See corresponding editorial on page 489.

Long-term coffee consumption, caffeine metabolism genetics, and risk of cardiovascular disease: a prospective analysis of up to 347,077 individuals and 8368 cases

Ang Zhou[1] *and Elina Hyppönen*[1,2,3]

[1]Australian Centre for Precision Health, University of South Australia Cancer Research Institute, Adelaide, Australia; [2]Population, Policy and Practice, UCL Institute of Child Health, London, United Kingdom; and [3]South Australian Health and Medical Research Institute, Adelaide, Australia

ABSTRACT

Background: Coffee is one of the most widely consumed stimulants worldwide and is generally considered to be safe or even beneficial for health. However, increased risk of myocardial infarction and hypertension has been suggested for individuals who carry a functional variant at cytochrome P450 1A2 (*CYP1A2*), which makes them less effective at metabolizing caffeine.

Objectives: The aim of this study was to examine if the *CYP1A2* genotype or a genetic score for caffeine metabolism (caffeine-GS) modifies the association between habitual coffee consumption and the risk of cardiovascular disease (CVD).

Methods: Genetic data and information on habitual coffee intake and relevant covariates were available for 347,077 individuals in the UK Biobank, including 8368 incident CVD cases. We used logistic regression to test for the association between coffee intake and CVD risk, and whether the association varies with *CYP1A2* genotype or caffeine-GS.

Results: The association between habitual coffee intake and CVD risk was nonlinear, and, compared with participants drinking 1–2 cups/day, the risk of CVD was elevated for nondrinkers, drinkers of decaffeinated coffee, and those who reported drinking >6 cups/day (increase in odds by 11%, 7%, and 22%, respectively, *P*-curvature = 0.013). *CYP1A2* genotype and caffeine-GS were not associated with CVD ($P \geq 0.22$ for all comparisons). There was no evidence for an interaction between the *CYP1A2* genotype or caffeine-GS and coffee intake with respect to risk of CVD ($P \geq 0.53$).

Conclusions: Heavy coffee consumption was associated with a modest increase in CVD risk, but this association was unaffected by genetic variants influencing caffeine metabolism. *Am J Clin Nutr* 2019;109:509–516.

Keywords: habitual coffee consumption, caffeine metabolism genetics, *CYP1A2*, cardiovascular disease, gene-by-coffee interaction, UK Biobank

Introduction

Coffee is one of the most widely consumed beverages worldwide (1), with an estimated 3 billion cups drunk daily (2). Hence, even relatively modest individual health effects of coffee would have important public health implications. The health benefits and risks of coffee have long been debated, with various outcomes reported (3). In terms of cardiovascular disease (CVD), although randomized controlled trials demonstrate that intake of caffeine, the most prominent bioactive compound of coffee, leads to acute increases in blood pressure (BP) (4, 5), evidence from observational studies looking at habitual coffee consumption typically does not indicate any increase in risk. A recent meta-analysis summarizing evidence from 36 prospective studies, involving a total of 1,279,804 participants and 36,352 cases of CVD (6), suggested a U-shaped association between habitual coffee intake and CVD risk, with the authors concluding that 3–5 cups/day is most beneficial, and further that heavy

This study was financially supported by the JJ Mason and HS Williams Memorial Foundation (CT23158) and the National Health and Medical Research Council, Australia (project 1123603). The funders had no role in the design, implementation, analysis, and interpretation of the data.

Supplemental Figures 1–4, Supplemental Tables 1–5, and Supplemental Material are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Address correspondence to EH (e-mail: Elina.Hypponen@unisa.edu.au).

Abbreviations used: *AHR*, aryl hydrocarbon receptor; BP, blood pressure; CAD, coronary artery disease; caffeine-GS, genetic score for caffeine metabolism; *CYP1A2*, cytochrome P450 1A2; CVD, cardiovascular disease; DBP, diastolic blood pressure; SBP, systolic blood pressure; SNP, single-nucleotide polymorphism.

Received June 26, 2018. Accepted for publication October 2, 2018.

First published online March 6, 2019; doi: https://doi.org/10.1093/ajcn/nqy297.

consumption is not associated with elevated risk. The possible cardioprotective effect could reflect the components of coffee that have antioxidant and anti-inflammatory actions (7), such as phenolic compounds, diterpenes, trigonelline, and melanoidins, possibly acting through epigenetic modifications (8).

In addition to studies looking into the direct associations with habitual coffee consumption, some gene-coffee interaction studies have suggested that the effect of coffee on CVD risk may be modified by a functional variant (rs762551) of the cytochrome P450 1A2 gene (*CYP1A2*), its encoded protein metabolizing ~95% of caffeine in the liver and exhibiting a wide inherited interindividual variability in activity (9). The *CYP1A2* variant is associated with the inducibility of CYP1A2, such that *CYP1A2* C allele carriers (~32% of the European population) metabolize caffeine less effectively than AA carriers (10, 11). In a matched case-control study involving 2014 cases of nonfatal myocardial infarction (MI), greater habitual coffee intake was associated with increased risk of MI but only among the "slow" caffeine metabolizers (12), with a trend toward a protective effect in the "fast" caffeine metabolizers. Evidence for a similar coffee-by-*CYP1A2* interaction has been observed with respect to BP and the risk of hypertension (13, 14), with the authors of 1 of these studies recommending that the slow caffeine metabolizers should abstain from coffee drinking (14).

Although case-control and high-risk cohort studies are effective in accumulating sufficient numbers of cases and boosting statistical power, they are more prone to selection bias, and results generated from these studies may not be directly applicable to the general population. The aim of the current study was to replicate the coffee-by-*CYP1A2* interaction with respect to CVD in a large population-based cohort, the UK Biobank, consisting of 347,077 individuals and 8368 CVD cases. In addition, with the recent discovery of further genetic determinants affecting caffeine metabolism (16), we also investigated the gene-coffee interaction using a genetic score constructed based on 8 genome-wide significant variants.

