

# Appraisal of the causal effect of *Chlamydia trachomatis* infection on epithelial ovarian cancer risk: a two-sample Mendelian randomisation study

Sarah L. Perrott<sup>\*1,2</sup>, Siddhartha P. Kar<sup>\*1</sup>

1. Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom
2. Early Cancer Institute, University of Cambridge, Cambridge, United Kingdom

**\* For correspondence:**

Sarah Perrott – [sp2176@cam.ac.uk](mailto:sp2176@cam.ac.uk)

Siddhartha Kar – [sk718@cam.ac.uk](mailto:sk718@cam.ac.uk)

## Abstract

**Background:** History of *Chlamydia trachomatis* infection has previously been associated with epithelial ovarian cancer (EOC) in observational studies. We conducted a two-sample univariable Mendelian randomisation (MR) study to examine whether genetically predicted seropositivity to the *C. trachomatis* major outer membrane protein (momp) D is causally associated with EOC.

**Methods:** MR analyses employed genetic associations derived from UK Biobank as proxies for momp D seropositivity in 25 509 EOC cases and 40 941 controls that participated in the Ovarian Cancer Association Consortium. Findings were replicated using a GWAS meta-analysis of global biobanks including the UK Biobank, FinnGen and BioBank Japan.

**Results:** Genetically predicted momp D seropositivity was associated with overall and high-grade serous EOC risk in inverse-variance weighted (IVW) and MR-Egger univariable MR analysis (odds ratio (OR) 1.06; 95% confidence interval (CI) 1.02—1.10, and OR 1.08; 95%CI 1.01—1.16, respectively). Replication yielded similar results for overall EOC (OR 1.11; 95%CI 1.01—1.22).

**Conclusion:** This MR study supports a causative link between *C. trachomatis* infection and overall and high-grade serous EOC.

## Introduction

*Chlamydia trachomatis* is the most prevalent sexually transmitted infection (STI), affecting 4% of women worldwide.<sup>1</sup> *C. trachomatis* infections are asymptomatic in up to 70%, and consequently often go undiagnosed and untreated, resulting in onward transmission and sequelae including pelvic inflammatory disease (PID) and infertility in women.<sup>2</sup> Both *C. trachomatis* and PID have been associated with an increased risk of ovarian cancer, the most lethal gynaecological cancer.

3-5

Ovarian cancer is a global problem; although central Europe has the highest incidence, rates in low- and middle-income countries have increased by approximately 70% since 2007.<sup>6</sup> Traditional risk factors include increased age, inherited germline mutations, family history, and infertility treatments. Parity, lactation and oral contraceptives are recognised as protective factors.<sup>7</sup> Ovarian tumours originate from either epithelial cells (90%), stromal cells (5-6%), or germ cells (2-3%). Epithelial ovarian cancer (EOC) is a heterogeneous group and can be further subdivided by histotype into high grade serous; low grade serous; mucinous; clear cell; and endometrioid EOC. High-grade serous EOC is not only the most common, accounting for 75% of all ovarian cancers, but also the most aggressive and deadly.<sup>8</sup> The conventional EOC risk factors, once accepted for all subtypes, are strongly associated with non-serous rather than serous EOCs.<sup>7</sup> Recent molecular and epidemiological evidence suggest that high-grade serous EOCs do not even originate from the ovary, but precancerous lesions of the fimbriated end of the fallopian tubes called serous tubal intraepithelial carcinomas (STICs).<sup>9</sup> STICs have previously been associated with pelvic inflammatory conditions including endometriosis and PID.<sup>10</sup>

Fallopian tube inflammation and damage are often secondary to PID, prompting consideration of the potential involvement of infectious agents in both STICs and EOC. The relationship between pathogenic organisms and cancer is well-established; helicobacter pylori has been long

associated with gastric cancer and human papilloma virus (HPV) with cervical cancer.<sup>11,12</sup> The associations between PID and EOC, and several infectious agents including *C. trachomatis*, *Neisseria gonorrhoeae*, *Mycoplasma genitalium*, HPV, and herpes simplex virus 2 (HSV-2) have been previously examined in observational studies, namely case-control or prospective cohort.<sup>3-5,13-18</sup> When considered overall, evidence supporting the role of *C. trachomatis* in tubal inflammation, damage, and potentially ovarian carcinogenesis, is more compelling compared to that of other STIs.<sup>7,19</sup> Despite this, a few studies have failed to demonstrate any association, or have suggested HSV-2 or *M. genitalium* may have a role instead.<sup>5,14,16-18</sup> The existing evidence base is deficient due to challenges in study design, influenced by residual confounding factors and relatively small study populations. *C. trachomatis* is a small obligate intracellular bacterium that has developed mechanisms to promote bacterial survival, which potentially lead to host cell genomic instability, cell cycle disruption, and inhibition of apoptosis – all of which are hypothetical hallmarks of a carcinogenic pathogen.<sup>10</sup> However, it has not been possible to determine whether observed associations are causal, or whether other external factors influence future EOC risk.

