Title

Assessment of moderate coffee consumption and risk of epithelial ovarian cancer: a Mendelian randomization study

Journal

International Journal of Epidemiology, 47(2)

ISSN

0300-5771

Authors

Ong, Jue-Sheng
Hwang, Liang-Dar
Cuellar-Partida, Gabriel
et al.

Publication Date

2018-04-01

DOI

10.1093/ije/dyx236

Peer reviewed

International Journal of Epidemiology, 2018, 450–459
doi: 10.1093/ije/dyx236
Advance Access Publication Date: 25 November 2017
Original article

Cancer

Assessment of moderate coffee consumption and risk of epithelial ovarian cancer: a Mendelian randomization study

Jue-Sheng Ong[1,2]* Liang-Dar Hwang,[1,2] Gabriel Cuellar-Partida,[1] Nicholas G Martin,[1] Georgia Chenevix-Trench,[1] Michael CJ Quinn,[1] Marilyn C Cornelis,[3] Puya Gharahkhani,[1] Penelope M Webb[4] and Stuart MacGregor;[1] Ovarian Cancer Association Consortium[#]

[1]Genetics and Computational Biology Department, QIMR Berghofer Medical Research Institute, Herston, QLD, Australia, [2]School of Medicine, University of Queensland, St Lucia, QLD, Australia, [3]Northwestern University Feinberg School of Medicine, Preventive Medicine, Chicago, IL, USA and [4]Population Health Department, QIMR Berghofer Medical Research Institute, Herston, QLD, Australia

*Corresponding author. Genetics and Computational Biology Department, QIMR Berghofer Medical Research Institute, 300 Herston Road, Herston, QLD 4006, Australia. E-mail: juesheng.ong@qimrberghofer.edu.au
[#]The full list of authors is to be found at the end of the article.

Editorial decision 12 October 2017; Accepted 21 October 2017

Abstract

Background: Coffee consumption has been shown to be associated with various health outcomes in observational studies. However, evidence for its association with epithelial ovarian cancer (EOC) is inconsistent and it is unclear whether these associations are causal.

Methods: We used single nucleotide polymorphisms associated with (i) coffee and (ii) caffeine consumption to perform Mendelian randomization (MR) on EOC risk. We conducted a two-sample MR using genetic data on 44 062 individuals of European ancestry from the Ovarian Cancer Association Consortium (OCAC), and combined instrumental variable estimates using a Wald-type ratio estimator.

Results: For all EOC cases, the causal odds ratio (COR) for genetically predicted consumption of one additional cup of coffee per day was 0.92 [95% confidence interval (CI): 0.79, 1.06]. The COR was 0.90 (95% CI: 0.73, 1.10) for high-grade serous EOC. The COR for genetically predicted consumption of an additional 80 mg caffeine was 1.01 (95% CI: 0.92, 1.11) for all EOC cases and 0.90 (95% CI: 0.73, 1.10) for high-grade serous cases.

Conclusions: We found no evidence indicative of a strong association between EOC risk and genetically predicted coffee or caffeine levels. However, our estimates were not statistically inconsistent with earlier observational studies and we were unable to rule out small protective associations.

Key words: Mendelian randomization, coffee, caffeine, causality, ovarian cancer

Key Messages

- Evidence for association between coffee and ovarian cancer is inconsistent and it is unclear whether the relationship is causal
- Results from this study indicate no evidence for a strong causal association between coffee intake and ovarian cancer susceptibility.
- A subsequent analysis on caffeine intake also found no causal link between caffeine intake and ovarian cancer.
- The Mendelian randomization estimates were consistent with observational findings of non-causality, but are unable to rule out small protective effects.

Introduction

Coffee is one of the most consumed beverages globally. A conventional cup of coffee can contain up to 1000 types of bioactive compounds including various kinds of antioxidants, aromatic compounds and, most importantly, caffeine. Caffeine has been found to suppress tumour growth in various animal models,[1,2] making it a potentially relevant therapeutic agent in cancer studies. Other compounds present in coffee are also found to have anti-inflammatory and anti-carcinogenic effects such as the induction of enzymes responsible for carcinogen detoxification, inhibition of carcinogen activation activities and stimulation of intracellular antioxidant defence.[1–3] Observational studies have investigated coffee and caffeine intake in relation to type 2 diabetes,[4,5] depression[6] and insomnia[7] as well as various cancers,[8,9] but the directions of association have been inconsistent across diseases.[10]

There are growing concerns regarding coffee consumption in relation to women's health. Epithelial ovarian cancer (EOC) is a gynaecological malignancy with a high fatality rate. Approximately 151 900 women worldwide die of the disease annually.[11] High-grade serous histology defines the most common EOC subtype.[12] Many individual studies have found conflicting directions of association with coffee consumption and EOC risk, but subsequent meta-analysis studies found no evidence for an association.[13–18] A more recent Danish study[19] suggested that moderate increase in daily caffeine intake (by one cup of coffee per day) might be protective against invasive EOC. Inconsistencies observed in the literature may be due to the lack of compatibility of categorical definitions (size of cup, content, caffeine intensity, method of brewing) and differences in definitions for baseline groups (i.e. non-drinkers). Some studies further combined consumption of tea and coffee to investigate caffeine intake specifically. However, more importantly, all studies to date examining the link between coffee/caffeine and EOC risk are observational studies where bias due to confounding may make it difficult to draw reliable conclusions.[20] For example, we can hypothesize that women diagnosed with EOC may have temporal nutritional awareness and develop aversion to caffeinated beverages (such as coffee and cola), which may distort the true underlying association in case-control studies. Since randomized trials examining coffee consumption in relation to ovarian risk have not been conducted, to work around these potential biases we can apply an instrumental variable technique, Mendelian randomization (MR),[21] to draw causal inferences on coffee consumption.

