ((−0.004)−0.003), 0.006 mmol/l (0.003–0.010), 0.000 mmol/l ((−0.002)−0.001) and 0.003 mmol/l ((−0.001)−0.007), respectively.

When stratifying participants into those with and without coffee intake, associations between coffee-intake alleles and the various endpoints were similar in the two strata (Table 3); however, among those without coffee intake, blood pressure was lower with higher coffee-intake allele score (Table 3), in accordance with the above-mentioned lower use of antihypertensive medication (Supplementary Table 1).

Coffee intake, obesity, metabolic syndrome and type 2 diabetes: genetic analyses

In genetic cross-sectional and prospective analyses among 93 179 individuals, 15 245 and 29 107 prevalent cases of obesity and metabolic syndrome and 3963 incident cases of type 2 diabetes from the CGPS and the CCHS, were included in analyses according to number of coffee-intake alleles. In meta-analyses of the five individual coffee-intake alleles, up to 171 200 individuals and 26 632 cases with type 2 diabetes from the CGPS, CCHS and the DIAGRAM consortium combined were included.

Assuming that a high coffee intake is causally associated with low risk of obesity, metabolic syndrome and type 2 diabetes, elevated coffee intake due to genetic variants should confer a similar low risk of obesity, metabolic syndrome and type 2 diabetes compared with observational estimates. On the basis of this assumption, we estimated that the increment in coffee intake associated with number of coffee-intake alleles would theoretically predict odds ratios for obesity of 1.00 (1.00–1.00) for 4 alleles, of 1.00 (1.00–1.01) for 5 alleles, of 1.00 (1.00–1.01) for 6 alleles, of 1.01 (1.00–1.01) for 7 alleles, of 1.01 (1.00–1.01) for 8 alleles, and of 1.01 (1.00–1.02) for 9–10 alleles compared with individuals with 0–3 alleles (Figure 5). Corresponding theoretically predicted odds ratios for metabolic syndrome were 1.00 (1.00–1.00), 1.00 (0.99–1.00), 1.00 (0.99–1.00), 0.99 (0.99–1.00), 0.99 (0.99–1.00) and 0.99 (0.99–0.99), and hazard ratios for type 2 diabetes were 0.99 (0.98–1.00), 0.98 (0.97–0.99), 0.98 (0.97–0.99), 0.97 (0.96–0.99), 0.97 (0.95–0.99) and 0.96 (0.94–0.98). However, corresponding observed odds and hazard ratios for obesity, metabolic syndrome and type 2 diabetes did not differ from 1.0.

When stratifying participants into those with and without coffee intake, associations between coffee-intake alleles and the various endpoints were similar in the two strata (Table 3).

To test our method, we examined the association between *FTO*, *MC4R* and *TMEM18* genotypes and risk of type 2 diabetes. Number of body mass index alleles according to *FTO*,