Mendelian randomisation (MR) is an epidemiological strategy aimed at removing potential biases which exist within conventional observational studies.<sup>20</sup> This method uses single nucleotide polymorphisms (SNPs) as instrumental variables, enabling a potential causal relationship between an exposure and outcome to be determined. To our knowledge, MR has never been used to explore this association. We therefore used a two-sample univariable MR approach to investigate the causal relationship between seropositivity to the *C. trachomatis* major outer membrane protein (momp) D antibody and EOC. By elucidating the causal relationship between *C. trachomatis* seropositivity and EOC, this study aims to contribute to our understanding of EOC aetiology – particularly with regard to serous EOC, whereby any modifiable risk factors yet remain undetermined.<sup>7</sup>

## Methods

### Study design

The current study has been reported according to the STROBE-MR guidelines (Reporting Standards Document). This study employed a two-sample univariable MR design to investigate the causal relationship between *C. trachomatis* seropositivity and EOC. To establish causality, MR relies on three key assumptions as illustrated in Figure 1.

The MR approach utilises SNPs as instrumental variables for the exposure, allowing the lifelong and causal effects of the exposure on the outcome to be examined. Since SNPs are randomly assigned at conception, an MR study theoretically mirrors the design of a randomised controlled trial, minimising effects of confounding through randomisation. Compared to conventional epidemiological studies, this makes MR studies less susceptible to reverse causation and unmeasured confounding, therefore providing stronger evidence of causal inference. This study uses the two-sample MR approach; this means summary statistics were obtained for the exposure and outcome from two independent genome-wide association study (GWAS) samples. In all analyses, we used univariable MR approach to evaluate if genetic liability to *C. trachomatis* phenotypes were causally associated with EOC. We utilised previously validated sets of single nucleotide polymorphisms (SNPs) associated at genome-wide significance ( $p < 5 \times 10^{-8}$ ) with *C. trachomatis* momp D seropositivity and EOC (overall, and by histotype). Additionally, we conducted several supplementary analyses to assess the robustness of our MR results, including sensitivity analyses and assessments for potential horizontal pleiotropy (i.e. ensuring SNPs influence the outcome only through the allocated exposure, and not through any other pathway). Publicly available summary statistics were retrieved from GWAS conducted predominantly in Europeans. All included GWASs and biobanks were approved by corresponding ethics committees and participants provided written informed consent.

### *C. trachomatis* serology (exposure) data

The exposure evaluated was *C. trachomatis* seropositivity level based on the serum momp D antibody. Momp D was chosen over other serological markers as there were multiple genome-wide significant ( $p < 5 \times 10^{-8}$ ) SNPs associated with it, allowing for the use of the powerful inverse-variance weighted MR method and for robust sensitivity analyses to evaluate the exclusion restriction assumption of MR (Figure 1). We used genome-wide significant SNPs associated with serum momp D antibody level in the UK Biobank from a GWAS by Butler-Laporte *et al.*<sup>21</sup> The UK Biobank is a prospective cohort study that recruited approximately 500,000 individuals of predominantly European ancestry, aged 40 to 69 years, at 22 assessment centres across England, Scotland, and Wales during 2006–2010. Of these, a subset of 9,724 individuals provided serum for serological evaluation of 20 common infectious microorganisms based on fluorescent bead multiplex technology. Butler-Laporte *et al* undertook GWAS for infections with a seroprevalence of over 15% in this subset of the UK Biobank, performing genetic association analyses on log-transformed mean fluorescence intensity (MFI) levels (a standardised measure of antibody level) in European ancestry individuals. A total of 1,371 individuals were seropositive for momp D (defined as MFI > 100) and included in the momp D GWAS. Ten independent  $p < 5 \times 10^{-8}$  SNPs from the momp D GWAS (linkage disequilibrium  $r^2 = 0.001$  and clumping distance = 10,000 kilobase pairs used to defined independence) were employed collectively as a genetic instrument for momp D seropositivity level.