Twin studies have shown that coffee consumption has a substantial genetic component, with an estimated heritability ranging from 0.37 to 0.77.[22–24] This suggests that coffee consumption may be a suitable trait for MR studies. In this study, we aim to refine the relationship between coffee and EOC susceptibility. We hypothesize that genetic predisposition towards higher coffee intake is inversely associated with i) overall EOC susceptibility and ii) high-grade serous EOC susceptibility, and draw inference on causality via MR.

Methods

Data source

Participants for this study were drawn from the Ovarian Cancer Association Consortium (OCAC). Genotyping was performed using the customized Infinium OncoArray-500 K array (Illumina)[25] consisting of ~322 000 variants. OncoArray data were available for 59 115 samples across 71 study cohorts worldwide, of which 56 479 samples passed initial quality control protocols. Each individual was assigned values to indicate the proportion of European, African or Asian ancestry they inherited based on genetic make-up, using principal component analysis. These values sum up to 1 and are used to categorize the subjects into one of the intercontinental ancestry groups. Following that, imputation into the 1000 Genomes Project reference panel was carried out with pre-phasing using

SHAPEIT and IMPUTE2.[26,27] First-degree related individuals and duplicated samples ($n = 1732$) were removed. DNA samples from women of non-European ancestry were excluded for this study. The total sample size used in this study was 44 062 women of European ancestry (Table 1 shows a breakdown of the sample size by EOC histology). Baseline characteristics of our study samples from OCAC according to weight, age, smoking status and other potential confounders are summarized in Supplementary Table 1, available as Supplementary data at *IJE* online.

Genetic variants for the MR analyses were identified through an extensive review of published genome-wide association studies (GWAS) findings for coffee, tea and/or caffeine consumption.[28–33] Single nucleotide polymorphisms (SNPs) associated with coffee consumption (measured as cups/day) which were considered for use were rs1481012 in the *ABCG2* gene, rs6968554 in the *AHR* gene, rs2470893 in the *CYP1A2* gene, rs17685 in the *POR* gene and rs6265 in the *BDNF* gene. In our subsequent analysis, we investigated whether the association of coffee intake (in cups per day) with ovarian cancer was driven mainly by genetic predisposition for altered caffeine intake. SNPs reported to show association with caffeine and considered for use here were rs6968865 from the *AHR* gene and rs2472297 from the *CYP1A2* gene. All of the SNPs investigated were either directly genotyped or imputed with high quality (info-score > 0.9). Although these variants are different SNPs in AHR and CYP1A2, they are in high linkage-disequillibrium ($r^2 = 0.8$), see Discussion below for more detail. In order to ensure that our SNPs of interest were strong instruments, we examined the statistical evidence in the literature for their association with coffee and with caffeine consumption, respectively. The variance on coffee consumption explained by a particular SNP can be derived using $r^2_{SNP} = 2p(1-p)\beta^2/\sigma^2$, where r^2_{SNP} refers to the variance explained by the SNP, p refers to the minor allele frequency (MAF) of the SNP, β is the measured magnitude of association per effect allele and σ^2 is the coffee trait variance. The variance explained by our SNP instruments can hence be obtained by linearly summing up r^2_{SNP} across each independent SNP instrument. We subsequently tested each SNP against several potential confounders. For each of age at menarche, measures of glycaemia, education attainment, BMI, waist-hip ratio, body fat and smoking behaviour, we extracted previously published results from publicly available GWAS datasets (full details plus references in Supplementary Table 3, available as Supplementary data at *IJE* online).

Table 1. Distribution of EOC cases among European participants in OCAC

Nature/subtype	European cases
Invasive	17779
All serous[¥]	11213
Endometrioid	2199
Clear-cell	1121
Mucinous	1125
All mucinous[¥]	2023
High-grade serous	7488
Low-grade serous	880
All EOC cases[¥]	20683

A complete breakdown of the EOC cases by each participating study is provided in Supplementary material, available at *IJE* online.

[¥]Including unclassified and unknown serous/mucinous ovarian tumours.

Causal effect estimation

To perform MR, we used a two-sample statistical model to estimate the magnitude of association between coffee consumption and ovarian cancer, using summary statistics.[34] We fitted an additive model in SNPTEST[35] to test for association between each SNP and ovarian cancer status. Within-ancestry principal components (PC1-PC9) were fitted to remove potential bias arising from intra-ethnic population difference. Additional covariates that might be confounders, such as BMI, smoking status and alcohol consumption, were not available for all the genotyped OCAC participants and hence were not included as covariates (although subject to the assumptions of MR: not including these potential confounders as covariates will not bias our results) to maximize sample size. The genomic control lambda value was computed using 483 972 SNPs genome-wide to assess the possibility of population stratification biasing the association between allele frequencies and phenotype.

For both coffee and caffeine consumption, we used the Wald-type ratio estimator[36] to combine the SNP estimates, which uses the SNP-risk factor and SNP-cancer magnitude of association estimates to calculate the aggregated causal effect. We estimated a causal odds ratio (COR) for all ovarian cancer and for the high-grade serous subtype. High-grade serous was the only histological subtype with sufficient numbers for sub-set analysis.

Results

SNP selection

We shortlisted a total of four independent SNPs (rs1481012, rs6968554, rs2470893, rs17685) as proxies for genetically determined coffee consumption behaviour.[31] For the analysis on caffeine, we used two SNPs (rs6968865,rs2472297)[33] as genetic proxies for total

caffeine consumption per day (in mg). Each of these SNPs is robustly associated with *P*-values less than $P < 5 \times 10^{-8}$ for coffee consumption in the original coffee GWAS. Due to the smaller sample size in the published analysis for caffeine consumption, the published *P*-values for the effects of rs6968865 and rs2472297 on caffeine consumption were not as strong as those for the SNP-coffee associations, but both of the SNPs combined associate with caffeine consumption with a *P*-value $= 3.74 \times 10^{-14}$,[33] with its direction of association verified in an Australian sample (Supplementary A1, available as Supplementary data at *IJE* online). Each of the SNPs thus satisfies the strong MR instrument criterion ($F \gg 10$).