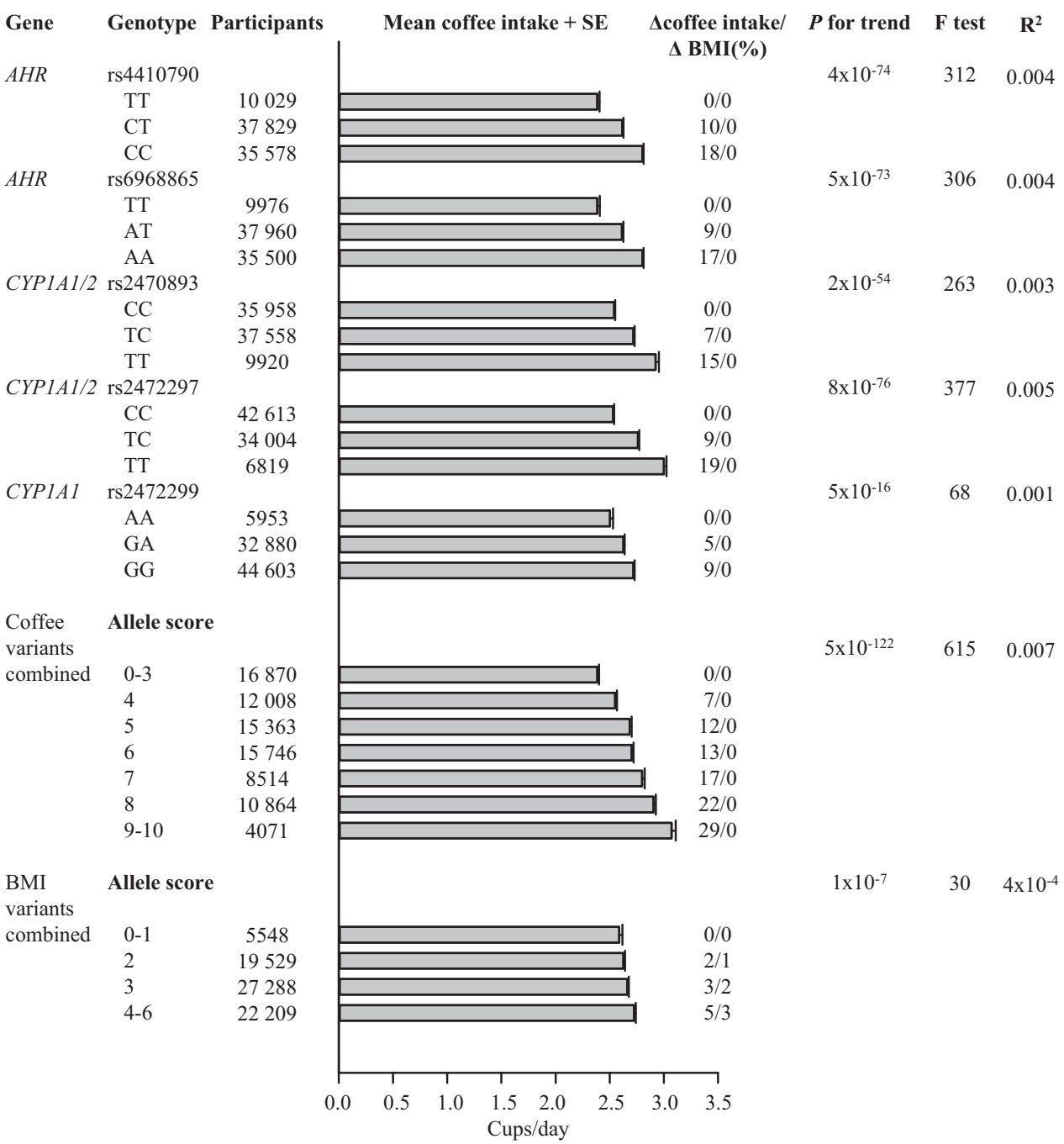

Figure 3. Coffee intake and body mass index according to coffee-intake genotypes and to number of coffee-intake and body-mass-index alleles. Analyses included 83 436 individuals from the Copenhagen General Population Study. Coffee-intake allele score was combinations of *CYP1A1*, *CYP1A2* and *AHR* genotypes, and body mass index allele score was combination of *FTO*, *MC4R* and *TMEM18* genotypes. Groupings of 0–3 and 9–10 coffee-intake alleles and 0–1 and 4–6 body mass-index alleles were made in the allele score to secure sufficient number of individuals in the extreme groups. Values represent means plus standard errors. *P*-values are for trend. F-test is for statistical strength of the genotypes or allele score as instruments, and R^2 is a measure of explained variation. SE, standard error; BMI, body mass index.

MC4R and *TMEM18* genotypes were associated with up to 3.1% lower mean body mass index (0–1 versus 4–6 BMI allele score), resulting in a theoretically predicted risk of type 2 diabetes of 0.90 (0.89–0.90) (Figure 5). In line with this, the observed hazard ratios for type 2 diabetes according to number of body mass index alleles were similarly lower, indicating a causal relationship. Importantly, the strength of the instruments evaluated as F-tests were better for coffee-intake alleles than for body mass index alleles, and the coffee-intake allele score explained more variation in coffee intake (R^2) than the body mass index allele did for variation in body mass index (Figure 5).