### *Epithelial ovarian cancer risk (outcome) data*

To evaluate the association between genetically predicted serum momp D level and risk of overall and histological subtype-specific EOC, we obtained summary genetic association data from a GWAS meta-analysis by Phelan *et al* and the Ovarian Cancer Association Consortium (OCAC).<sup>22</sup> This GWAS meta-analysis included 25,509 EOC cases and 40,941 controls. Associations were reported for EOC overall (n = 25,509 cases), and for seven histotypes: high-

grade serous (n = 13,037 cases), low-grade serous (n = 1,012 cases), invasive mucinous (n = 1,417 cases), clear cell (n = 1,366 cases), endometrioid (n = 2,810 cases), serous low malignant potential (n = 1,954 cases), and mucinous low malignant potential (n = 1,149 cases), with each histotype compared against 40,941 controls. The GWAS meta-analysis was based on individuals of genetically inferred European ancestry.<sup>23,24</sup> Given the large sample size of this OCAC GWAS meta-analysis and inclusion of major EOC histotypes, these outcome data were used for the MR primary analyses.

Findings were replicated using a GWAS meta-analysis of two independent biobanks, the UK Biobank and FinnGen, that together included 1,588 EOC cases and 244,932 controls of genetically inferred European ancestry.<sup>25</sup> This study by Sakaue *et al* combined genotype and phenotype data from these global biobanks to identify genetic associations with EOC risk (not subdivided by histotype). Disease endpoint data were derived from electronic medical records and health registries in the two biobanks. This GWAS meta-analysis provided a distinct, independent population that was ideal for replication purposes.

### **Statistical analyses**

Analyses were performed using the *TwoSampleMR* package on R version 4.3.2.<sup>20</sup> This package facilitates semi-automated data formatting, harmonisation, and application of MR methods.

Genetic associations for EOC from the OCAC GWAS were imputed using the *1000 Genomes Project* reference panel.<sup>22</sup> The global biobanks GWAS data were imputed using a combination of reference panels, including the *Haplotype Reference Consortium*, *UK10K*, *1000 Genomes Project*, and a Finnish-specific whole genome sequenced reference panel.<sup>26</sup> Linkage disequilibrium clumping was performed to ensure instruments for the exposure are independent, removing one correlated SNP (rs140290448). Summary statistics for the 10

remaining momp D-associated lead SNPs were available in both EOC data sets (OCAC and global biobanks). These exposure and outcome summary statistics were then subject to allele harmonisation by aligning the beta-coefficients to ensure that the effect allele in the exposure data set is the same effect allele in the outcome data set. For rs140290448, no proxy SNPs in linkage disequilibrium were identified.

For the primary analysis, effect estimates for EOC association by histotype with momp D seropositivity were generated using the inverse-variance weighted median (IVW) method with random-effects using the OCAC GWAS data. The IVW method is widely used in MR; it combines SNP-exposure and SNP-outcome associations using inverse-variance weights, maximising statistical power of the analysis – particularly when multiple instrumental variables are available. It is consistent and easily interpreted, however can be sensitive to outliers and horizontal pleiotropy. Therefore, we conducted four sensitivity analysis (MR-Egger, weighted-mean, simple mode, and weighted mode) to examine the consistency of results and correct for any directional pleiotropy. These alternative MR methods are more robust to pleiotropic variants and were performed for overall EOC and high-grade serous ovarian cancer. We also performed the MR-Egger intercept; this can indicate any potential pleiotropic effects of the genetic instruments.

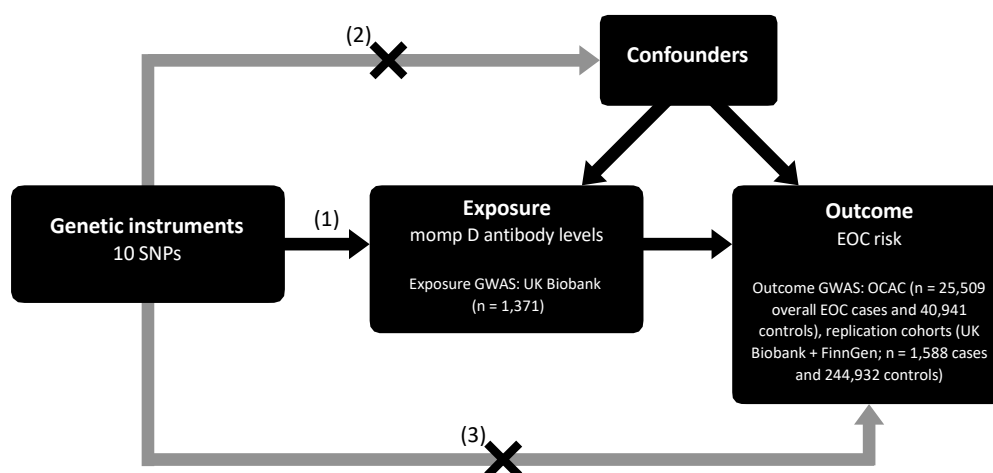
For the secondary (replication) analysis, overall EOC associations with momp D seropositivity was investigated using the global biobanks dataset. Finally, meta-analysis of effect estimates (both by IVW method) from both OCAC and global biobanks were performed, providing an overall effect estimate. The strength of the 10-SNPs was also evaluated using the F-statistic, with  $F < 10$  considered an indicator of a weak instrument.<sup>27</sup> All statistical tests were 2-sided and  $p < 0.05$  was considered statistically significant unless otherwise specified. Odds ratios (OR) and 95% confidence intervals (95% CI) were presented as forest plots using the *ggplot2* R package, version