In our pleiotropy assessment, the SNP rs6265 in the *BDNF* gene was found to have pleiotropic effects on other traits of relevance to ovarian cancer (BMI and age of menarche, see supplementary material available as Supplementary data at *IJE* online), so it was excluded from our analyses. After removing BDNF, the four coffee SNPs combined explain ~1.2% of the variation in coffee intake,[31] whereas the two SNPs combined for our MR caffeine study explain ~1.3% of the variation in caffeine intake.[33] We also tested the association between established ovarian cancer risk factors (oral contraceptive use, estrogen use, parity) and our SNPs of interest. The results of our pleiotropy assessment are available in Supplementary Table 3 (publicly available GWAS) and Supplementary Table 4 (OCAC dataset; available as Supplementary data at *IJE* online). In brief, no associations were found above chance level, and we conclude that the assumptions of no-pleiotropy are not violated. In particular, coffee consumption and cigarette consumption are correlated in some populations, but our chosen SNPs are not associated with smoking (Supplementary Table 3).

Instrumental variable analysis

The SNP-cancer association results for each genetic instrument used are available in Supplementary Table 2, available as Supplementary data at *IJE* online. We estimated the causal odds ratio associated with a genetically predicted one cup per day change in coffee consumption. For all EOC cases, the COR for consuming one additional cup of coffee per day was 0.92 (95% CI: 0.79, 1.06). For high-grade serous EOC, the COR was 0.90 (95% CI: 0.73, 1.10). We also performed an additional analysis to investigate caffeine consumption, with the COR scaled in terms of an 80-mg increase (the approximate caffeine content in a conventional cup of coffee). The COR for consuming an additional 80 mg of caffeine was 1.01 (CI: 0.92, 1.11) for all EOC cases and 0.90 (CI: 0.73, 1.10) for high-grade serous cases. The CORs derived from individual SNP instruments are shown in Figure 1 for coffee consumption and Figure 2 for caffeine intake.

Population stratification and confounding

Due to the missing covariate data on some OCAC participants (see Supplementary Table 1), the analyses were performed by only fitting the first nine genetic (ancestral) principal components as covariates. In a sensitivity analysis using participants with confounder data available ($n \sim 11\,400$), adjustment for potential confounders (age of menarche, education level, number of pregnancies, oral contraceptive use, estrogen use, smoking and BMI) did not change the magnitude of the SNP-disease associations (See Supplementary Table 6, available as Supplementary data at *IJE* online). The genomic control lambda was 1.076 ($\lambda_{1000} = 1.007$, LD-score intercept = 1.032), demonstrating that there is little evidence for inflation of the genome-wide association statistics due to population stratification. Plots of the ancestral principal components (PC1 against PC2) between cases and controls indicate that the cases and controls are homogeneous (See Supplementary Figures 1 and 2, available as Supplementary data at *IJE* online).

Discussion

In our study sample of 44 062 European participants from OCAC, we found no evidence suggestive of a large causal association between (genetically predicted) coffee consumption and overall EOC risk or risk of high-grade serous EOC. Similarly, our findings consistently suggest no causal link between caffeine intake and EOC susceptibility.

Research in context

Most epidemiological studies in the past investigated the association of EOC with coffee consumption by assessing the difference in EOC risk among non-coffee drinkers and strong coffee drinkers. Consumption of > 3 cups of coffee per day was used as a benchmark to indicate strong coffee drinking behaviour. To compare our results, we rescaled findings from these observational studies to reflect an averaged moderate change in daily coffee consumption (1 cup of coffee per day) using Equation 1 in Supplementary material, available as Supplementary data at *IJE* online. The resultant estimates from our study were broadly compatible with results of previous meta-analyses (Figure 3).

Although some individual observational studies have found associations between coffee consumption and risk of EOC, meta-analyses have found no evidence to show that coffee consumption protects against EOC.[13] However, a

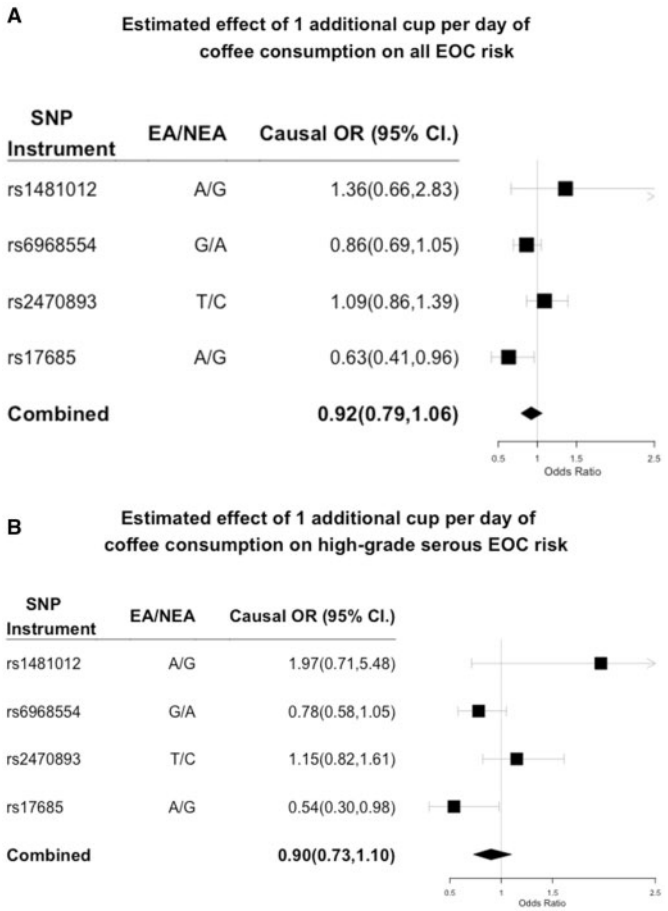

Figure 1. (A) Instrumental variable estimate for coffee consumption on EOC susceptibility. (B) Instrumental variable for coffee consumption on high-grade serous EOC susceptibility.