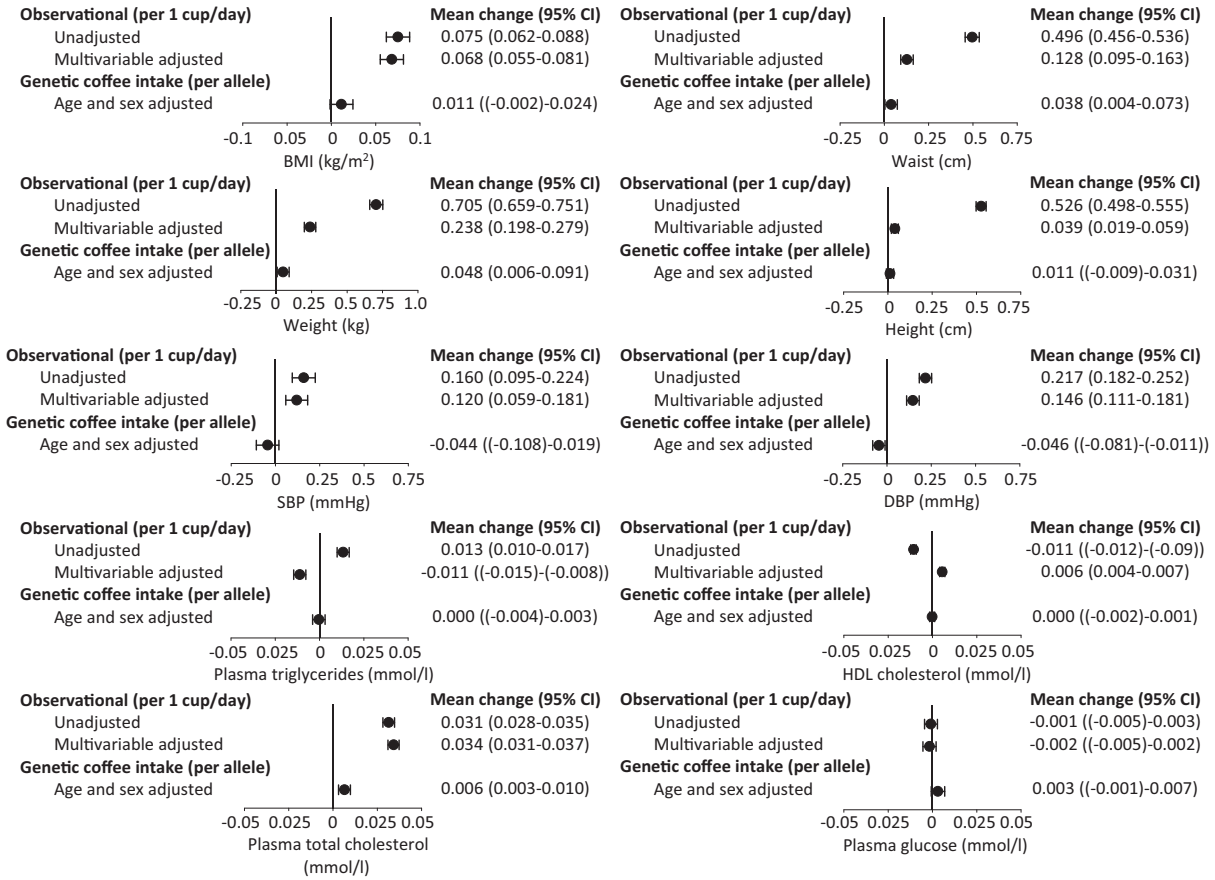

Figure 4. Linear regressions coefficients for body mass index, waist circumference, weight, height, systolic and diastolic blood pressure, plasma triglycerides and high-density lipoprotein cholesterol, plasma total cholesterol and plasma glucose according to coffee intake and allele score. Observational analyses were unadjusted and multivariable adjusted for age, sex, smoking status, physical inactivity and use of antihypertensive and lipid-lowering medication, and genetic analyses were adjusted for age and sex. Analyses included 93 179 individuals from the Copenhagen General Population Study and the Copenhagen City Heart Study combined, with slight variation in number of individuals according to availability of data. CI, confidence interval; BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density lipoprotein.

Odds ratios for type 2 diabetes for each coffee-intake allele separately did not differ from 1.0 in the different studies, and results from meta-analyses were similar (Figure 6): per allele meta-analysed odds ratios for type 2 diabetes were 1.01 (0.98 to 1.04) for *AHR* rs4410790, 0.98 (0.95 to 1.01) for *AHR* rs6968865, 1.01 (0.99 to 1.03) for *CYP1A1/2* rs2470893, 1.01 (0.98 to 1.03) for *CYP1A1/2* rs2472297 and 0.98 (0.95 to 1.01) for *CYP1A1* rs2472299. Random effect model estimates were similar. Also, the odds ratios for type 2 diabetes, per each coffee-intake allele increase from 0 to 10 alleles, were 1.00 (0.99 to 1.02) for CGPS + CCHS, 0.97 (0.88 to 1.07) for DIAGRAM and, in meta-analysis, 1.00 (0.99 to 1.01) for CGPS + CCHS + DIAGRAM.

Sensitivity analyses

When individuals with coffee intake of 1.1–2 cups/day were used as reference group (data not shown) or when prospective analyses were expanded to include registry-based diagnosis of types 1 and 2 diabetes combined—as some type 2 diabetes may be registered as type 1 (data not shown)—results were similar to those presented. Also, when only individuals from the CGPS with information on coffee intake were included in the genetic analysis on risk of obesity, metabolic syndrome and type 2 diabetes (Supplementary Figure 5, available as Supplementary data at *IJE* online), the results were similar to the results including also the CCHS (Figure 5) or both the CCHS and the DIAGRAM (Figure 6). Finally, when the genetic analyses on risk of type 2 diabetes were conducted using 789 incident cases of type 2 diabetes instead of 3963 incident and prevalent diabetes cases combined, the results were similar to those in Figure 5, but with wider confidence intervals (data not shown).