Methods

Participants

The UK Biobank is a large prospective cohort study with >500,000 participants aged 37–73 y recruited from 22 assessment centers across the United Kingdom between 2006 and 2010 with the aim of improving the prevention, diagnosis, and treatment of diseases of middle and old age (17, 18). The study included genome-wide genotyping of all participants and a wide range of data from health and lifestyle questionnaires; analyses of blood, urine, and saliva samples; and clinical assessments. Information on health outcomes has been enhanced by the linkage to electronic health records and mortality registrations. Details of the recruitment and data collection in the UK Biobank have been described elsewhere (19, 20). All participants provided informed consent (20). Participants have the ability to withdraw their consent at any time, and these analyses exclude 23 individuals who have done so to date. We have further restricted the analyses to individuals who were identified as white British Caucasians based on self-report and genetic profiling and further excluded participants with mismatched information between self-reported and genetic sex. Genetic profiling identified patterns of relatedness in the UK Biobank (21), and in this study we allow for a maximum of 2 members from each family. Final analyses were conducted among individuals with complete information on cardiovascular traits, coffee intake, and relevant covariates (n = 347,077, **Supplemental Figure 1**). The present study was conducted under UK Biobank application number 20175. The UK Biobank study was approved by the National Information Governance Board for Health and Social Care and the North West Multicentre Research Ethics Committee.

Habitual coffee intake

Information on coffee consumption (cups/day) was obtained via a question in which participants were asked to report "How many cups of coffee do you drink each day? (include decaffeinated coffee)." Among the coffee drinkers, a further question was asked about the types of coffee. We grouped the participants according to their daily coffee intake into 7 categories, namely nondrinkers, those drinking decaffeinated coffee, and those drinking caffeinated coffee, with further subgrouping based on the average daily amount (<1, 1–2, 3–4, 5–6, and >6 cups). Information on tea intake (cups/day) was also available. We approximated habitual caffeine intakes (mg/day) by combining information on coffee and tea consumption and assuming that 1 cup of coffee and tea contained approximately 75 and 40 mg of caffeine, respectively (22). Participants who indicated that they drank <1 cup of coffee or tea per day were assigned a value of 0.5. Individuals who reported drinking ≥15 cups of tea or coffee per day were grouped together as 15 cups/day. We examined the influence of participants reporting very high coffee or tea intakes (≥16 cups/day, N = 1437), but the treatment of these outliers did not alter the findings (data not shown). Further information on other sources of caffeine was not available in the UK Biobank and hence could not be included in the estimation of habitual caffeine intake. As a result, caffeine intake was only used as a secondary exposure.

Cardiovascular traits

Our primary outcome was incident CVD, with prevalent CVD and systolic and diastolic blood pressure (SBP and DBP) measured during the baseline assessment used as secondary outcomes. Incident CVD events (i.e., those who had their first CVD-related diagnosis after the baseline visit), including coronary artery disease (CAD), stroke, and peripheral artery disease, were identified using data linkage to Hospital Episodes Statistics and mortality data (23). Fatal cases of CVD were identified using mortality data. Cases were classified using International Classification of Diseases (versions 9 and 10), as well as Classification of Interventions and Procedures (**Supplemental Table 1**). Prevalent CVD cases were participants with a recording of a CVD diagnosis in the Hospital Episodes Statistics prior to the baseline visit. BP was measured using digital monitors (HEM-7015IT; Omron Healthcare Inc.). We calculated the mean SBP and DBP values from the 2 BP measurements, and according to common practice in studies of this kind (23, 24), accounted for medication use by adding 15 and 10 mm Hg to SBP and DBP, respectively, for participants reported to be taking antihypertensive medication (20.8% of individuals) (23).

Sensitivity analyses were conducted by adjusting for medication use in the models using BPs as the outcome, with no differences in findings compared to those reported (data not shown).

Covariates

A diverse range of covariates was considered, with all measures obtained during the baseline assessment. These covered basic demographics (age, sex, and location); anthropometric measures (BMI and waist circumference); lifestyle factors, including smoking (nonsmokers, ex-smokers, and those smoking <1–5, 6–10, 11–15, 16–20, 21–25, or >25 cigarettes/day), alcohol intake (never, special occasion only, and those drinking 1–3 times/month, 1 or 2 times/week, 3–4 times/week, or ≥5 times/week), intensity of physical activity (light, moderate, or vigorous), and habitual tea intake (<1, 1–2, 3–4, or >4 cups/day); general health indicators, including self-reported health status (poor, fair, good, or excellent), and long-term illness (no, or yes). Socioeconomic status was approximated using the Townsend deprivation index reflecting area deprivation (25) and education [none, National Vocational Qualification (NVQ)/Certificate of Secondary Education (CSE)/Advanced levels (A-levels), or degree/professional]. For smoking, 68% of current smokers did not indicate the number of cigarettes that they smoked each day. For those participants, we assigned the median of the number of cigarettes/day, 15.