common criticism of observational studies is inconsistency in the definition of categorized consumption (i.e. different studies adopt different definitions of heavy drinkers) and the variability in types of coffee beverages, which may differ strongly in terms of nutritional content (most importantly, caffeine). These systematic differences can make the interpretation of meta-analysed findings difficult. Moreover, it is difficult to rule out the potential effects of selection bias in case-control studies and of unmeasured or uncontrolled confounding in observational studies in general. In contrast, here we use genetically predicted coffee intake to provide more uniform estimates of coffee consumption in a large sample size (coffee GWAS,[31] $n > 80\,000$). Our two-sample MR design allows us to investigate the underlying association without the issue of potential confounders such as education level, alcohol use and smoking behaviour, which were established by earlier studies to be strongly correlated to coffee consumption. In our pleiotropy assessment, the SNP instruments we employ are not associated with these potential confounders (Supplementary Table 3).

Even though the MR analyses were performed separately for coffee consumption and caffeine intake with independent SNPs within each study, the inference we draw from these findings are not independent. This is due to the fact that for each study the most important single SNPs (rs2470893 in *CYP1A2*, which explains ~0.5% of the variance in coffee consumption,[31] and rs2472297 in *CYP1A2*, which explains ~0.8% of the variance in caffeine consumption[33]) are in high linkage disequilibrium ($r^2 = 0.7$ between the two SNPs). Hence, the effect of those SNPs (rs2470893, rs2472297) on coffee and caffeine consumption may not be separable (i.e. *CYP1A2* is involved in metabolizing common bioactive compounds in coffee). The same applies for SNPs rs6968865 and rs6968554 in *AHR*.

Previous studies have highlighted a potential role of caffeine in inducing p53-dependent (tumour suppression gene) apoptosis.[37] Since *TP53* mutations are found in almost all high-grade serous EOC,[38] an analysis of high-grade serous EOC alone was of particular interest.

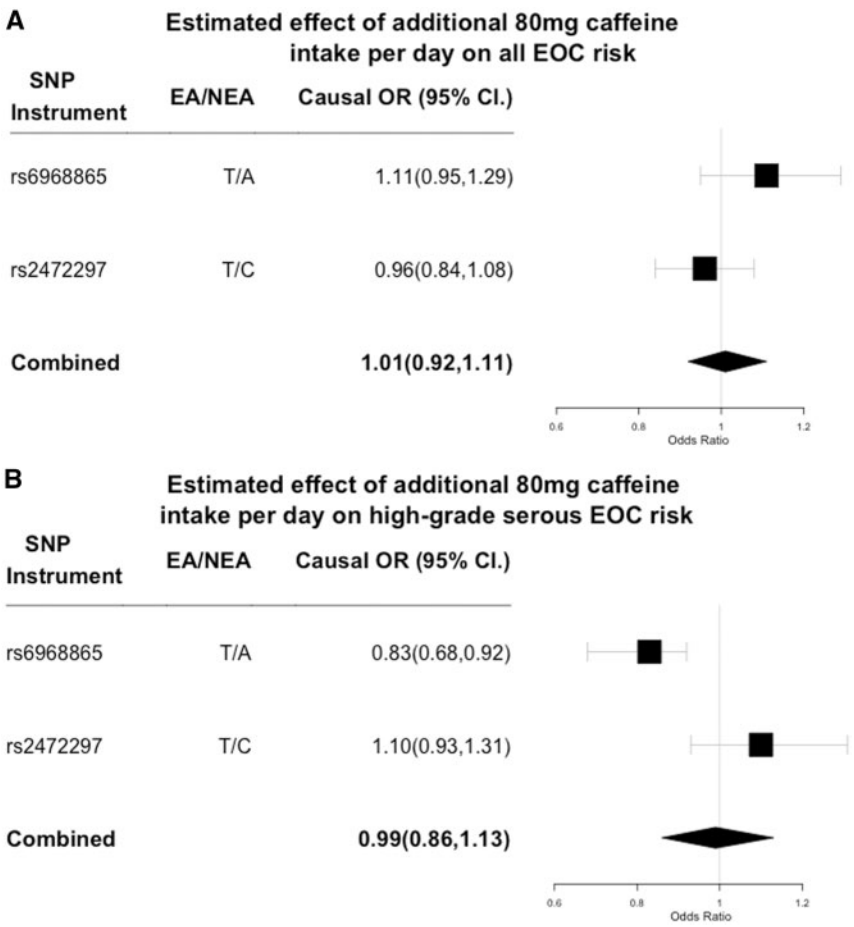

Figure 2. (A) Instrumental variable estimate for caffeine intake on EOC susceptibility. (B) Instrumental variable estimate for caffeine intake on high-grade serous EOC susceptibility.