Discussion

High coffee intake was associated with low risk of obesity, metabolic syndrome and type 2 diabetes, and was

Table 3. Linear regression coefficients for continuous outcomes, hazard ratios for type 2 diabetes, and odds ratios for obesity and metabolic syndrome per coffee intake allele from 0 to 10 alleles for all participants, participants with no coffee intake, and participants with coffee intake seperately in the Copenhagen General Population Study

	Number of individuals	Cases	Coefficient type (unit)	Coefficient value	95 % CI	*P* for interaction
Body mass index						
All	83 436	–	β (kg/m^2)	0.012	(−0.002)–0.025	
No coffee intake	8948	–	β (kg/m^2)	0.010	(−0.037)–0.056	0.93
With coffee intake	74 488	–	β (kg/m^2)	0.013	(−0.002)–0.027	
Waist circumference						
All	83 082	–	β (cm)	0.038	0.002–0.075	
No coffee intake	8894	–	β (cm)	0.065	(−0.056)–0.186	0.64
With coffee intake	74 188	–	β (cm)	0.038	(−0.001)–0.076	
Weight						
All	83 315	–	β (kg)	0.048	0.006–0.091	
No coffee intake	8934	–	β (kg)	0.045	(−0.097)–0.187	0.92
With coffee intake	74 381	–	β (kg)	0.049	0.005–0.093	
Height						
All	83 320	–	β (cm)	0.012	(−0.009)–0.032	
No coffee intake	8936	–	β (cm)	0.016	(−0.047)–0.079	0.99
With coffee intake	74 384	–	β (cm)	0.010	(−0.011)–0.031	
Sytolic blood pressure						
All	83 436	–	β (mmHg)	−0.047	(−0.110)–0.016	
No coffee intake	8948	–	β (mmHg)	−0.192	(−0.378)-(−0.006)	0.18
With coffee intake	74 488	–	β (mmHg)	−0.035	(−0.103)–0.032	
Diastolic blood pressure						
All	83 436	–	β (mmHg)	−0.044	(−0.080)-(−0.007)	
No coffee intake	8948	–	β (mmHg)	−0.118	(−0.230)-(−0.007)	0.13
With coffee intake	74 488	–	β (mmHg)	−0.038	(−0.077)–0.000	
Plasma triglycerides						
All	83 436	–	β (mmol/L)	0.001	(−0.003)–0.004	
No coffee intake	8948	–	β (mmol/L)	0.005	(−0.007)–0.017	0.44
With coffee intake	74 488	–	β (mmol/L)	0.000	(−0.004)–0.004	
High-density liporpotein cholesterol						
All	83 093	–	β (mmol/L)	−0.001	(−0.002)–0.001	
No coffee intake	8916	–	β (mmol/L)	−0.002	(−0.007)–0.002	0.58
With coffee intake	74 177	–	β (mmol/L)	−0.001	(−0.003)–0.001	
Plasma total cholesterol						
All	83 436	–	β (mmol/L)	0.006	0.002–0.009	
No coffee intake	8948	–	β (mmol/L)	0.000	(−0.010)–0.010	0.19
With coffee intake	74 488	–	β (mmol/L)	0.006	0.002–0.009	
Plasma glucose						
All	83 436	–	β (mmol/L)	0.003	(−0.001)–0.007	
No coffee intake	8911	–	β (mmol/L)	0.001	(−0.010)–0.013	0.71
With coffee intake	74 125	–	β (mmol/L)	0.003	(−0.001)–0.007	
Type 2 diabetes						
All	83 436	3051	HR	1.01	0.99–1.02	
No coffee intake	8948	295	HR	1.03	0.98–1.09	0.40
With coffee intake	74 488	2756	HR	1.01	0.99–1.02	
Obesity						
All	83 436	13 806	OR	1.00	0.99–1.01	
No coffee intake	8948	1547	OR	1.02	0.99–1.04	0.30
With coffee intake	74 488	12 259	OR	1.00	0.99–1.01	
Metabolic syndrome						
All	82 740	26 046	OR	1.00	0.99–1.01	
No coffee intake	8862	2616	OR	1.01	0.98–1.03	0.70
With coffee intake	73 878	23 430	OR	1.00	0.99–1.01	

Only participants from the Copenhagen General Population Study were included in the analyses since information on coffee intake was not available for the Copenhagen City Heart Study. Analyses were adjusted for age and sex. Number of participants in individual analyses varied slighty due to availability of data. HR = hazard ratio. OR = odds ratio. CI = confidence interval.

