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

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>International Journal of Epidemiology</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>IJE-2017-05-0516</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>04-May-2017</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Ong, Jue Sheng; QIMR Berghofer Medical Research Institute, Statistical Genetics  
Hwang, Liang-Dar; QIMR Berghofer Medical Research Institute, Genetics and Computational Biology Department  
Cuellar-Partida, Gabriel; University of Queensland Diamantina Institute for Cancer Immunology and Metabolic Medicine, Genomic Medicine; QIMR Berghofer Medical Research Institute,  
Bryne, Enda; University of Queensland Institute for Molecular Bioscience  
Fasching, Peter; University of California at Los Angeles, David Geffen School of Medicine, Department of Medicine, Division of Hematology and Oncology; ; University Hospital Erlangen, Department of Gynecology and Obstetrics, Friedrich-Alexander-University Erlangen-Nuremberg Comprehensive Cancer Center Erlangen Nuremberg,  
Hein, Alexander; Universitatsklinikum Erlangen, Department of Gynecology and Obstetrics  
Burghaus, Stefanie; Universitatsklinikum Erlangen, Department of Gynecology and Obstetrics  
Beckmann, Matthias; University Hospital Erlangen, Department of Gynecology and Obstetrics, Friedrich-Alexander-University Erlangen-Nuremberg Comprehensive Cancer Center Erlangen Nuremberg,  
Lambrechts, Diether; Vesalius Research Center, VIB,  
Van Nieuwenhuysen, Els; Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven,  
Vergote, Ignace; Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven,  
Vanderstichele, Adriaan; Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven  
Swerdlow, Anthony; Institute of Cancer Research, Section of Epidemiology  
Jones, Michael; Institute of Cancer Research, Section of Epidemiology  
Orr, Nicholas; Institute of Cancer Research  
Schoemaker, Minouk; Institute of Cancer Research, Section of Epidemiology  
Velez Edwards, Digna; Vanderbilt Epidemiology Center  
Brenton, James; Cancer Research UK Cambridge Research Institute  
Benitez, Javier; Centro Nacional de Investigaciones Oncologicas, Human |
<table>
<thead>
<tr>
<th>Names</th>
<th>Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Genetics group</td>
<td>García, María J.; Centro Nacional de Investigaciones Oncologicas, Human</td>
</tr>
<tr>
<td>Cancer Genetics group</td>
<td>Rodríguez-Antona, Cristina; Centro Nacional de Investigaciones Oncologicas, Hereditary Endocrine Cancer group</td>
</tr>
<tr>
<td>Oncology Science, Program in Epidemiology, Division of Public Health</td>
<td>Rossing, Mary Anne; Program in Epidemiology, Division of Public Health</td>
</tr>
<tr>
<td>Fortner, Renée T.; Deutsches Krebsforschungszentrum, Department of</td>
<td>Cancer Epidemiology</td>
</tr>
<tr>
<td>Cancer Epidemiology</td>
<td>Riboli, Elio; Imperial College</td>
</tr>
<tr>
<td>Chang-Claude, Jenny; German Cancer Research Center (DKFZ), Division of</td>
<td>Cancer Epidemiology</td>
</tr>
<tr>
<td>Cancer Epidemiology</td>
<td>Eilber, Ursula; German Cancer Research Center (DKFZ), Division of Cancer</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>Rudolph, Anja; Deutsches Krebsforschungszentrum, Division of Cancer Epidemiology</td>
</tr>
<tr>
<td>Wang-Gohrke, Shan; University of Ulm</td>
<td>Yannoukakos, Drakoulis; National Centre for Scientific Research-Demokritos</td>
</tr>
<tr>