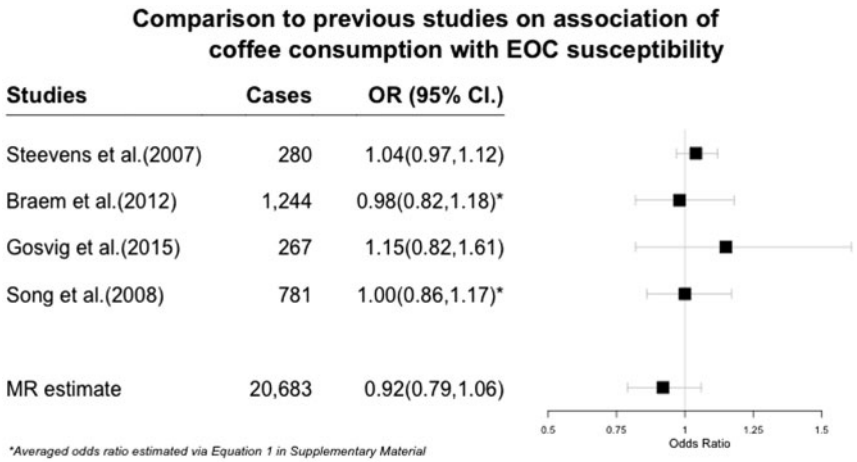

Figure 3. Comparison of Instrumental variable findings with observational studies.

However in our study, coffee and caffeine intake did not appear to be associated with any risk of high-grade serous carcinoma among Europeans.

Strengths and limitations

One of the strengths of our study is that participants used in our analyses were all of European ancestry, limiting potential bias due to population heterogeneity. Furthermore, the use of ancestral principal components to define ethnicity also prevents heritage-reporting errors (i.e. ethnicity was determined based on SNP profiles, as summarized by ancestry principal components, to avoid self-reporting biases). In our MR study, the use of GWAS findings to predict coffee/caffeine consumption rather than relying on self-reports of consumption should remove misclassification biases that can plague self-reported studies and contribute to statistical heterogeneity in meta-analyses of observational studies. Since coffee consumption generally stabilizes during adulthood, our two-sample MR approach is protected by potential biases due to apparent age differences between the SNP-coffee samples and the OCAC samples. In other words, the estimated SNP-coffee association during adulthood remains a robust genetic predisposition to lifetime coffee intake behaviour.

For our MR to infer about causality, several MR assumptions have to be met. First, the instruments (SNPs) used here were robustly associated (with $F >> 10$) with coffee and caffeine intake, respectively. Second, the SNPs used in this study showed no evidence for any pleiotropic effects that may confound the association with EOC susceptibility. The third MR assumption, that the genetic variants used in our study only influence EOC susceptibility through mediating coffee consumption, can be difficult to test directly. However, previous studies have examined the role of *CYP1A1*, *CYP1A2* and *AHR* in detail.[28,29,32,39] In each case, an SNP in or near the gene has been implicated by GWAS, and we assume that the action of the SNP on coffee consumption is via the specified gene. Taking each in turn, *CYP1A2* encodes the primary enzyme that metabolizes caffeine in the liver, whereas *CYP1A1* encodes protein that metabolizes polycyclic aromatic hydrocarbons, which are more commonly found in coffee beans. The *AHR* gene is known to induce both *CYP1A1* and *CYP1A2* via a DNA-binding mechanism,[29] and is also has a role in detection of toxic chemicals.[39] Despite coffee intake being strongly correlated with smoking, our pleiotropy assessment indicated that none of the SNPs appear to be associated (Bonferroni corrected P-value > 0.05) with smoking behaviours. Moreover, the lack of a main effect of the SNPs on smoking makes a coffee-smoking interaction less likely. Thus, it seems very improbable that these SNPs directly influence ovarian cancer through other independent biological processes.

Although we found no evidence supportive of an association between the SNPs used and common risk factors for EOC[40,41] (e.g. smoking, oral contraceptive use, parity etc.), it is hard to rule out directly possibilities of residual pleiotropy. However, suppose that a SNP has a strong pleiotropic effect which biases our results; for us to observe the null causal odds ratio we find here, the other SNPs (or some combination of SNPs) must act pleiotropically in the opposite direction and with similar magnitude to the first SNP. Since this is unlikely, it is unlikely that pleiotropic effects have a considerable influence on our non-causality conclusion.

There are some limitations that should also be considered in our analyses. First, our study was performed using only women of European ancestry and our findings may not generalize to other populations. Even though our SNPs greatly exceed the traditional strong instruments criteria ($F > 10$), our SNPs combined only account for a relatively small proportion of variation ($\sim$1.2%) in coffee consumption (cups per day), potentially leading to problems in power when applying MR. With a relatively small proportion of variance explained, we must extrapolate from small changes in predicted coffee consumption. If the sample size in our data linking genotype to ovarian cancer risk were small, the overall estimates of the causal odds ratio would be too large to be useful. However, as we have available a large dataset from an international consortium, the overall standard error in our causal odds ratio is relatively small, allowing us to make clear statements on the likely limits of the causal effect of coffee consumption on ovarian cancer (e.g. for all histologies, the causal OR is 0.92 with 95% confidence interval 0.79, 1.06).

The precision of our estimates is good for the most common subtype high-grade serous (COR 0.90 with 95% CI: 0.73, 1.10), but for the less common subtypes taken individually, our power is low; we similarly have insufficient power to perform stratified analyses (e.g. based on groups with particular smoking or BMI status).

The difference in coffee consumption as quantified in our MR analysis can be hard to interpret. In our analysis, CORs are calculated based on one additional cup of coffee per day averaging across all possible quantities of coffee consumption among regular coffee drinkers (including non-drinkers). This made it difficult to compare our estimates reliably with those from studies that investigated extreme ends of the trait distribution [ranging from heavy coffee drinking (> 5 cups) and/or coffee drinkers to non-drinkers]. Here, it is difficult for our study to completely rule out previous findings that showed positive associations of EOC when comparing very heavy coffee-drinkers

with other categories.[13] That is, our findings only infer that moderate differences in coffee consumption (averaging over the entire trait distribution) do not influence risk of EOC, as the MR framework assumes that modifiable exposures linearly affect the underlying risk factor; this might be violated if the outcome to exposure relationship is non-linear (follows a J-shaped curve).