CYP1A2 genotyping and genetic score for caffeine metabolism

UK Biobank was genotyped by 2 arrays that share 95% common content (UK BiLEVE array for 50,000 participants and UK Biobank Axiom array for 450,000 participants), covering ~800,000 single-nucleotide polymorphisms (SNPs) and insertion-deletion markers. The genotype data underwent centralized quality controls and ~96 million SNPs were imputed using the UK10K haplotype reference panel merged with the 1000 Genomes Phase 3 reference panel. For more detail, see http://biobank.ctsu.ox.ac.uk/showcase/label.cgi?id=100314). *CYP1A2* (rs762551) was directly genotyped, with a call rate of 99.8% and a minor allele frequency of 0.27, and its genotype distribution follows the Hardy-Weinberg equilibrium ($P = 0.75$). The genetic score for metabolism (caffeine-GS) was constructed by combining 8 genome-wide significant variants near aryl hydrocarbon receptor (*AHR*) and *CYP1A2* for caffeine metabolites, including rs4410790, rs6968554, rs10275488, rs2892838, rs12909047, rs35107470, rs2470893, and rs2472297 (16). *CYP1A2* (rs762551) was in modest linkage disequilibrium with the other *CYP1A2* variants (R^2 = 0.05–0.29), and it was excluded from the caffeine-GS as it was not among the top hits of the genome-wide study (16). A higher caffeine-GS score indicates slower caffeine metabolism. Information regarding the variants for the caffeine-GS and construction of the caffeine-GS is detailed in the **Supplemental Material** and **Supplemental Table 2**. Because *CYP1A2* rs762551 variant and caffeine-GS were highly correlated ($P < 1.0 \times 10^{-300}$), and the functional effect on caffeine metabolism has been determined only for *CYP1A2* in a caffeine challenge test (10, 11), our primary analyses were based on *CYP1A2*, and caffeine-GS was considered as a secondary genetic determinant of caffeine metabolism.

Statistical analyses

The associations between habitual coffee intake and incident CVDs were examined using multiple logistic regression, adjusting for baseline SBP, as well as a range of covariates covering demographic, anthropometric, lifestyle, general health, and socioeconomic aspects of participants (see the Covariates section above). Due to the nonlinear association between coffee intake and CVD, we used a continuous coffee intake indicator (cups/day) in the interaction analyses and included the quadratic term coffee^2 in the model to allow for the nonlinearity. The first model to test for the SNP-by-coffee interaction allowed for interaction with both the linear and quadratic terms of coffee.

$$\text{Model 1 : CVD} = \text{coffee} + \text{coffee}^2 + \text{SNP} + \text{coffee} \times \text{SNP} + \text{coffee}^2 \times \text{SNP} + \text{other covariates} \quad (1)$$

In the absence of interaction with the quadratic term, the model was then simplified to only allow for interaction with the linear term.

$$\text{Model 2 : CVD} = \text{coffee} + \text{coffee}^2 + \text{SNP} + \text{coffee} \times \text{SNP} + \text{other covariates} \quad (2)$$

The significance of the interaction was determined based on the highest level interaction term in the model, and here, lack of interaction was inferred when neither coffee^2 × SNP (Model 1) nor coffee × SNP (Model 2) were significant at the 5% level. Interactions by smoking and age were examined in the corresponding manner, replacing "SNP" above with smoking or age as relevant and by adding further 3-way interactions with coffee-SNP-smoking where relevant.

All models were weighted by (1 – kinship coefficient) (26) to account for relatedness, and models including genetic variants were additionally adjusted for the SNP genotyping array (UK BiLEVE array/UK Biobank Axiom array) and for the top 15 principal components accounting for population structure. All analyses were performed using STATA version 14.1 (StataCorp LP).

Results

Table 1 shows the patterns of coffee consumption and lists the incidence of CVD by participant characteristics. In contrast to notable variations seen with respect to these phenotypic indicators, distributions of demographics, lifestyle, general health, and socioeconomic factors were reasonably balanced across the *CYP1A2* rs762551 genotype and caffeine-GS quartiles (**Supplemental Table 3**). An exception was tea consumption, which slow caffeine metabolizers tended to drink less of than fast caffeine metabolizers ($P = 1.1 \times 10^{-30}$ for *CYP1A2* rs762551, $P = 2.5 \times 10^{-170}$ for caffeine-GS quartiles, Supplemental Table 3). The numbers of incident and prevalent CVD cases are shown in **Supplemental Table 4**. *CYP1A2* rs762551 C allele carriers (slow metabolizers) tended to drink less coffee than the

TABLE 1 Patterns of habitual coffee consumption and incidence of CVD by baseline characteristics in the UK Biobank[1]