4.3.2. Finally, this study has been reported as per the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)-MR guideline.<sup>28</sup>

**Figure 1:** Study design and assumptions of two-sample Mendelian randomisation analysis.

- (1) Relevance assumption: Genetic instruments (i.e., SNPs) are associated with the exposure.
- (2) Independence assumption: These SNPs are not associated with known or unknown confounding factors that affect the exposure-outcome relationship.
- (3) Exclusion restriction assumption: The impact of SNPs on the outcome occurs solely through the designated exposure and not through an alternative pathway (i.e., not via horizontal pleiotropy).



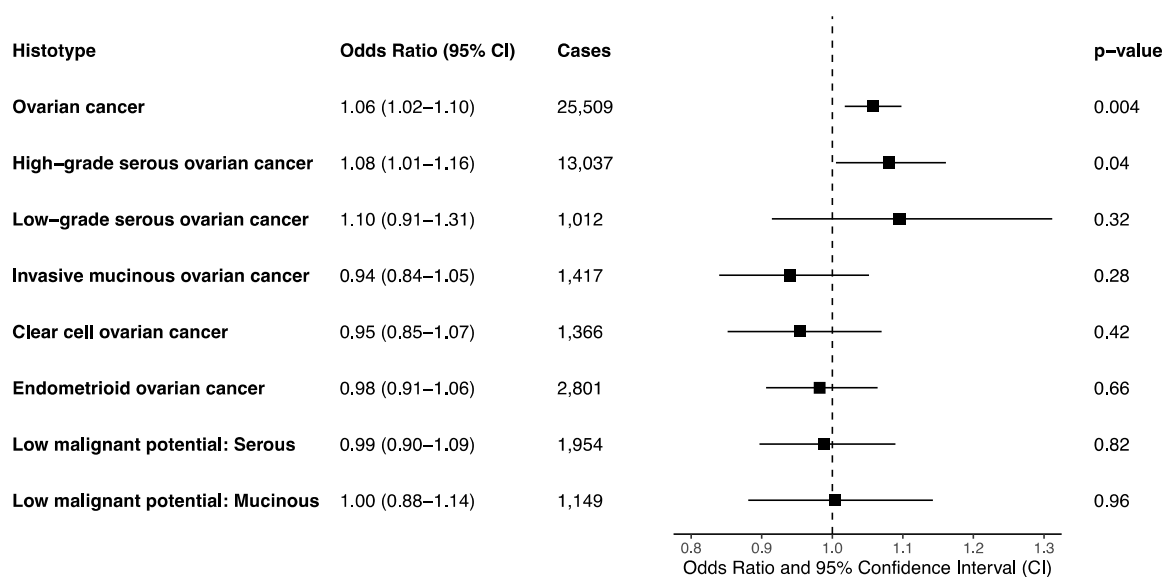
## Results

Associations between individual SNPs and the risk of EOC overall and by histotype in the OCAC and in the combined UK Biobank and FinnGen GWAS meta-analyses are provided in Supplementary Tables 1 and 2, respectively. Each SNP was associated at genome-wide significance ( $p < 5 \times 10^{-8}$ ) with momp D antibody seropositivity level and collectively, they yielded an F-statistic of 112, indicating adequate instrument strength given that F was  $> 10$ .