An additional consideration is how to handle non-coffee drinkers. For caffeine this is not an issue, because non-users are included in the SNP association studies. For coffee consumption, in our main analysis we focus on 'cups per day' coffee consumption. However, the GWASs to date on 'cups per day' in coffee consumers also found[31] that the same SNPs were also strongly associated with drinking status 'high' versus 'low/no' coffee consumption). Hence our findings in support of non-causality of 'cups per day' probably extend to alternative definitions such as 'high' versus 'low/no' status.

We found no evidence indicative of a strong association between EOC risk and genetically predicted coffee or caffeine levels. However, our estimates were not statistically inconsistent with earlier observational studies, and we were unable to rule out small protective associations. Our MR results are more readily interpretable than previous observational studies, because they are unlikely to be adversely affected by confounding biases which can invalidate the conclusions from observational studies.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

This work was supported by the COGS project which is funded through: a European Commission Seventh Framework Programme Grant (agreement number 223175 HEALTH F2 2009–223175); and the Genetic Associations and Mechanisms in Oncology (GAME- ON), an NCI Cancer Post-GWAS Initiative (U19-CA148112). The Ovarian Cancer Association Consortium is supported by a grant from the Ovarian Cancer Research Fund, thanks to donations by the family and friends of Kathryn Sladek Smith (PPD/RPCI.07). Funding of the constituent studies was provided by: the California Cancer Research Program (00–01389V-20170, N01-CN25403, 2II0200); the Canadian Institutes of Health Research (MOP-86727); Cancer Australia; Cancer Council Victoria; Cancer Council Queensland; Cancer Council New South Wales; Cancer Council South Australia; Cancer Council Tasmania; Cancer Foundation of Western Australia; the Cancer Institute of New Jersey; Cancer Research UK (C490/A6187, C490/A10119, C490/A10124); the Danish Cancer Society (94–222–52); the ELAN Program of the University of Erlangen-Nuremberg; the Eve Appeal; the Helsinki University Central Hospital Research Fund; Helse Vest; grants from the Swedish Cancer Society and European Research Council Advanced Grant ERC-2011–294576; the Norwegian Cancer Society; the Norwegian Research Council; the Ovarian Cancer Research Fund; Nationaal Kankerplan of Belgium; the L & S Milken Foundation; the Polish Ministry of Science and Higher Education (4 PO5C 028 14, 2 PO5A 068 27); the Roswell Park Cancer Institute Alliance Foundation; Breast Cancer Now, the Institute of Cancer Research, NHS funding to the Royal Marsden/ICR NIHR Biomedical Research Centre; the US National Cancer Institute (K07-CA095666, K07-CA80668, K07-CA143047, K22-CA138563, N01-CN55424, N01-PC67001, N01-PC067010, N01-PC035137, P01-CA017054, P01-CA087696, P30-CA072720, P30-CA15083, P30-CA008748, P50-CA159981, P50-CA105009, P50-CA136393, R01-CA149429, R01-CA014089, R01-CA016056, R01-CA017054, R01-CA049449, R01-CA050385, R01-CA054419, R01-CA058598, R01-CA058860, R01-CA061107, R01-CA061132, R01-CA063678, R01-CA063682, R01-CA067262, R01-CA071766, R01-CA074850, R01-CA080978, R01-CA083918, R01-CA087538, R01-CA092044, R01-CA095023, R01-CA122443, R01-CA112523, R01-CA114343, R01-CA126841, R01-CA136924, R03-CA113148, R03-CA115195, U01-CA069417, U01-CA071966, UM1-CA186107, UM1-CA176726 and intramural research funds); the NIH/National Center for Research Resources/General Clinical Research Center (MO1-RR000056); the US Army Medical Research and Material Command (DAMD17–01–1–0729, DAMD17–02–1–0666, DAMD17–02–1–0669, W81XWH-07–0449, W81XWH-10–1–02802); the US Public Health Service (PSA-042205); the National Health and Medical Research Council of Australia (199600, 400281 and 1025142); the German Federal Ministry of Education and Research of Germany Programme of Clinical Biomedical Research (01GB 9401); the State of Baden-Wurttemberg through Medical Faculty of the University of Ulm (P.685); the German Cancer Research Center; the Minnesota Ovarian Cancer Alliance; the Mayo Foundation; the Fred C. and Katherine B. Andersen Foundation; the Lon V. Smith Foundation (LVS-39420); the Oak Foundation; Eve Appeal; the OHSU Foundation; the Mermaid I project; the Rudolf-Bartling Foundation; the UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge, Imperial College London, University College Hospital 'Womens' Health Theme' and the Royal Marsden Hospital; and Work Safe BC 14. Investigator-specific funding: J.S. and G.C. thank the University of Queensland and QIMR Berghofer Medical Research Institute for scholarship support. G.C.T. and P.M.W. are supported by the National Health and Medical Research Council (NHMRC). S.M. acknowledges funding support from an Australian Research Council Future Fellowship. This work is also supported by the NHMRC project grant 1123248.

Conflict of interest: M.T.G. is a consultant for Johnson and Johnson Ltd. U.M. has stock ownership and research funding from Abcodia Ltd, a UCL spin-out company with interest in biomarkers and ovarian cancer screening. P.F. conducts research with grants from Amgen and Novartis, and received honoraria from Amgen, Novartis, Roche, Celgene, Nanostring, Genomic Health and TEVA. J.D.B. holds stock in Inivata, the makers of DNA assays, but declares no conflict of interest regarding this work. All remaining authors declare no competing interest.