		Coffee intake (cups/day)								Incident CVD[2]	
	N (%)	None, %	Decaf, %	<1, %	1–2, %	3–4, %	5–6, %	>6, %	*P*	Case, %	*P*
Age											
<65 y	278,548 (80.3)	21.3	15.3	5.8	30.7	17.4	7.2	2.4	0.53	2.1	1.60×10^{-265}
≥65 y	68,529 (19.7)	17.5	17.8	5.4	34.8	17.0	5.8	1.7		4.7	
Sex											
Male	160,108 (46.1)	18.8	12.7	5.9	32.2	19.3	8.3	2.9	$<1.0 \times 10^{-300}$	3.9	$<1.0 \times 10^{-300}$
Female	186,969 (53.9)	22.1	18.4	5.5	30.8	15.6	5.7	1.8		1.5	
BMI											
<18.5 kg/m^2	1729 (0.5)	25.8	15.2	6.1	30.5	13.7	5.7	3.0	6.51×10^{-30}	2.6	7.57×10^{-62}
18.5–<25 kg/m^2	114,924 (33.1)	21.0	15.9	6.1	33.8	16.0	5.4	1.9		1.8	
25–<30 kg/m^2	149,184 (43.0)	19.7	15.7	5.6	31.4	18.1	7.3	2.3		2.8	
≥30 kg/m^2	81,240 (23.4)	21.5	15.7	5.4	28.3	17.8	8.4	2.9		3.4	
Smoking											
Nonsmokers	190,951 (55.0)	21.7	17.3	5.9	31.8	16.4	5.6	1.5	$<1.0 \times 10^{-300}$	2.0	1.38×10^{-158}
Ex-smokers	122,555 (35.3)	19.2	15.0	5.6	32.6	18.2	7.2	2.1		3.0	
<1–5 cigs/d	2432 (0.70)	18.8	11.6	5.1	32.2	19.7	9.9	2.8		2.3	
6–10 cigs/d	5659 (1.63)	20.8	10.8	4.8	24.0	20.5	12.7	6.5		3.6	
11–15 cigs/d	16,408 (4.73)	17.8	10.6	5.3	28.4	20.5	11.8	5.7		4.1	
16–20 cigs/d	5866 (1.69)	21.0	8.9	4.0	19.3	17.5	17.6	11.8		5.9	
21–25 cigs/d	1420 (0.41)	21.2	7.9	3.7	19.2	17.0	18.5	12.7		7.5	
>25 cigs/d	1786 (0.51)	24.8	7.1	4.3	18.3	13.1	15.0	17.5		7.8	
Alcohol intake											
Nondrinkers	21,635 (6.2)	35.8	17.7	4.3	21.1	11.7	6.3	3.3	$<1.0 \times 10^{-300}$	3.4	2.51×10^{-33}
Special occasions or 1–3 times/mo	73,529 (21.2)	26.3	18.7	5.8	26.3	13.3	6.7	2.9		2.6	
1 or 2 times/wk	91,583 (26.4)	20.7	16.8	5.9	31.3	16.4	6.7	2.1		2.5	
3 or 4 times/wk	85,113 (24.5)	16.4	15.1	5.9	34.5	19.5	6.9	1.9		2.3	
Daily or almost daily	75,217 (21.7)	15.3	11.8	5.5	36.3	21.4	7.7	2.1		2.8	
Tea intake											
<1 cup/d	60,581 (17.5)	12.9	17.5	1.8	16.8	25.5	18.1	7.4	$<1.0 \times 10^{-300}$	2.6	8.14×10^{-10}
1–2 cups/d	75,010 (21.6)	10.0	16.4	3.0	35.2	24.7	8.7	1.9		2.4	
3–4 cups/d	102,641 (29.6)	19.0	16.4	6.3	37.4	17.2	3.3	0.6		2.5	
>4 cups/d	108,845 (31.4)	33.7	13.8	9.2	31.5	7.8	2.9	1.2		2.8	
Physical activity											
Light	105,271 (30.3)	21.4	15.9	5.9	29.3	17.4	7.6	2.6	1.25×10^{-12}	2.8	1.15×10^{-10}
Moderate	172,294 (49.6)	19.6	15.9	5.7	32.7	17.6	6.5	2.0		2.4	
Vigorous	69,512 (20.0)	21.7	15.4	5.3	31.9	16.5	6.8	2.6		2.8	
Education											
None	58,108 (16.7)	26.0	15.7	5.0	29.3	14.6	6.6	2.8	7.49×10^{-240}	4.2	1.49×10^{-34}
NVQ/CSE/A-levels	125,856 (36.3)	22.1	15.6	5.7	30.2	16.7	7.2	2.5		2.4	
Degree/professional	163,113 (47.0)	17.5	15.9	6.0	33.2	18.7	6.8	2.0		2.2	
Townsend deprivation index quartiles											
Q1 (lowest)	86,709 (25.0)	18.2	18.1	5.6	32.0	18.0	6.5	1.8	2.34×10^{-8}	2.4	8.16×10^{-41}
Q2	86,784 (25.0)	19.6	16.9	5.6	31.6	17.6	6.8	1.9		2.4	
Q3	86,811 (25.0)	20.6	15.6	5.9	31.6	17.2	6.9	2.3		2.6	
Q4 (highest)	86,773 (25.0)	24.0	12.5	5.8	30.7	16.4	7.4	3.2		3.0	
Self-rated health											
Excellent	59,487 (17.1)	18.4	15.2	5.6	33.8	18.7	6.4	1.9	1.09×10^{-86}	1.6	1.15×10^{-229}
Good	206,750 (59.6)	19.9	16.1	5.7	32.0	17.5	6.7	2.1		2.3	
Fair	68,256 (19.7)	23.3	15.6	5.6	28.9	16.1	7.7	2.9		3.8	
Poor	12,584 (3.63)	28.4	13.9	5.9	25.9	13.7	7.8	4.5		6.3	
Long-term illness											
No	237,405 (68.4)	19.8	15.3	5.7	32.4	17.9	6.8	2.1	3.16×10^{-109}	2.1	6.31×10^{-95}
Yes	109,672 (31.6)	22.4	16.8	5.7	29.5	15.9	7.1	2.6		3.8	

[1]P values were adjusted for age, sex, and assessment center, and also accounted for relatedness closer than second-degree relatives. A levels, Advanced levels; cig, cigarette; CSE, Certificate of Secondary Education; CVD, cardiovascular disease; decaf, decaffeinated; NVQ, National Vocational Qualification; Q, quartile.

[2]Includes mortality and all CVD-related hospitalizations.

others, and a similar strong pattern was seen with increasing quartiles in caffeine-GS (**Supplemental Table 5**). There was no association between *CYP1A2* rs762551 or caffeine-GS with any CVD incident or prevalent outcome ($P \geq 0.22$ for all, **Table 2** and Supplemental Table 5). In cross-sectional analyses, we observed a clear pattern for an association between slow caffeine metabolism (indexed by *CYP1A2* rs762551 C allele or a higher caffeine-GS quartile) and higher BP (Supplemental Table 5).

Coffee intake and incident CVD

In models adjusting for demographic, anthropometric, lifestyle, general health, and socioeconomic covariates in addition to baseline BP, compared with light coffee drinkers (1–2 cups/day), the odds of incident CVD were 11%, 7%, and 22% higher for nonhabitual, decaffeinated, and heavy drinkers (>6 cups/day), respectively (P-curvature = 0.013, Table 2 and **Figure 1**). This U-shaped pattern was mainly reflective of the

TABLE 2 Phenotypic and genetic associations with incident CVD[1]