<td>Goodman, Marc; Cancer Prevention and Control, Samuel Oschin</td>
<td>Comprehensive Cancer Institute, Cedars-Sinai Medical Center</td>
</tr>
<tr>
<td>Antonenkova, Natalia; N.N. Alexandrov National Cancer Centre of Belarus</td>
<td>Dörk, Thilo; Clinics of Obstetrics and Gynaecology, Hannover Medical School,</td>
</tr>
<tr>
<td>Dürst, Matthias; Jena-University Hospital-Friedrich Schiller University, Department of Gynecology</td>
<td>Hillemanns, Peter; Clinics of Obstetrics and Gynaecology, Hannover Medical School,</td>
</tr>
<tr>
<td>Butzow, Ralf; Department of Pathology, University of Helsinki and Helsinki University Hospital</td>
<td>Leminen, Arto; Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital,</td>
</tr>
<tr>
<td>Nevanlinna, Heli; Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital</td>
<td>Pelttari, Liisa; Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital,</td>
</tr>
<tr>
<td>Edwards, Robert; Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine,</td>
<td>Ovarian Cancer Center of Excellence, Women's Cancer Research Program, Magee-Women's Research Institute and University of Pittsburgh Cancer Institute,</td>
</tr>
<tr>
<td>Kelley, Joseph; Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine,</td>
<td>Modugno, Francesmary; Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine,</td>
</tr>
<tr>
<td>Moysich, Kirsten; Roswell Park Cancer Institute, University of Texas School of Public Health,</td>
<td>Ness, Roberta; Roswell Park Cancer Institute Department of Cancer Prevention and Population Sciences,</td>
</tr>
<tr>
<td>Kjær, Susanne; Inst. of Cancer epidemiology, Virus, Hormones and Cancer Jensen, Allan; Danish Cancer Society, Institute of Cancer Epidemiology</td>
<td>Giles, Graham; Cancer Council Victoria, Cancer Epidemiology Centre,</td>
</tr>
<tr>
<td>Bruinsma, Fiona; Cancer Epidemiology Centre, Cancer Council Victoria,</td>
<td>Hildebrandt, Michelle; The University of Texas MD Anderson Cancer Center,</td>
</tr>
<tr>
<td>Epidemiology</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Liang, Dong; College of Pharmacy and Health Sciences, Texas Southern University,</td>
<td></td>
</tr>
<tr>
<td>Wu, X; The University of Texas, Department of Epidemiology</td>
<td></td>
</tr>
<tr>
<td>LeMarchand, Loic; University of Hawaii Cancer Center,</td>
<td></td>
</tr>
<tr>
<td>Setiawan, Veronica; University of Southern California, Preventive Medicine</td>
<td></td>
</tr>
<tr>
<td>Permut, Jennifer B.; H Lee Moffitt Cancer Center and Research Institute, Department of Cancer Epidemiology</td>
<td></td>
</tr>
<tr>
<td>Bisogna, Maria; Gynecology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center,</td>
<td></td>
</tr>
<tr>
<td>Dao, Fanny; Memorial Sloan-Kettering Cancer Center, Department of Surgery</td>
<td></td>
</tr>
<tr>
<td>Levine, Douglas A.; Memorial Sloan-Kettering Cancer Center, Department of Surgery</td>
<td></td>
</tr>
<tr>
<td>Cramer, Daniel; Harvard University,</td>
<td></td>
</tr>
<tr>
<td>Terry, Kathryn; Brigham and Women's Hospital, Obstetrics and Gynecology Epidemiology Center</td>
<td></td>
</tr>
<tr>
<td>Tworoger, Shelley; Harvard Medical School and Brigham and Women's Hospital, Channing Division of Network Medicine; Harvard School of Public Health, Department of Epidemiology</td>
<td></td>
</tr>
<tr>
<td>Stampfer, Meir; Harvard School of Public Health, Epidemiology</td>
<td></td>
</tr>
<tr>
<td>Willett, W; Harvard University T H Chan School of Public Health</td>
<td></td>
</tr>
<tr>