Acknowledgements

We acknowledge with appreciation all the individuals who participated in the OCAC studies, the many hospital directors and staff, gynaecologists, general practitioners and pathology staff who

provided assistance with confirmation of diagnoses and the many research assistants and interviewers for assistance with the studies. For the OCAC studies, we thank: D Bowtell, A deFazio, D Gertig, A Green, P Parsons, N Hayward, and D Whiteman (AUS); G Peuteman, T Van Brussel and D Smeets (BEL); L Gacucova (HMO); P Schurmann, F Kramer, W Zheng, T W Park, Simon, K Beer-Grondke and D Schmidt (HJO); S Windebank, C Hilker and J Vollenweider (MAY); the state cancer registries of AL, AZ, AR, CA, CO, CT, DE, FL, GA, HI,ID,IL,IN,IA,KY,LA,ME, MD,MA,MI,NE,NH,NJ,NY,NC,ND, OH,OK,OR,PA,RI,SC,TN,TX,VA,WA and WYL(NHS); L Paddock, M King, L Rodriguez-Rodriguez, A Samoila and Y Bensman (NJO); M Sherman, A Hutchinson, N Szeszenia-Dabrowska, B Peplonska, W Zatonski, A Soni, P Chao and M Stagner (POL); C Luccarini, P Harrington, the SEARCH team and ECRIC (SEA); I Jacobs, M Widschwendter, E Wozniak, N Balogun, A Ryan, C Karpinskyj and J Ford (UKO); Carole Pye (UKR); A Amin Al Olama, J Dennis, E Dicks, K Michilaidou and K Kuchenbaker (COGS).

*Full list of authors: Jue-Sheng Ong, Liang-Dar Hwang, Gabriel Cuellar-Partida, Australian Ovarian Cancer Study, Enda Bryne, Peter A Fasching, Alexander Hein, Stefanie Burghaus, Matthias W Beckmann, Diether Lambrechts, Els Van Nieuwenhuysen, Ignace Vergote, Adriaan Vanderstichele, Anthony J Swerdlow, Michael Jones, Nicholas Orr, Minouk Schoemaker, Digna Velez Edwards, James Brenton, Javier Benítez, María J García, Cristina Rodriguez-Antona, Mary Anne Rossing, Renée T Fortner, Elio Riboli, Jenny Chang-Claude, Ursula Eilber, Shan Wang-Gohrke, Drakoulis Yannoukakos, Marc T Goodman, Natalia Bogdanova, Thilo Dörk, Matthias Duerst, Peter Hillemanns, Ingo B Runnebaum, Natalia Antonenkova, Ralf Butzow, Heli Nevanlinna, Liisa M Pelttari, Robert P Edwards, Joseph L Kelley, Francesmary Modugno, Kirsten B Moysich, Roberta B Ness, Rikki Cannioto, Florian Heitz, Beth Karlan, Håkan Olsson, Susanne K Kjaer, Allan Jensen, Graham G Giles, Fiona Bruinsma, Michelle A T Hildebrandt, Dong Liang, Xifeng Wu, Loic Le Marchand, V Wendy Setiawan, Jennifer B Permuth, Maria Bisogna, Fanny Dao, Douglas A Levine, Daniel W Cramer, Kathryn L Terry, Shelley S Tworoger, Meir Stampfer, Walter Willet, Stacey Missmer, Line Bjorge, Reidun K Kopperud, Katharina Bischof, Liv Cecilie Vestrheim Thomsen, Lambertus A Kiemeney, Leon FAG Massuger, Tanja Pejovic, Angela Brooks-Wilson, Sara H Olson, Valerie McGuire, Joseph H Rothstein, Weiva Sieh, Alice S Whittemore, Linda S Cook, Nhu D Le, C Blake Gilks, Jacek Gronwald, Anna Jakubowska, Jan Lubiński, Tomasz Kluz, Nicolas Wentzensen, Louise Brinton, Britton Trabert, Jolanta Lissowska, Estrid Høgdall, Claus K Høgdall, Dale P Sandler, Alicja Wolk, Jonathan P Tyrer, Honglin Song, Diana Eccles, Ian Campbell, Ros Glasspool, Iain McNeish, James Paul, Nadeem Siddiqui, Rebecca Sutphen, John R McLaughlin, Catherine Phelan†, Hoda Anton-Culver, Argyrios Ziogas, Taymaa May, Simon A Gayther, Aleksandra Gentry-Maharaj, Usha Menon, Susan J Ramus, Anna H Wu, David Huntsman, Anna deFazio, Agnieszka Dansonka-Mieszkowska, Jolanta Kupryjanczyk, Joanna Moes-Sosnowska, Lukasz Michal Szafron, Julie M Cunningham, Stacey J Winham, Harvey A Risch, Ellen L Goode, Joellen M Schildkraut, Celeste L Pearce, Andrew Berchuck, Paul D P Pharoah, Nicholas G Martin, Georgia Chenevix-Trench, Michael C J Quinn, Marilyn C Cornelis, Puya Gharahkhani, Penelope M Webb and Stuart MacGregor.

*Full author details are available in a supplementary file at *IJE* online. †Deceased.