		Incident CVD[2]		Incident CAD		Incident stroke		Fatal CVD	
	N (%)	Case (%)	OR (95% CI)	Case (%)	OR (95% CI)	Case (%)	OR (95% CI)	Case (%)	OR (95% CI)
Coffee intake									
Nondrinkers	66,140 (20.5)	1767 (2.7)	1.11 (1.04, 1.18)	1356 (2.1)	1.13 (1.06, 1.23)	523 (0.8)	1.10 (0.98, 1.24)	439 (0.7)	1.14 (0.99, 1.31)
Decaf drinkers	50,861 (15.8)	1266 (2.5)	1.07 (1.0, 1.15)	956 (1.9)	1.10 (1.01, 1.19)	406 (0.8)	1.11 (0.98, 1.25)	269 (0.5)	1.07 (0.92, 1.24)
<1 cup/d	18,023 (5.6)	453 (2.5)	1.01 (0.91, 1.12)	364 (2.0)	1.08 (0.97, 1.22)	130 (0.7)	0.98 (0.81, 1.20)	99 (0.6)	1.0 (0.80, 1.25)
1–2 cup/d	102,208 (31.7)	2512 (2.5)	Referent	1863 (1.8)	Referent	759 (0.8)	Referent	552 (0.6)	Referent
3–4 cups/d	56,237 (17.4)	1461 (2.6)	1.04 (0.97, 1.12)	1114 (2.0)	1.07 (0.99, 1.16)	425 (0.8)	1.01 (0.89, 1.14)	307 (0.6)	0.98 (0.85, 1.14)
5–6 cups/d	22,217 (6.9)	636 (2.9)	1.05 (0.95, 1.15)	499 (2.3)	1.11 (0.99, 1.23)	159 (0.7)	0.88 (0.74, 1.06)	161 (0.7)	1.07 (0.88, 1.30)
>6 cups/d	7246 (2.2)	273 (3.8)	1.22 (1.07, 1.40)	199 (2.8)	1.22 (1.04, 1.42)	78 (1.1)	1.19 (0.93, 1.52)	65 (0.9)	1.03 (0.78, 1.36)
P			0.012		0.014		0.13		0.50
P-curvature			0.013		0.0022		0.21		0.015
Caffeine intake[3]									
quintiles									
Q1 (lowest)	66,387 (20.5)	1641 (2.5)	Referent	1222 (1.9)	Referent	526 (0.8)	Referent	370 (0.6)	Referent
Q2	67,926 (21.0)	1696 (2.5)	0.95 (0.88, 1.02)	1286 (1.9)	0.96 (0.89, 1.04)	517 (0.8)	0.91 (0.80, 1.02)	372 (0.6)	0.90 (0.78, 1.04)
Q3	62,475 (19.3)	1484 (2.4)	0.87 (0.81, 0.94)	1148 (1.8)	0.90 (0.83, 0.98)	433 (0.7)	0.80 (0.70, 0.91)	346 (0.6)	0.89 (0.77, 1.03)
Q4	62,903 (19.4)	1596 (2.5)	0.91 (0.85, 0.98)	1223 (2.0)	0.93 (0.86, 1.01)	450 (0.7)	0.82 (0.72, 0.93)	366 (0.6)	0.90 (0.77, 1.04)
Q5 (highest)	64,025 (19.8)	1951 (3.0)	0.99 (0.92, 1.06)	1472 (2.3)	0.99 (0.92, 1.07)	554 (0.9)	0.92 (0.82, 1.05)	438 (0.7)	0.89 (0.77, 1.02)
P			0.0005		0.085		0.004		0.45
P-curvature			0.013		0.0083		0.16		0.010
CYP1A2 rs762551									
Fast	170,954 (52.9)	4439 (2.6)	Referent	3367 (2.0)	Referent	1329 (0.8)	Referent	987 (0.6)	Referent
Slow	152,176 (47.1)	3915 (2.6)	0.98 (0.94, 1.03)	2974 (2.0)	0.99 (0.94, 1.04)	1145 (0.8)	0.96 (0.88, 1.04)	904 (0.6)	1.02 (0.93, 1.12)
P			0.49		0.58		0.31		0.69
Caffeine-GS[4]									
quartiles									
Q1 (lowest)	80,725 (25.0)	2140 (2.7)	Referent	1613 (2.0)	Referent	638 (0.8)	Referent	504 (0.6)	Referent
Q2	69,252 (21.4)	1743 (2.5)	0.96 (0.90, 1.03)	1334 (1.9)	0.98 (0.91, 1.06)	497 (0.7)	0.92 (0.81, 1.03)	393 (0.6)	0.93 (0.82, 1.07)
Q3	87,657 (27.1)	2257 (2.6)	0.98 (0.92, 1.04)	1715 (2.0)	0.98 (0.92, 1.06)	667 (0.8)	0.97 (0.87, 1.08)	498 (0.6)	0.93 (0.82, 1.05)
Q4 (highest)	85,496 (26.5)	2214 (2.6)	1.0 (0.94, 1.06)	1679 (2.0)	1.01 (0.94, 1.08)	672 (0.8)	1.01 (0.91, 1.13)	496 (0.6)	0.97 (0.85, 1.1)
P			0.53		0.83		0.36		0.65
P-trend			0.93		0.79		0.68		0.59

[1] *P* values are from Wald test unless otherwise indicated. All models were weighted by (1 – kinship coefficient) to account for relatedness and adjusted for basic demographics (age, sex, and location), anthropometric measures (BMI and waist circumference), lifestyle factors (smoking, alcohol intake, habitual tea intake, and physical activity), general health indicators (self-reported health and long-term illness), socioeconomic status (Townsend deprivation index and education), and baseline systolic blood pressure. Models including *CYP1A2* rs762551 or caffeine-GS additionally adjusted for the top 15 principal components and SNP genotyping array. CAD, coronary artery disease; caffeine-GS, genetic score for caffeine metabolism; CVD, cardiovascular disease; *CYP1A2*, cytochrome P450 1A2; decaf, decaffeinated; Q, quintile; SNP, single-nucleotide polymorphism.

[2] Includes mortality and all CVD-related hospitalizations.

[3] Estimated based on daily consumption of coffee (75 mg/cup) and tea (40 mg/cup).