<td>Winham, Stacey; Department of Health Science Research, Division of Biomedical Statistics and Informatics, Mayo Clinic,</td>
<td></td>
</tr>
<tr>
<td>Bjorge, Line; Haukeland Universitetssjukehus, Department of Gynecology and Obstetrics</td>
<td></td>
</tr>
<tr>
<td>Kopperud, Reidun K.; Department of Gynecology and Obstetrics, Haukeland University Hospital</td>
<td></td>
</tr>
<tr>
<td>Bischof, Katharina; Department of Gynecology and Obstetrics, Haukeland University Hospital</td>
<td></td>
</tr>
<tr>
<td>Kiemenej, Lambertus; Radboud University Medical Center, Radboud Institute for Health Sciences,</td>
<td></td>
</tr>
<tr>
<td>Massuger, Leon; Radboud University Medical Center, Radboud Institute for Molecular Life Sciences, Department of Gynaecology,</td>
<td></td>
</tr>
<tr>
<td>Pejovic, Tanja; Department of Obstetrics &amp; Gynecology, Oregon Health &amp; Science University, ; Knight Cancer Institute, Oregon Health &amp; Science University,</td>
<td></td>
</tr>
<tr>
<td>Brooks-Wilson, Angela; Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, ; Department of Biomedical Physiology and Kinesiology, Simon Fraser University,</td>
<td></td>
</tr>
<tr>
<td>Olson, Sara; Memorial Sloan-Kettering Cancer Center,</td>
<td></td>
</tr>
<tr>
<td>McGuire, Valerie; Stanford University,</td>
<td></td>
</tr>
<tr>
<td>Rothstein, Joseph; Department of Health Research and Policy - Epidemiology, Stanford University School of Medicine,</td>
<td></td>
</tr>
<tr>
<td>Sieh, Weiva; Icahn School of Medicine at Mount Sinai, Department of Genetics and Genomic Sciences</td>
<td></td>
</tr>
<tr>
<td>Whittemore, A; &lt;none&gt;,</td>
<td></td>
</tr>
<tr>
<td>Cook, L; University of New Mexico,</td>
<td></td>
</tr>
<tr>
<td>Le, Nhu; BC Cancer Agency, Cancer Control Research</td>
<td></td>
</tr>
<tr>
<td>Gilks, C. Blake ; University of British Columbia, Pathology and Laboratory Medicine</td>
<td></td>
</tr>
<tr>
<td>Gronwald, Jacek; International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University,</td>
<td></td>
</tr>
<tr>
<td>Jakubowska, Anna; International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University</td>
<td></td>
</tr>
<tr>
<td>Lubiński, Jan; International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University,</td>
<td></td>
</tr>
<tr>
<td>Kluż, Tomasz; Universytet Rzeszowski, Institute of Midwifery and Emergency Medicine</td>
<td></td>
</tr>
</tbody>
</table>
| Wentzensen, Nicolas; Division of Cancer Epidemiology and Genetics, National Cancer Institute,
Risch, HA; <none>,
Goode, Ellen; Mayo Clinic College of Medicine,
Schildkraut, Joellen; Duke Cancer Institute, Cancer Prevention, Detection &
Control Research Program
Pearce, Celeste Leigh; Department of Preventive Medicine, Keck School of
Medicine, University of Southern California Norris Comprehensive Cancer
Center,
Berchuck, Andrew; Duke University,
Pharoah, Paul; Cancer Research UK, Department of Oncology, Cambridge
University, Cambridge, United Kingdom,
Martin, Nicholas; QIMR Berghofer Department of Genetics and
Computational Biology
Chenevix-Trench, Georgia; Cancer Genetics Group, QIMR Berghofer
Medical Research Institute,
Quinn, Michael; QIMR Berghofer Department of Genetics and
Computational Biology
Cornelis, Marilyn; Northwestern University Feinberg School of Medicine,
Preventive Medicine
Gharahkhani, Puya; QIMR Berghofer Medical Research Institute, Statistical
Genetics
Webb, Penelope M.; QIMR Berghofer Medical Research Institute, Population
Health
MacGregor, Stuart; QIMR Berghofer Medical Research Institute, Statistical
Genetics