### *Primary Mendelian randomisation and sensitivity analyses*

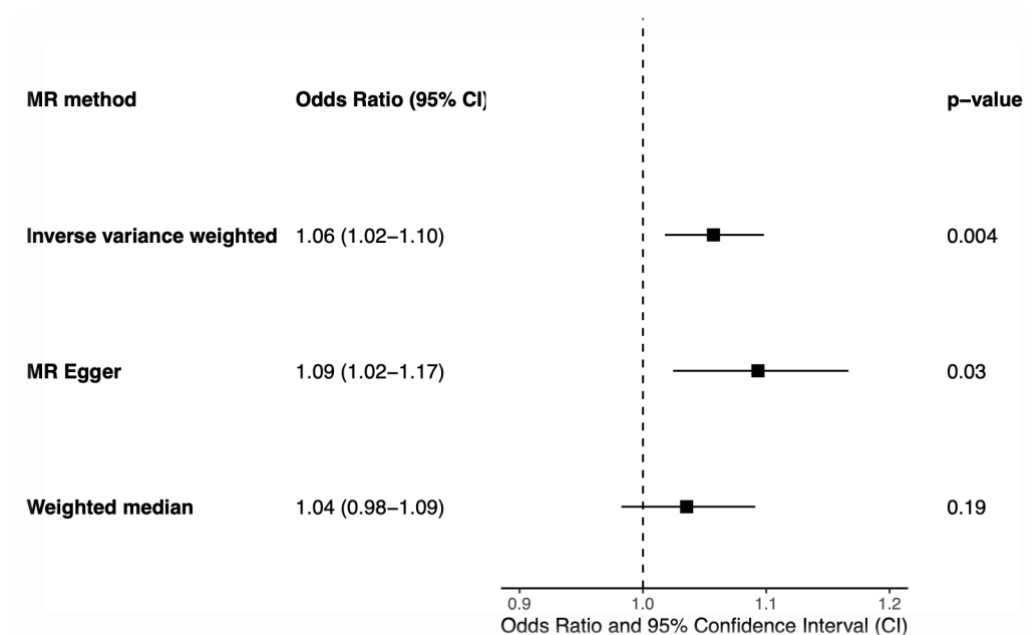
Genetically predicted elevated momp D antibody seropositivity level was associated with increased risk of EOC overall and EOC of the high-grade serous histotype in the primary inverse variance weighted MR analysis (overall EOC OR 1.06 (95% CI 1.02–1.10;  $p = 0.004$ ; Figure 2; all odds ratios and confidence intervals are in units of per standard deviation increase in log-transformed antibody mean fluorescence intensity), and high-grade serous EOC OR 1.08 (95% CI 1.01–1.16;  $p = 0.04$ ; Figure 2)). The point estimate of the effect size for the association with low-grade serous EOC risk was consistent with the high-grade serous EOC association (1.10 (95% CI 0.91–1.31;  $p = 0.32$ ; Figure 2)). There was no evidence of an association with risk of invasive mucinous, clear cell, endometrioid, serous low malignant potential, and mucinous low malignant potential EOC (Figure 2). The inverse variance weighted estimates for risk of overall EOC and high-grade serous EOC were consistent in direction with the results of the MR-Egger and the weighted median methods in sensitivity analyses (Figure 3). Sensitivity analyses results for all histotypes are provided in Supplementary Table 3. The MR-Egger intercept test ( $p = 0.25$  for both overall and high-grade serous EOC) suggested absence of directional horizontal pleiotropy (i.e., it provided statistical evidence for the absence of an effect of the 10 SNPs collectively on overall or high-grade serous EOC risks via a pathway other than momp D serology; Supplementary Table 4).

**Figure 2:** Inverse variance weighted Mendelian randomisation analysis results for the association between genetically predicted momp D antibody seropositivity level and overall and histotype-specific epithelial ovarian cancer risk. All odds ratios and confidence intervals are in units of per standard deviation increase in log-transformed antibody mean fluorescence intensity.