[4] A higher caffeine-GS score indicates slower caffeine metabolism.

association between coffee consumption and incident CAD (6078 cases), although the notably wider confidence intervals around effect estimates for associations with stroke (2092 cases) and fatal CVD (1892 cases) covered those seen for CAD in all groups (Table 2). A similar U-shaped association was observed for habitual caffeine intake combining information on coffee and tea intake (Table 2, *P*-curvature = 0.013 for incident CVD). In contrast, for prevalent cases, the odds of CVD was increased for nonhabitual and decaffeinated drinkers, but no increases in risk were seen for higher intake groups (Supplemental Table 5 and **Supplemental Figure 3**). There was a strong inverse association between habitual coffee consumption with SBP and DBP among coffee drinkers (**Supplemental Figure 4** and Supplemental Table 5). Compared with participants drinking 1–2 cups of caffeinated coffee per day, those drinking decaffeinated coffee had on average lower BP (Supplemental Figure 4 and Supplemental Table 5).

Effect modification by caffeine metabolism variants

In models investigating possible effect modification by *CYP1A2* rs762551 genotype or caffeine-GS on the association between coffee intake and incident CVD, we found no evidence for interaction, with all patterns between habitual coffee consumption and CVD being practically identical regardless of genotype ($P \geq 0.53$ for both, Figure 1 and **Supplemental Figure 2**). There also was no evidence for interaction between coffee consumption and *CYP1A2* genotype with respect to prevalent CVD and any BP indicator ($P \geq 0.78$ for all, Supplemental Figures 3 and 4). We also observed no evidence of gene-by-coffee-by-smoking or gene-by-coffee-by-age interactions with respect to any CVD outcome or BP ($P \geq 0.11$ for all, data not shown).

Discussion

In line with an earlier meta-analyses of prospective studies (6), we observed a nonlinear association between habitual coffee intake and CVD risk, with light coffee consumers (1–2 cups/day) having the lowest risk and an elevated risk seen for nonhabitual and decaffeinated coffee drinkers, as well as for heavy drinkers (>6 cups/day). However, unlike earlier smaller gene-by-coffee interaction studies (12, 14), we saw no evidence to suggest that the association between coffee consumption and long-term risk