Key Words: Mendelian randomization, coffee, ovarian cancer, causality, instrumental
variable
Assessment of moderate coffee consumption and risk of epithelial ovarian cancer: a Mendelian randomization study

Jue-Sheng Ong, Liang-Dar Hwang, Gabriel Cuellar-Partida, (more OCAC authors here), Nicholas G. Martin, Georgia Chenevix-Trench, Michael C. J. Quinn, Marilyn C. Cornelis, Puya Gharahkhani, Penelope M. Webb, Stuart MacGregor

[Please refer to the attached author list document for the complete list of authors and affiliations.]

Number of Words (main text): 3 250 words

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 SNPs associated with (i) coffee and (ii) caffeine consumption to perform Mendelian randomisation 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: 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.

Conclusion:
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 Message**

- 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 to observational finding 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 1,000 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 stimulating intracellular antioxidant defence (1-3). Observational studies have investigated coffee and caffeine intake in relation to type 2 diabetes (4, 5), depression (6), 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). The high-grade serous histology is the most
common EOC subtype (12). Whilst many individual studies have found conflicting directions of
association with coffee consumption and EOC risk, 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 – 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 customised Infinium OncoArray-500K 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 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 makeup, using principal component analysis. These values sum up to 1 and are used to categorise 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=1,732) 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 summarised in Supplementary Table 1.

Genetic variants for the MR analyses were identified through an extensive review of published GWAS findings for coffee, tea and/or caffeine consumption (28-33). SNPs associated with coffee consumption (measured as cups/day) that 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 with coffee intake (in cups per day) onto 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-disequilibrium
($r^2=0.8$), see discussion for more detail. In order to ensure that our SNPs of interest are 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 $\tau^2_{SNP} = 2p(1-p)\beta^2/\sigma^2$ where $\tau^2_{SNP}$ refers to the variance
explained by the SNP, $p$ refers to the MAF of the SNP, $\beta$ is the measured magnitude of association
per effect allele and $\sigma^2$ is the coffee trait variance. The variance explained by our SNP instruments
can hence be obtained by linearly summing up $\tau^2_{SNP}$ across each independent SNP instrument. We
subsequently tested each SNP against potential confounders using publicly available GWAS datasets.
The possible confounders tested were age at menarche, measures of glycaemia, education
attainment, BMI, waist-hip ratio, body fat and smoking behaviour (supplementary table 3).

Causal Effect estimation

To perform MR we utilised 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 OR (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 4 independent SNPs (rs1481012, rs6968554, rs2470893, rs17685) as proxies for genetically determined coffee consumption behaviour (31). For the analysis on caffeine, we used 2 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). Each of the SNPs thus satisfies the strong MR instrument criterion (F>>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, supplementary material) so it was excluded from our analyses. After removing BDNF, the 4 coffee SNPs combined explain about ~1.2% of the variation in coffee intake (31), whereas the 2 SNPs combined for our MR caffeine study explain about ~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). In brief, no associations were
found above chance level and we conclude that the assumptions of no-pleiotropy is 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. 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% confidence interval, 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 80mg increase (the approximate caffeine content
in a conventional cup of coffee). The COR for consuming an additional 80mg 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. 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 9 genetic (ancestral) principal components as
covariates. In a sensitivity analysis using participants with confounder data available (n~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). The genomic control lambda was 1.076
($\lambda_{A1000} = 1.007$, 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 Figure 1 and 2).

**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
nor on 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. 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 common criticism of observational studies is
inconsistency in the definition of categorised 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 2-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 was 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 SNP 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 alone was of particular interest. 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 summarised 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 2-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 remain a robust genetic predisposition to lifetime coffee intake behaviour.

For our MR to infer about causality, several MR assumptions have to be met. Firstly, the instruments (SNPs) used here were robustly associated (with F>>10) to coffee and caffeine intake respectively. Secondly, 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, a 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 metabolises caffeine in the liver, while *CYP1A1* encodes protein that metabolises 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 responsible 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. Firstly, our study was performed using only European ancestry women and our findings may not generalize to other populations. Even though our SNPs are strong instruments, taken from large GWAS studies, all our SNPs combined only account for a relatively small proportion of variation (~1.2%) in coffee consumption (cups per day). The variance explained directly influences the precision of our estimates and with a larger proportion of variance explained (or a larger sample size), the confidence interval on our causal odds ratio would be smaller. However, for both all histologies and high grade serous histology alone, the confidence intervals from our study are small enough to exclude all but small (<10%) true effects of coffee consumption on cancer risk. Our study does not have good power for the individual subtypes other than high grade serous and we 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 (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 to 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; which 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 based 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.