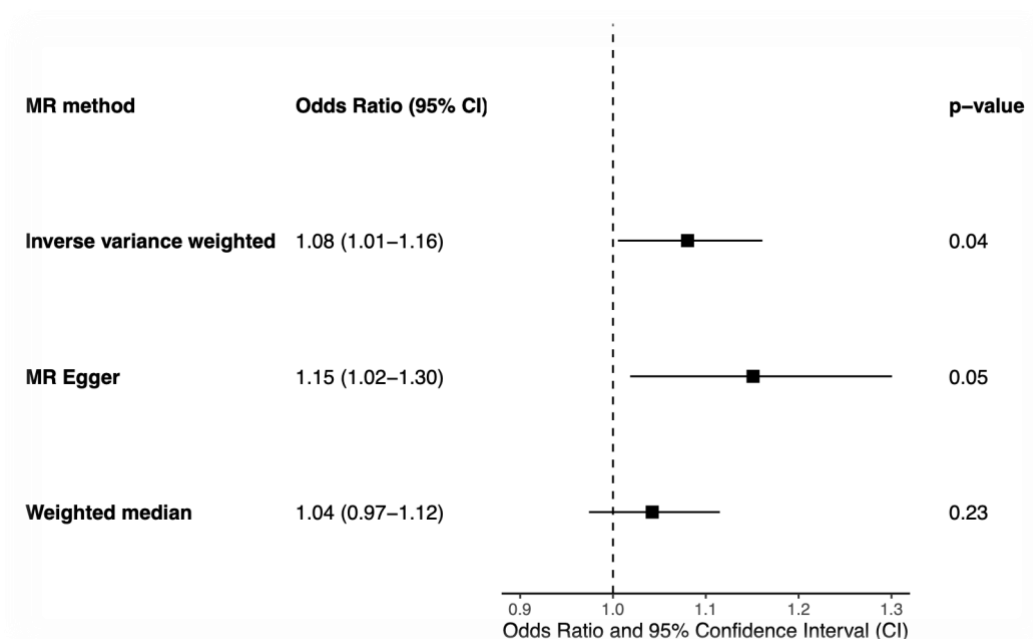


**Figure 3:** Mendelian randomisation (MR) sensitivity analysis results for the association between genetically predicted momp D antibody seropositivity level and (A) overall and (B) high-grade serous epithelial ovarian cancer risk. All odds ratios and confidence intervals are in units of per standard deviation increase in log-transformed antibody mean fluorescence intensity.

(A)



(B)

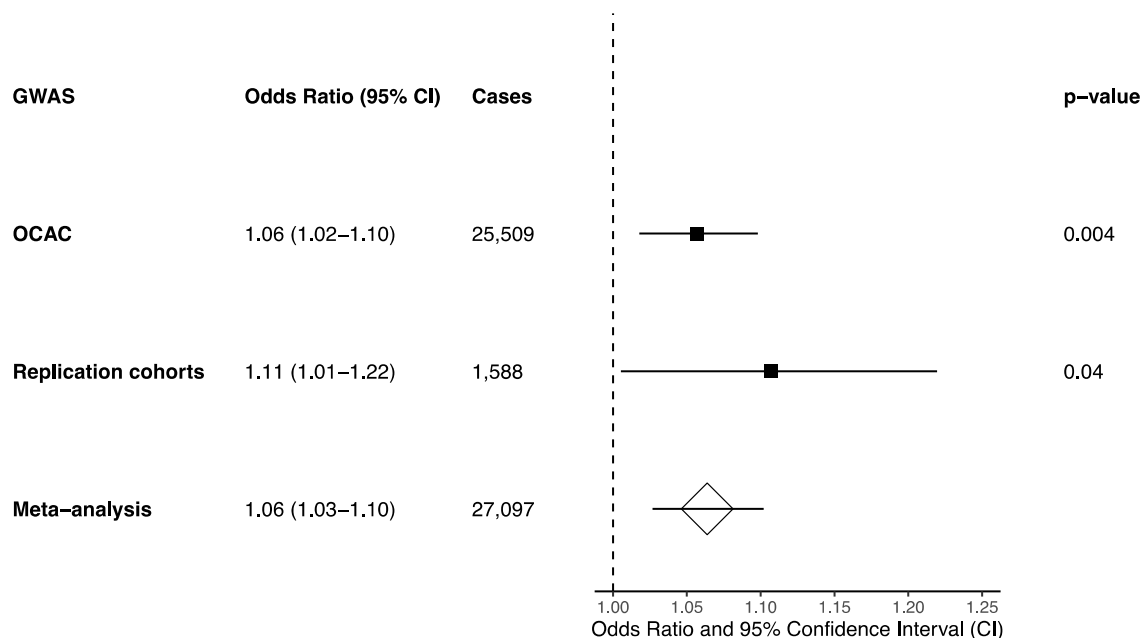


### *Replication Mendelian randomisation and meta-analysis*

We replicated our primary MR result for overall EOC risk using an independent GWAS data set that comprised the UK Biobank and FinnGen biobank combined, observing an association between genetically predicted elevated momp D antibody seropositivity level and increased EOC

risk (OR 1.11 (95% CI 1.01—1.22;  $p=0.04$ ; Figure 4), by the inverse variance weighted method. The point estimate of the effect size in the replication analysis was consistent between the inverse variance weighted method and the MR Egger and weighted median sensitivity analyses (Supplementary Tables 3 and 4). Fixed effect meta-analysis of the MR estimate from OCAC with that from the two biobanks yielded a combined OR of 1.06 (95% CI 1.03—1.10;  $p=0.0006$ ; Figure 4).

**Figure 4:** Replication Mendelian randomisation analysis (inverse variance weighted method) and meta-analysis results for the association between genetically predicted momp D antibody seropositivity level and ovarian cancer risk. The replication cohorts were UK Biobank and FinnGen combined. The meta-analysis combined the estimate from the replication cohorts with that from the Ovarian Cancer Association Consortium (OCAC) case-control studies. All odds ratios and confidence intervals are in units of per standard deviation increase in log-transformed antibody mean fluorescence intensity.