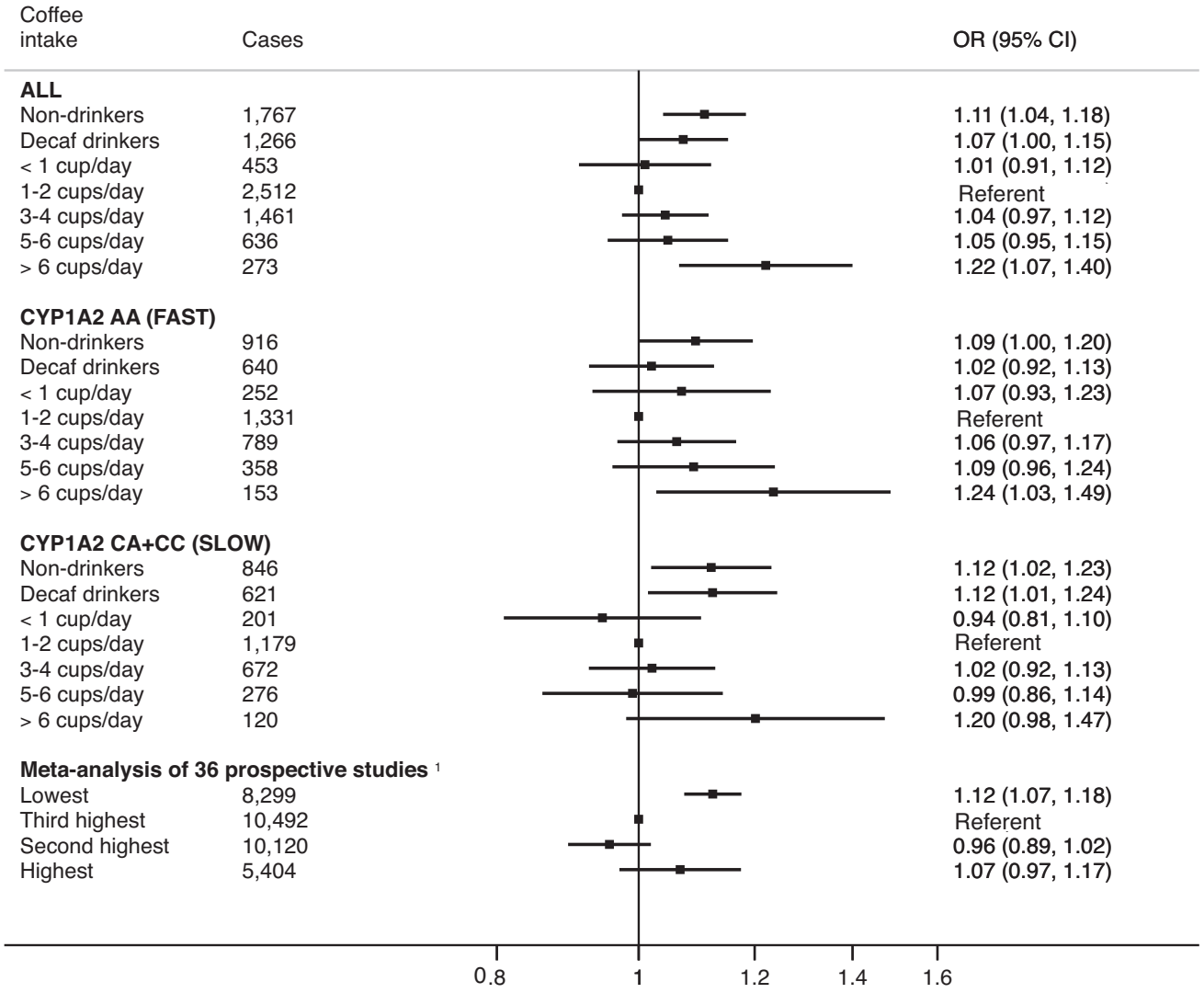

FIGURE 1 Association of habitual coffee intake with incident of CVD in the UK Biobank and in an earlier meta-analysis of 36 prospective studies involving 1,279,804 individuals and 36,352 cases (6). In the UK Biobank, results are presented for the whole cohort and stratified by *CYP1A2* genotype. All models were weighted by (1 – kinship coefficient) to account for relatedness and adjusted for basic demographics (age, sex, and location), anthropometric measures (BMI and waist circumference), lifestyle factors (smoking, alcohol intake, physical activity, and habitual tea intake), general health indicators (self-reported health and long-term illness), socioeconomic status (Townsend deprivation index and education), and baseline systolic blood pressure. For the meta-analysis by Ding et al. 2014 (6), adjustments vary between studies. Median and range (cups/day) for the lowest, third highest, second highest, and highest intake groups are 0 (0–2.5), 1.5 (0.5–3.5), 3.5 (1–6.5), and 5.5 (1.25–11.25), respectively. CVD, cardiovascular disease; *CYP1A2*, cytochrome P450 1A2; decaf, decaffeinated.

of CVD is affected by genetic variants influencing the rate of caffeine metabolism.

The U-shaped association between habitual coffee intake and CVD risk is consistent with a meta-analysis of 36 prospective studies involving 1,279,804 participants and 36,352 CVD cases (6), although the associations observed for moderate and heavy drinkers (compared with the light drinkers) in our study were stronger. In contrast, our study and the earlier meta-analysis (6) do not support findings from a recent study using the UK Biobank, which suggested a potential protective effect even for very high levels of coffee consumption with respect to mortality from CVD-related causes (15). However, this earlier study used nondrinkers as the reference group, and given that any benefit in their analyses was only seen for mortality outcomes in the first 3 y after the measurement of coffee intake, the proposed benefit is likely to reflect influences by reverse causation. The observed coffee-CVD associations resemble the U-shaped association often reported for alcohol intake and CVD risk (27, 28). Rather than reflecting increases in CVD risk by lack of alcohol (or coffee) consumption, it appears more likely that the observed elevations in risk reflect reverse causation and reductions in habitual drinking patterns by people with already compromised cardiohealth (29). Indeed, coffee consumption is known to be one of the first behaviors to be altered if our health status declines (30), highlighting the need for caution with respect to interpretation of these types of observational coffee-disease associations.

Caffeine is known to acutely increase BP (4, 5). The observed strong inverse association between coffee intake and BP in

cross-sectional analyses may suggest self-regulation in drinking behavior where people adjust coffee intake to maintain biological exposure of caffeine at levels supportive of normal BP. As previously reported (16), and confirmed in our study, similar self-regulation of behavior is seen in relation to genetic variants affecting caffeine metabolism: slow caffeine metabolizers tend to drink less coffee than fast metabolizers, potentially seeking to take advantage of the psychostimulant effects of caffeine (16). In the UK Biobank, we also found that slow caffeine metabolizers (indexed by *CYP1A2* rs762551 C allele or a higher caffeine-GS quartile) on average have a higher BP than fast metabolizers, suggesting that caffeine metabolism and BP may act in concert to affect coffee drinking behavior. This is consistent with the observation that there is an overlap in genetic architecture underlying BP regulation and caffeine metabolism, with *CYP1A2* being discovered in genome-wide association analyses for both traits (16, 24).

Unlike earlier studies suggesting that the *CYP1A2* rs762551 variant modulates the effect of habitual coffee intake on cardiovascular outcomes, including myocardial infarction and hypertension (12, 14), we observed no evidence to support related effect modification. In addition, smoking is known to induce CYP1A2 (31) and is strongly positively associated with coffee drinking (32). It is plausible that smoking behaviors may have moderated *CYP1A2* influences on CVD outcomes, but we saw no evidence for gene-by-coffee-by-smoking interactions. Given that earlier studies had much smaller sample sizes, if the *CYP1A2*-by-coffee interaction were real, we would have expected to see some evidence to support related influences in our notably larger study. Although it could be argued that the consistency between earlier studies is a testament to the robustness of the *CYP1A2*-by-coffee interaction (12, 14), it is perhaps more likely that the perceived consistency may merely reflect publication bias toward positive results, which is a well-established phenomenon in studies on gene-environment interaction (33).

Our study has several strengths and limitations to consider. First, the sample size available to us is unprecedented, and we were able to conduct analyses on incident CVD with adjustment for a comprehensive set of health, demographic, and lifestyle data in our analyses. Second, in addition to the *CYP1A2* rs76255 genotype used in the earlier studies, we included information for several other caffeine-related SNPs to construct a more comprehensive caffeine-GS. An important limitation of our study is that the UK Biobank is not representative of the general population (34) despite its large sample size. However, earlier publications from the UK Biobank have replicated expected exposure-disease associations (35, 36) and well-established gene-by-exposure interactions [such as *FTO*-by-physical activity interaction (37)], suggesting that this lack of representativeness may have had little influence on the results of the current study. Generalizability of our findings to other ethnic groups will need to be done with caution as analyses were restricted to participants of white British origin. We also cannot exclude information bias, as information on habitual coffee intakes was self-reported, and therefore, some degree of nondifferential misclassification bias is inevitable, which would bias the coffee-CVD associations and *CYP1A2*/caffeine-GS-by-coffee interactions toward null. In contrast, CVD incident cases were classified using linked hospital inpatient and death registry data, avoiding problems relating to self-report.

In conclusion, heavy coffee consumption was associated with a modest increase in CVD risk, and we found no evidence that this association is modified by genetic variants affecting caffeine metabolism.

The authors' responsibilities were as follows—EH: conceived the study and designed the research question; AZ: analyzed the data and wrote the first draft; and both authors: interpreted the results and drafted the manuscript, revised the manuscript critically for important intellectual content, and read and approved the final manuscript. Neither author has a conflict of interest to declare.