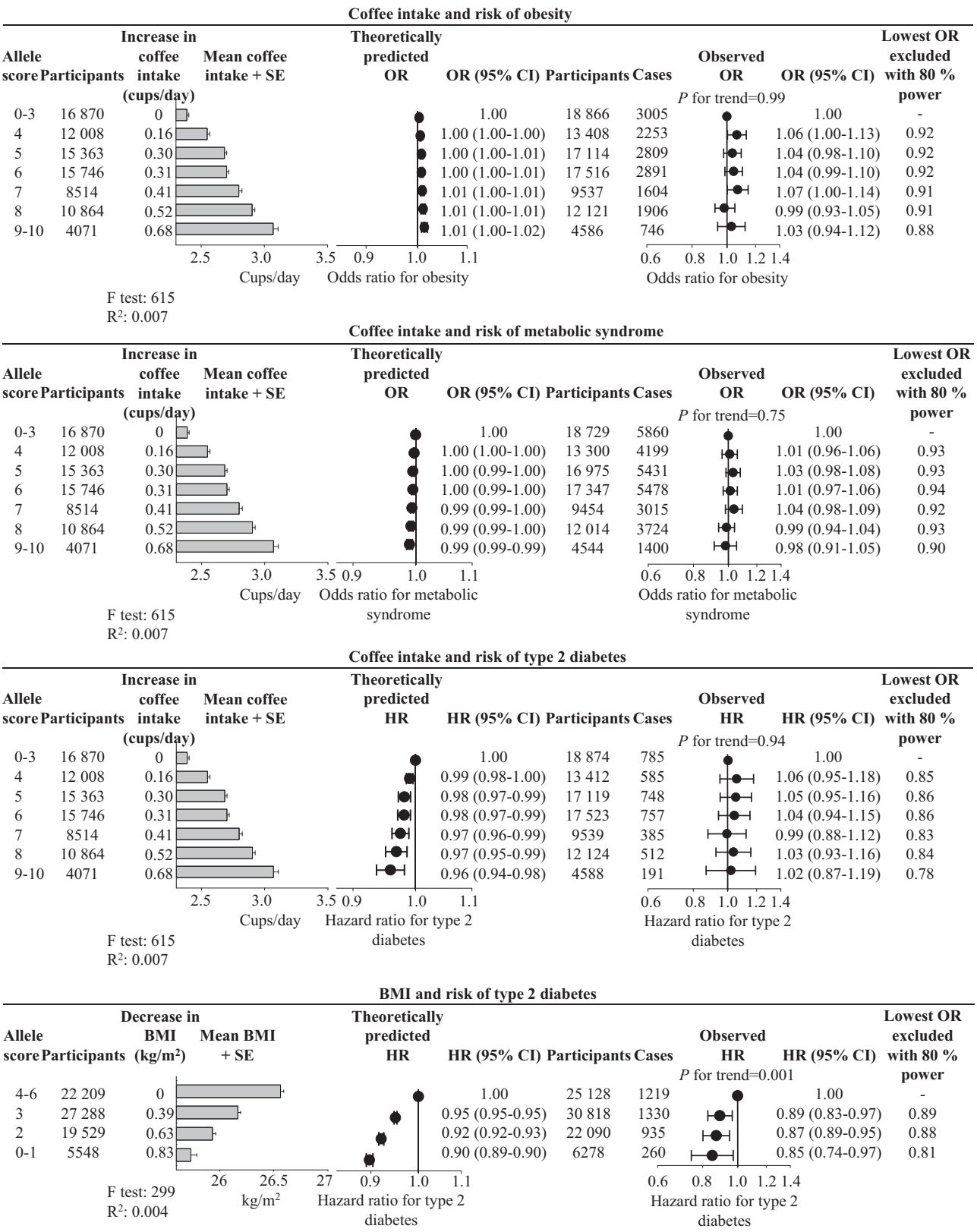

Figure 5. Theoretically predicted and observed odds ratios for obesity and metabolic syndrome, and hazard ratios for incident type 2 diabetes by number of coffee-intake alleles and body mass index alleles. Analyses of theoretically predicted risks were based on 83 436 individuals from the Copenhagen General Population Study with available information on coffee intake, and the observed risk was based on all 93 179 individuals from the Copenhagen General Population Study and the Copenhagen City Heart Study combined. F-test is for statistical strength of the allele score as an instrument, and R^2 is a measure of explained variation. Odds/hazard ratios were adjusted for sex and age. OR, odds ratio; HR, hazard ratio; CI, confidence interval; BMI, body mass index.; SE, standard error.