## Discussion

In this two-sample MR study, genetically predicted seropositivity to the *C. trachomatis* major outer membrane protein (momp) D antibody was causally associated with overall and high-grade serous EOC. This association with overall EOC remained when replicated in an independent population. There was no evidence of an association existing with other EOC histotypes.

This finding is partially consistent with a number of published observational studies investigating the relationship between exposure to STIs and EOC. A UK-based prospective analysis by Idahl *et al* using the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, including 791 EOC cases and 1669 controls, found *C. trachomatis* infection history (assessed by antibodies to chlamydia heat shock protein 60 (cHSP60-1)) was associated with higher risk of overall (1.36 [1.13-1.64]) and high-grade serous (1.44 [1.12-1.85]) EOC.<sup>5</sup> Similarly, Fortner *et al* found *C. trachomatis* Pg3p antibody seropositivity was associated with higher risk of ovarian cancer overall (RR = 2.07 [1.25–3.43]) using 337 cases from the Nurses' Health Studies. Their results were similar for invasive and low-malignant potential tumours, and no associations were observed for other STIs.<sup>3</sup> Trabert *et al* conducted two case-control studies of 224 and 160 EOC cases in Poland and the USA respectively, finding *C. trachomatis* Pg3p antibody seropositivity to be associated with double the risk of EOC after adjustment. Again, no significant associations with other STIs were observed.<sup>4</sup> However, a Finnish prospective study of 484 cases by Skarga *et al* found no association between *C. trachomatis* Pgp3 seropositivity and EOC, although did between *M. genitalium* and mucinous EOC (RR, 1.66 [95% CI, 1.09-2.54]).<sup>14</sup> Two further case-control studies from Italy and Canada found no significant association between *C. trachomatis* and EOC.<sup>17,18</sup>

Interestingly, findings from observational studies exploring the association between PID and EOC are also mixed. Jonsson *et al* conducted a case-control study of 15 072 Swedish EOC cases, reporting associations with PID (OR of 1.39; 95% CI 1.17—1.66).<sup>13</sup> Meanwhile, a prospective study of 5356 women in Denmark found no associations with PID (hazard ratio (HR), 1.05; 95% CI, 0.92—1.20).<sup>16</sup> A prospective study of 454 Australian women found high-grade serous EOC was associated with previous PID (HR 1.47; 95% CI 1.04-2.07), but not endometriosis or infertility.<sup>15</sup> Conversely, a study of 600 African-American women found PID to not be associated with EOC, but endometriosis was.<sup>29</sup> A further pooled analysis of 13 case-control studies concluded the relationship to be uncertain and liable to uncontrolled confounding.<sup>30</sup>

Some small-scale non-observational studies support the role of *C. trachomatis* in EOC carcinogenesis, whereby *C. trachomatis* DNA has been detected in 84% of high-grade serous and 17% of invasive EOC.<sup>7,19</sup> Although limited by sample size, these studies may elucidate potential causative mechanisms underlying the observed association. Due to the high prevalence of asymptomatic *C. trachomatis* infections, exact incidence is unclear. It has been suggested that without treatment, approximately half of *C. trachomatis* infected women will develop chronic infection.<sup>31</sup> Therefore, *C. trachomatis* infections often persist, and the link between chronic infection, inflammation and cancer has been extensively described.<sup>32</sup> Additionally, *C. trachomatis* bacteria have the capability to enter a viable but non-replicative persistent state, effectively evading the immune response of the host cell.<sup>33,34</sup> During this persistent state, the bacteria increase the production of cHSP60-1, an anti-apoptotic protein, while decreasing the synthesis of structural and membrane proteins.<sup>34</sup> *C. trachomatis* can inhibit the release of mitochondrial cytochrome C and caspase 3, thus enabling the infected cell to evade intrinsic apoptosis.<sup>35</sup> This ability to resist apoptosis prolongs the survival of the infected host cell and allows potentially DNA-damaged cells to persist, heightening the risk of STIC lesions and carcinogenesis.

This study has several strengths. First, we utilised MR to assess the causal association of genetic susceptibility to *C. trachomatis* with EOC, which can provide reliable evidence on the effect of modifiable risk factors for disease and overcomes some limitations of traditional observational epidemiology – such as study feasibility, ethical constraints, and confounding. Second, there were ten momp D seropositivity-associated lead SNPs, allowing us to conduct multiple sensitivity analyses to ensure robustness of our findings and determine the absence of horizontal pleiotropy. Third, our findings of an association have been further strengthened by positive associations holding when replicated using an independent Global biobank population. Finally, we used meta-analysis of our findings to give an overall effect estimate which may be useful to inform clinical or public health decision making.