References

1. Fredholm BB, Battig K, Holmen J, Nehlig A, Zvartau EE. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. Pharmacol Rev 1999;51(1):83–133.
2. International Coffee Organization. Annual Review 2015–2016. [Internet]. Available at: http://www.ico.org/documents/cy2016-17/annual-review-2015-16-e.pdf (accessed 18 May 2018).
3. Cano-Marquina A, Tarin JJ, Cano A. The impact of coffee on health. Maturitas 2013;75(1):7–21.
4. Noordzij M, Uiterwaal CS, Arends LR, Kok FJ, Grobbee DE, Geleijnse JM. Blood pressure response to chronic intake of coffee and caffeine: a meta-analysis of randomized controlled trials. J Hypertens 2005;23(5):921–8.
5. 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(4):1113–26.
6. 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(6):643–59.
7. Grosso G, Godos J, Galvano F, Giovannucci EL. Coffee, caffeine, and health outcomes: an umbrella review. Annu Rev Nutr 2017;37:131–56.
8. Chuang YH, Quach A, Absher D, Assimes T, Horvath S, Ritz B. Coffee consumption is associated with DNA methylation levels of human blood. Eur J Hum Genet 2017;25(5):608–16.
9. Rasmussen BB, Brix TH, Kyvik KO, Brosen K. The interindividual differences in the 3-demthylation of caffeine alias CYP1A2 is determined by both genetic and environmental factors. Pharmacogenetics 2002;12(6):473–8.
10. Sachse C, Brockmoller J, Bauer S, Roots I. Functional significance of a C→A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine. Br J Clin Pharmacol 1999;47(4):445–9.
11. Castorena-Torres F, Mendoza-Cantu A, de Leon MB, Cisneros B, Zapata-Pérez O, López-Carrillo L, Salinas JE, Albores A. CYP1A2 phenotype and genotype in a population from the Carboniferous region of Coahuila, Mexico. Toxicol Lett 2005;156(3):331–9.
12. Cornelis MC, El-Sohemy A, Kabagambe EK, Campos H. Coffee, CYP1A2 genotype, and risk of myocardial infarction. JAMA 2006;295(10):1135–41.
13. Soares RN, Schneider A, Valle SC, Schenkel PC. The influence of CYP1A2 genotype in the blood pressure response to caffeine ingestion is affected by physical activity status and caffeine consumption level. Vasc Pharmacol 2018;106:67–73.
14. Palatini P, Ceolotto G, Ragazzo F, Dorigatti F, Saladini F, Papparella I, Mos L, Zanata G, Santonastaso M. CYP1A2 genotype modifies the association between coffee intake and the risk of hypertension. J Hypertens 2009;27(8):1594–601.
15. Loftfield E, Cornelis MC, Caporaso N, Yu K, Sinha R, Freedman N. Association of coffee drinking with mortality by genetic variation in caffeine metabolism: findings from the UK Biobank. JAMA Intern Med 2018;178(8):1086–97.
16. Cornelis MC, Kacprowski T, Menni C, Gustafsson S, Pivin E, Adamski J, Artati A, Eap CB, Ehret G, Friedrich N, et al. Genome-wide association study of caffeine metabolites provides new insights to caffeine metabolism and dietary caffeine-consumption behavior. Hum Mol Genet 2016;25(24):5472–82.

17. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12(3):e1001779.
18. Littlejohns TJ, Sudlow C, Allen NE, Collins R. UK Biobank: opportunities for cardiovascular research. Eur Heart J 2017;0:1–10. doi:10.1093/eurheartj/ehx254.
19. Elliott P, Peakman TC. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. Int J Epidemiol 2008;37(2):234–44.
20. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12(3):e1001779.
21. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. bioRxiv 2017.
22. 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(7):e103448.
23. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, Ntalla I, Surendran P, Liu C, Cook JP, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat Genet 2017;49(3):403–15.
24. Ehret GB, Ferreira T, Chasman DI, Jackson AU, Schmidt EM, Johnson T, Thorleifsson G, Luan J, Donnelly LA, Kanoni S, et al. The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. Nat Genet 2016;48(10):1171–84.
25. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the north. London: Routledge; 1988.
26. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010;26(22):2867–73.
27. Corrao G, Rubbiati L, Fau-Bagnardi V, Bagnardi V, Fau-Zambon A, Zambon A, Fau-Poikolainen K, Poikolainen K. Alcohol and coronary heart disease: a meta-analysis. Addiction 2000;95(10):1505–23.
28. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. BMJ 2011;342:d671.
29. Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, Ye Z, Prieto-Merino D, Dehghan A, Trompet S, Wong A, et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. BMJ 2014;349:g4164.
30. Soroko S, Chang J, Barrett-Connor E. Reasons for changing caffeinated coffee consumption: the Rancho Bernardo Study. J Am Coll Nutr 1996;15(1):97–101.
31. Hukkanen J, Jacob P, 3rd, Peng M, Dempsey D, Benowitz NL. Effect of nicotine on cytochrome P450 1A2 activity. BMJ 2011;342:d671.
32. Swanson JA, Lee JW, Hopp JW. Caffeine and nicotine: a review of their joint use and possible interactive effects in tobacco withdrawal. Addict Behav 1994;19(3):229–56.
33. Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. Am J Psychiatry 2011;168(10):1041–9.
34. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol 2017;186(9):1026–34.
35. Ganna A, Ingelsson E. 5 year mortality predictors in 498,103 UK Biobank participants: a prospective population-based study. Lancet 2015;386(9993):533–40.
36. Di Angelantonio E, Kaptoge S, Wormser D, Willeit P, Butterworth AS, Bansal N, O'Keeffe LM, Gao P, Wood AM, Burgess S, et al. Association of cardiometabolic multimorbidity with mortality. JAMA 2015;314(1):52–60.
37. Young AI, Wauthier F, Donnelly P. Multiple novel gene-by-environment interactions modify the effect of FTO variants on body mass index. Nat Commun 2016;7:12724.