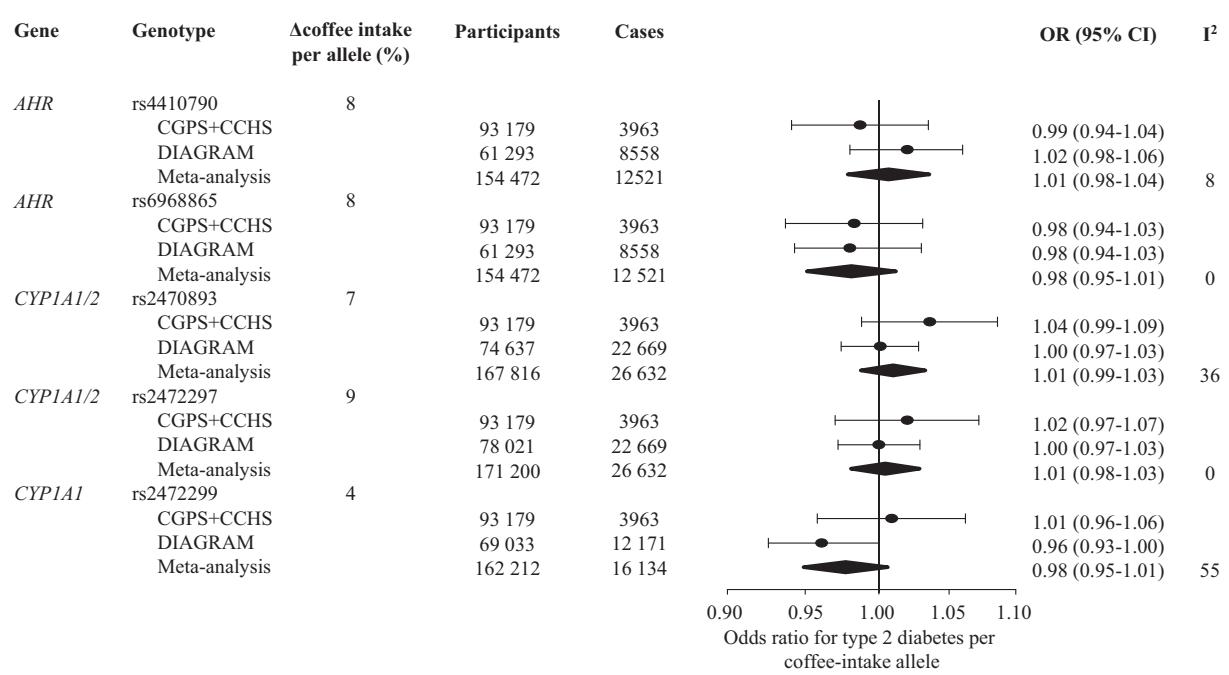

Figure 6. Odds ratios for type 2 diabetes per coffee-intake allele for each genetic variant. Analyses included up to 171 200 individuals from the Copenhagen General Population Study (CGPS), the Copenhagen City Heart Study (CCHS) and the DIAGRAM consortium, separately and combined as fixed effect meta-analysed odds ratios. Numbers of individuals and cases vary according to availability of data from the DIAGRAM consortium. Values of I^2 represent percentage heterogeneity between studies in the meta-analyses. Odds ratios were adjusted for sex and age in CGPS + CCHS but unadjusted in the DIAGRAM consortium. OR, odds ratio; CI, confidence interval.

associated with related components thereof, but with no genetic evidence to support corresponding causal relationships. These are novel genetic findings.

The hypothesis that coffee intake may have a protective effect on risk of type 2 diabetes, and possibly on obesity, metabolic syndrome and related outcomes, is based on several observational studies.[1–6,29,30] In line with these studies, we found a corresponding 6% lower risk of type 2 diabetes for a one cup/day higher coffee intake in observational analyses and also lower risks of obesity (up to 4 cups/day) and metabolic syndrome. Such findings have been supported by several mechanistic studies reporting an improvement in indicators of glucose metabolism such as higher insulin secretion, insulin sensitivity and β-cell function with higher coffee intake.[6,31] Some studies propose that the caffeine content in coffee may increase energy expenditure and decrease body weight.[32] Taken together, a large body of observational studies support a protective role of coffee consumption on risk of type 2 diabetes and possibly obesity, metabolic syndrome and related outcomes. Our genetic analyses suggest that these findings could be explained by confounding and/or reverse causation, rather than by direct biological mechanisms. As our genetic associations between *FTO*, *MC4R* and *TMEM18* genotypes and coffee intake suggested that high coffee intake is causally associated with high BMI (rather than with low BMI), and thus with obesity, metabolic syndrome and type 2 diabetes, it is unlikely that reverse causation could explain the observed associations between coffee intake and these outcomes. This makes confounding by unmeasured variables the most likely explanation for the observational findings.

Although the present analyses do not provide genetic evidence suggesting a causal relationship between coffee intake and obesity, metabolic syndrome, type 2 diabetes or related outcomes, it is not possible based on our data to entirely negate the hypothesis. The reason for this is that the associations of the genetic variants with coffee intake are relatively small, and therefore large numbers of cases and controls are required in order to reliably preclude any causal relation. Thus, although this study provides the first line of evidence arguing against a causal relation between coffee intake and risk of obesity, metabolic syndrome, type 2 diabetes and related outcomes, randomized controlled trials or even larger genetic studies of this type are required to definitively refute the hypothesis.