**Funding**

This work was supported by The COGS project which is funded through a European Commission’s Seventh Framework Programme Grant (agreement number 223175 HEALTH F2 2009-223175); the Genetic Associations and Mechanisms in Oncology (GAME-ON): A 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, 2I10200); 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; 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 POSC 028 14, 2 POSA 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 and 400281); 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: JS and GC thank the University of Queensland and QIMR Berghofer Medical Research Institute for scholarship support. YL was supported by the NHMRC Early Career Fellowship. GCT, REN and PMW are supported by the National Health and Medical Research Council. SM acknowledges funding support from an Australian Research Council Future Fellowship.

Acknowledgements

We acknowledge with appreciation all the individuals who participated in the OCAC studies, and the many hospital directors and staff, gynaecologists, general practitioners and pathology services 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, 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 (NJ); 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).

Reference


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

<table>
<thead>
<tr>
<th>Nature/Subtype</th>
<th>European Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasive</td>
<td>17,779</td>
</tr>
<tr>
<td>All serous(^x)</td>
<td>11,213</td>
</tr>
<tr>
<td>Endometrioid</td>
<td>2,199</td>
</tr>
<tr>
<td>Clearcell</td>
<td>1,121</td>
</tr>
<tr>
<td>Mucinous</td>
<td>1,125</td>
</tr>
<tr>
<td>All mucinous(^x)</td>
<td>2,023</td>
</tr>
<tr>
<td>High-grade serous</td>
<td>7,488</td>
</tr>
<tr>
<td>Low-grade serous</td>
<td>880</td>
</tr>
<tr>
<td>All EOC cases(^x)</td>
<td>20,683</td>
</tr>
</tbody>
</table>

\(^x\) Including unclassified and unknown serous/mucinous ovarian tumours.

Note: A complete breakdown of the EOC cases by each participating study is provided in Supplementary Material.
Figure 1A. Instrumental variable estimate for coffee consumption on EOC susceptibility.

237x159mm (72 x 72 DPI)
Figure 1B. Instrumental variable for coffee consumption on high-grade serous EOC susceptibility.

242x159mm (72 x 72 DPI)
Figure 2A. Instrumental variable estimate for caffeine intake on EOC susceptibility.

Estimated effect of additional 80mg caffeine intake per day on all EOC risk

<table>
<thead>
<tr>
<th>SNP Instrument</th>
<th>EA/NEA</th>
<th>Causal OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs6968865</td>
<td>T/A</td>
<td>1.11(0.95,1.29)</td>
</tr>
<tr>
<td>rs2472297</td>
<td>T/C</td>
<td>0.96(0.84,1.08)</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>1.01(0.92,1.11)</td>
</tr>
</tbody>
</table>

242x159mm (72 x 72 DPI)
**Figure 2B. Instrumental variable estimate for caffeine intake on high-grade serous EOC susceptibility.**

<table>
<thead>
<tr>
<th>SNP Instrument</th>
<th>EA/NEA</th>
<th>Causal OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs6968865</td>
<td>T/A</td>
<td>0.83 (0.68, 0.92)</td>
</tr>
<tr>
<td>rs2472297</td>
<td>T/C</td>
<td>1.10 (0.93, 1.31)</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td></td>
<td><strong>0.99 (0.86, 1.13)</strong></td>
</tr>
</tbody>
</table>

242x159mm (72 x 72 DPI)
Comparison to previous studies on association of coffee consumption with EOC susceptibility

<table>
<thead>
<tr>
<th>Studies</th>
<th>Cases</th>
<th>OR (95% CI.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steevens et al.</td>
<td>280</td>
<td>1.04 (0.9, 1.12)</td>
</tr>
<tr>
<td>Braem et al. (2012)</td>
<td>1,244</td>
<td>0.98 (0.82, 1.18)*</td>
</tr>
<tr>
<td>Gosvig et al. (2015)</td>
<td>267</td>
<td>1.15 (0.82, 1.61)</td>
</tr>
<tr>
<td>Song et al. (2008)</td>
<td>781</td>
<td>1.00 (0.86, 1.17)*</td>
</tr>
<tr>
<td>MR estimate</td>
<td>20.683</td>
<td>0.92 (0.79, 1.06)</td>
</tr>
</tbody>
</table>

*Averaged odds ratio estimated via Equation 1 in Supplementary Material

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

242x159mm (72 x 72 DPI)