References

1. Huber WW, Rossmanith W, Grusch M, Haslinger E *et al*. Effects of coffee and its chemopreventive components kahweol and cafestol on cytochrome P450 and sulfotransferase in rat liver. *Food Chem Toxicol* 2008;**46:**1230–38.
2. Huber WW, Scharf G, Nagel G, Prustomersky S, Schulte-Hermann R, Kaina B. Coffee and its chemopreventive components Kahweol and Cafestol increase the activity of O6-methylguanine-DNA methyltransferase in rat liver - comparison with phase II xenobiotic metabolism. *Mutat Res* 2003; **522:**57–68.
3. Boettler U, Volz N, Teller N *et al*. Induction of antioxidative Nrf2 gene transcription by coffee in humans: depending on genotype?. *Mol Biol Rep* 2012;**39:**7155–62.
4. 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.
5. Lee JK, Kim K, Ahn Y, Yang M, Lee JE. Habitual coffee intake, genetic polymorphisms, and type 2 diabetes. *Eur J Endocrinol* 2015;**172:**595–601.
6. Lucas M, Mirzaei F, Pan A *et al*. Coffee, caffeine, and risk of depression among women. *Arch Intern Med* 2011;**171:** 1571–78.
7. Skarupke C, Schlack R, Lange K *et al*. Insomnia complaints and substance use in German adolescents: did we underestimate the role of coffee consumption? Results of the KiGGS study. *J Neurol Transm (Vienna)* 2017;**174(Suppl 1):** 69–78.
8. 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.
9. Crippa A, Discacciati A, Larsson SC, Wolk A, Orsini N. Coffee consumption and mortality from all causes, cardiovascular disease, and cancer: a dose-response meta-analysis. *Am J Epidemiol* 2014;**180:**763–75.
10. Schoenfeld JD, Ioannidis JP. Is everything we eat associated with cancer? A systematic cookbook review. *Am J Clin Nutr* 2013;**97:** 127–34.
11. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* 2015;**65:** 87–108.
12. Koti M, Siu A, Clement I *et al*. A distinct pre-existing inflammatory tumour microenvironment is associated with chemotherapy resistance in high-grade serous epithelial ovarian cancer. *Br J Cancer* 2015;**112:**1215–22.
13. Braem MG, Onland-Moret NC, Schouten LJ *et al*. Coffee and tea consumption and the risk of ovarian cancer: a prospective cohort study and updated meta-analysis. *Am J Clin Nutr* 2012;**95:** 1172–81.
14. Steevens J, Schouten LJ, Verhage BA, Goldbohm RA, van den Brandt PA. Tea and coffee drinking and ovarian cancer risk: results from the Netherlands Cohort Study and a meta-analysis. *Br J Cancer* 2007;**97:**1291–94.
15. Song YJ, Kristal AR, Wicklund KG, Cushing-Haugen KL, Rossing MA. Coffee, tea, colas, and risk of epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2008;**17:** 712–16.

16. Lueth NA, Anderson KE, Harnack LJ, Fulkerson JA, Robien K. Coffee and caffeine intake and the risk of ovarian cancer: the Iowa Women's Health Study. *Cancer Causes Control* 2008;**19:** 1365–72.
17. Silvera SA, Jain M, Howe GR, Miller AB, Rohan TE. Intake of coffee and tea and risk of ovarian cancer: a prospective cohort study. *Nutr Cancer* 2007;**58:**22–27.
18. Goodman MT, Tung KH, McDuffie K, Wilkens LR, Donlon TA. Association of caffeine intake and CYP1A2 genotype with ovarian cancer. *Nutr Cancer* 2003;**46:**23–29.
19. Gosvig CF, Kjaer SK, Blaakaer J, Hogdall E, Hogdall C, Jensen A. Coffee, tea, and caffeine consumption and risk of epithelial ovarian cancer and borderline ovarian tumors: Results from a Danish case-control study. *Acta Oncol* 2015;**54:**1144–51.
20. Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998;**316:** 140–44.
21. 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.
22. Vink JM, Staphorsius AS, Boomsma DI. A genetic analysis of coffee consumption in a sample of Dutch twins. *Twin Res Hum Genet* 2009;**12:**127–31.
23. Laitala VS, Kaprio J, Silventoinen K. Genetics of coffee consumption and its stability. *Addiction* 2008;**103:**2054–61.
24. Luciano M, Kirk KM, Heath AC, Martin NG. The genetics of tea and coffee drinking and preference for source of caffeine in a large community sample of Australian twins. *Addiction* 2005; **100:**1510–17.
25. Phelan CM, Kuchenbaecker KB, Tyrer JP *et al*. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet* 2017;**49:**69–78.
26. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012;**44:** 955–59.
27. O'Connell J, Gurdasani D, Delaneau O *et al*. A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genet* 2014;**10:**e1004234.
28. Josse AR, Da Costa LA, Campos H, El-Sohemy A. Associations between polymorphisms in the AHR and CYP1A1-CYP1A2 gene regions and habitual caffeine consumption. *Am J Clin Nutr* 2012;**96:**665–71.
29. Amin N, Byrne E, Johnson J *et al*. Genome-wide association analysis of coffee drinking suggests association with CYP1A1/ CYP1A2 and NRCAM. *Mol Psychiatry* 2012;**17:**1116–29.
30. 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.
31. Coffee, Caffeine Genetics Consortium; Cornelis MC, Byrne EM, Esko T *et al*. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol Psychiatry* 2015;**20:**647–56.
32. 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.
33. 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.
34. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG; EPIC-Interact Consortium. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* 2015;**30:**543–52.
35. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007;**39:**906–13.
36. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;**37:**658–65.
37. He Z, Ma WY, Hashimoto T, Bode AM, Yang CS, Dong Z. Induction of apoptosis by caffeine is mediated by the p53, Bax, and caspase 3 pathways. *Cancer Res* 2003;**63:**4396–401.
38. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;**474:**609–15.
39. Lensu S, Tuomisto JT, Tuomisto J, Viluksela M, Niittynen M, Pohjanvirta R. Immediate and highly sensitive aversion response to a novel food item linked to AH receptor stimulation. *Toxicol Lett* 2011;**203:**252–57.
40. Kurman RJ, Shih Ie M. Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. *Int J Gynecol Pathol* 2008;**27:**151–60.
41. Salehi F, Dunfield L, Phillips KP, Krewski D, Vanderhyden BC. Risk factors for ovarian cancer: an overview with emphasis on hormonal factors. *J Toxicol Environ Health B Crit Rev* 2008;**11:** 301–21.