Strengths of our study include the large sample size and the fact that all participants in the Copenhagen studies were Whites of Danish descent, which ensures a homogeneous population and limits the risk of genetic admixture; however, population homogeneity may limit the generalizability of findings, although we are not aware of data to suggest that the present results should not apply to people of most races and in most countries with large

consumption of coffee. It is also reassuring that the results for the DIAGRAM consortium with participants from many different European populations were similar to those for the Copenhagen studies.

Limitations include pleiotropic effects of the genetic variants, since *CYP1A1*, *CYP1A2* and *AHR* genes are involved in the metabolism and regulation of several other components besides caffeine, e.g. clinical drugs, procarcinogenes and steroids.[14,33–35] Although we cannot exclude pleiotropic effects as potential problems, our use of five different genetic variants for coffee intake close to three different genes, but with similar results, nevertheless argue against pleiotropic effects as major problems in the present study.[14] To add further evidence with respect to the pleiotropy issue, we also examined baseline characteristics and endpoints in those with and without coffee intake, which introduces the possibility of collider stratification bias; indeed among individuals without coffee intake: age, use of antihypertensive medication and blood pressure all decreased with increasing coffee-intake allele score, likely explained by collider stratification bias. The collider stratification bias likely is generated because age influences coffee drinking and the allele score influences coffee drinking. Whereas in a source population these are not associated, they may become associated when stratifying on the collider, coffee drinking. The lower age among those with the highest coffee-intake allele score in individuals not drinking coffee, will then also explain the associated observed lower use of antihypertensive medication and blood pressure, as both of these are lower in younger vs older individuals. Indeed, when not adjusting for age, the association of higher coffee-intake allele score with systolic blood pressure was more pronounced ($\beta = -0.192$ with age adjustment compared with $\beta = -0.302$ without age adjustment).

The association between *FTO*, *MC4R* and *TMEM18* genotypes and coffee intake is a potential problem for the positive genetic control, as greater weight actually seems to influence coffee drinking slightly. However, since a 29 % genetically higher coffee intake is not associated with change in risk of type 2 diabetes, it is unlikely that the observed 5% higher coffee intake for *FTO*, *MC4R* and *TMEM18* genotypes would explain the association of high BMI with type 2 diabetes. Another potential limitation is that the genetic variants might not be valid as proxy variables for circulating caffeine levels. A theoretically plausible biological mechanism explaining the association between genetics and coffee intake is that people who slowly metabolize caffeine, because of genetic differences affecting their caffeine metabolising enzymes, will have higher circulating caffeine levels for the same amount of coffee consumed, and partly compensate for this by reducing their coffee intake, as suggested elsewhere.[36] If this was the case, caffeine levels would not be associated with genotype.

Linkage disequilibrium is another concern in Mendelian randomization analyses; however, none of the five genetic variants used were in linkage disequilibrium with nearby genetic variants associated with risk of type 2 diabetes. Completeness of diagnoses of type 2 diabetes in the Danish registries is also a potential problem; however, in the present study 4.3% of all Copenhagen participants were registered with type 2 diabetes, similar to that found overall for Denmark.[37] Finally, in our study there was a single question asking about average weekly coffee intake in cups, with no information on size of the cup, type or strength of the coffee, or period of coffee intake. In observational analyses, inconsistency in measurement of coffee intake would most likely bias hazard/odds ratios towards 1.0, and therefore is unlikely to explain the observed associations with obesity, metabolic syndrome, type 2 diabetes and related outcomes. Furthermore, in genetic analysis, inconsistency likely would affect mean coffee estimates of all genotypes equally, and thus could not explain the observed differences in mean coffee intake.

In conclusion, high coffee intake was associated with low risk of obesity, metabolic syndrome and type 2 diabetes, and was associated with related components, but with no genetic evidence to support corresponding causal relationships.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

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

Acknowledgements

A.T.N., M.T. and B.G.N. designed the study and gathered, analysed and interpreted the data. A.T.N. generated laboratory data, drafted the paper and prepared the figures. All authors had full access to the data. B.G.N. had the final responsibility for the decision to submit for publication. All authors revised and finally approved the manuscript before submission.

Conflict of interest: The authors have no conflicts of interest to declare.

References

1. Butt MS, Sultan MT. Coffee and its consumption: benefits and risks. *Crit Rev.Food Sci Nutr* 2011;**51**:363–73.

2. Nawrot P, Jordan S, Eastwood, J *et al*. Effects of caffeine on human health. *Food Addit Contam* 2003;**20:**1–30.
3. 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.
4. Huxley R, Lee CM, Barzi F *et al*. Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: a systematic review with meta-analysis. *Arch Intern Med* 2009;**169:**2053–63.
5. van Dam RM, Hu FB. Coffee consumption and risk of type 2 diabetes: a systematic review. *JAMA* 2005;**294:**97–104.
6. Natella F, Scaccini C. Role of coffee in modulation of diabetes risk. *Nutr Rev* 2012;**70:**207–17.
7. 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.
8. Jiang X, Zhang D, Jiang W. Coffee and caffeine intake and incidence of type 2 diabetes mellitus: a meta-analysis of prospective studies. *Eur J Nutr* 2014;**53**:25–38.
9. O'Keefe JH, Bhatti SK, Patil HR *et al*. Effects of habitual coffee consumption on cardiometabolic disease, cardiovascular health, and all-cause mortality. *J Am Coll Cardiol* 2013;**62:**1043–51.
10. Hino A, Adachi H, Enomoto M *et al*. Habitual coffee but not green tea consumption is inversely associated with metabolic syndrome: an epidemiological study in a general Japanese population. *Diabetes Res Clin Pract* 2007;**76:**383–89.
11. Grosso G, Stepaniak U, Micek A *et al*. Association of daily coffee and tea consumption and metabolic syndrome: results from the Polish arm of the HAPIEE study. *Eur J Nutr* 2014, Nov 4. PMID: 25367317. [Epub ahead of print.]
12. Matsuura H, Mure K, Nishio N *et al*. Relationship between coffee consumption and prevalence of metabolic syndrome among Japanese civil servants. *J Epidemiol* 2012;**22:**160–66.
13. Rebello SA, van Dam RM. Coffee consumption and cardiovascular health: getting to the heart of the matter. *Curr Cardiol Rep* 2013;**15:**403.
14. Davey Smith, G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;**32:**1–22.
15. 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.
16. 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.
17. 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.
18. 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 ischemic heart disease. *JAMA* 2008;**299:**2524–32.
19. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *JAMA* 2009;**301:**2331–39.
20. Zacho J, Tybjaerg-Hansen A, Jensen JS *et al*. Genetically elevated C-reactive protein and ischemic vascular disease. *N Engl J Med* 2008;**359:**1897–908.
21. Morris AP, Voight BF, Teslovich TM *et al*. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;**44:**981–90.
22. Hermansen K, Bech BH, Dragsted LO *et al*. 2012. http://www.vidensraad.dk/sites/default/files/kaffe_sundhed_og_sygdom_0.pdf (10 May 2015, date last accessed).
23. Nordestgaard BG, Palmer TM, Benn M *et al*. The effect of elevated body mass index on ischemic heart disease risk: causal estimates from a Mendelian randomization approach. *PLoS Med* 2012;**9:**e1001212.
24. Knowler WC, Fowler SE, Hamman RF *et al*. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. *Lancet* 2009;**374:**1677–86.
25. Knowler WC, Barrett-Connor E, Fowler SE *et al*. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;**346:**393–403.
26. Tuomilehto J, Lindstrom J, Eriksson JG *et al*. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 2001;**344:**1343–50.
27. Imamura M, Maeda S. Genetics of type 2 diabetes: the GWAS era and future perspectives. *[Review.]* EndocrJ 2011;**58:** 723–39.
28. Billings LK, Florez JC. The genetics of type 2 diabetes: what have we learned from GWAS? *Ann N Y Acad Sci U S A* 2010;**1212:** 59–77.
29. Ding M, Bhupathiraju SN, Chen, van Dam RM, Hu FB. Caffeinated and decaffeinated coffee consumption and risk of type 2 diabetes: a systematic review and a dose-response meta-analysis. *Diabetes Care* 2014;**37:**569–86.
30. Bhupathiraju SN, Pan A, Manson JE *et al*. Changes in coffee intake and subsequent risk of type 2 diabetes: three large cohorts of US men and women. *Diabetologia* 2014;**57:**1346–54.
31. Tunnicliffe JM, Shearer J. Coffee, glucoseand insulin resistance: physiological mechanisms and mediators. *Appl Physiol Nutr Metab* 2008;**33:**1290–300.
32. Greenberg JA, Boozer CN, Geliebter A. Coffee, diabetes, and weight control. *Am J Clin Nutr* 2006;**84:**682–93.
33. 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.
34. Puga A, Ma C, Marlowe JL. The aryl hydrocarbon receptor cross-talks with multiple signal transduction pathways. *Biochem Pharmacol* 2009;**77:**713–22.
35. Zhou SF, Wang B, Yang LP, Liu JP. Structure, function, regulation and polymorphism and the clinical significance of human cytochrome P450 1A2. *Drug Metab Rev* 2010;**42:**268–354.
36. McMahon G, Taylor AE, Davey Smith G, Munafo MR. Phenotype refinement strengthens the association of AHR and CYP1A1 genotype with caffeine consumption. *PLoS One* 2014; **9:**e103448.
37. Sivertsen SH. Diabetes in Denmark. 2014. http://diabetes.dk/presse/diabetes-i-tal/diabetes-i-danmark.aspx (10 May 2015, date last accessed).