Nonetheless, there are limitations. First, MR studies rely on a key assumption that the genetic instruments only act on the outcome via the exposure. Although this study yielded an insignificant MR-Egger intercept test, the possibility of horizontal pleiotropy cannot be completely ruled out. Third, the weighted median, simple mode, and weighted mode methods resulted in insignificant association in all the analyses, most likely attributable to the lower statistical power with reference to other MR methods.<sup>36</sup> For the primary analysis using the OCAC GWAS, case numbers among the less common histotypes were limited, potentially underpowering results. For the replication analyses using the Global biobanks GWAS, outcome data combined all EOC histotypes together – both serous and non-serous – for which the pathology and epidemiology are known to differ.<sup>7</sup> We used *C. trachomatis* momp D-associated lead SNPs as our genetically predicted exposure; Butler-Laporte et al found momp D to have a relatively high genomic inflation factor of 1.24.<sup>21</sup> This questions whether all the SNPs identified show a true association with the phenotype (i.e. momp D seropositivity).<sup>37</sup> However, the corrected inflation factor ( $\lambda_{1000}$ ) is 1.10 when the population size is considered, which is generally



considered acceptable for GWAS.<sup>38</sup> Furthermore, momp D is not the gold standard serological test for *C. trachomatis*. Although momp D antibody specificity is directly comparable to that of the Pgp3 antibody, Pgp3 is recognised as the gold standard as it is more sensitive and may be associated with longer persistence than momp D.<sup>39</sup> However, Pgp3 was only associated with one SNP, unlike momp D which was associated with eleven, therefore it was not possible to use this as our study exposure given inability to perform sensitivity analyses. Both Pgp3 and momp D antibodies are general markers of *C. trachomatis* infection, and do not necessarily indicate persistent, untreated *C. trachomatis* infections which we postulate are causally associated with EOC. Many people infected with *C. trachomatis* in their lifetime do not develop chronic infection due to natural immunity or treatment, and it is unlikely these individuals will develop any long-term consequences – such as EOC. cHSP60-1, when produced by *C. trachomatis* bacteria, is associated with a state of persistent, untreated chlamydia infection and its sequelae – including evasion of immune defence, inhibition of cell apoptosis, and survival of a DNA-damaged cell. Therefore, it is proposed that cHSP60-1 production is a potential pathogenic mechanism by which *C. trachomatis* could cause EOC, thus positive serology for cHSP60-1 antibodies would be a more useful marker than momp D or Pgp3 to investigate whether an association with EOC exists.<sup>5,7</sup> Unfortunately, no GWAS has yet included cHSP60-1 antibodies. It is also important to note that effect estimates, representing lifetime risk, may be somewhat underestimated due to prevalence of momp D seropositivity in individuals previously treated or naturally cleared of *C. trachomatis* infection.<sup>40</sup> Finally, the GWAS meta-analysis from which the summary statistics were retrieved for this study only included Europeans and a small sample from Japan; therefore the generalisability to other populations are unknown.

This study, shedding light on the causative link between *C. trachomatis* and EOC, has several public health implications. Affecting approximately one in twenty women aged 15-24 years in the USA, *C. trachomatis* is highly prevalent and often silent.<sup>41</sup> Our study provides compelling further

reasons why accessible sexual health education and services are vital for short and long-term population health protection targets. The first phase one trial for a *C. trachomatis* vaccination was successfully completed in 2019, with others in pre-clinical development.<sup>42,43</sup> Our study, highlighting this causative link, may further motivate stakeholders to continue supporting the development of *C. trachomatis* vaccines, with an aim to roll out in adolescence alongside HPV-vaccines.<sup>43</sup> The causal role of *C. trachomatis* in the pathogenesis of other cancers, including anorectal and cervical, also requires further evaluation,<sup>44-46</sup> as does EOC risk in the context of other STIs. *M. genitalium* is a small intracellular bacterium known to cause PID, and although less prevalent than *C. trachomatis*, it is also often asymptomatic, associated with PID, and frequently resistant to antibiotics. Observational studies have described associations with EOC, and with prevalence rising, it deserves attention.<sup>3,5,14</sup> A study by Shanmughapriya *et al* suggested associations exist between intrauterine device (IUD) use and subsequent PID and EOC due to IUD colonisation by actinomycetes and disruption of the normal genitourinary flora.<sup>47,48</sup> This is a further area requiring additional exploration, given the potential implications if found to be true. Finally, our study has highlighted STI exposure as an important aetiological aspect of EOC and therefore may have a role in cancer polygenic risk scores and stratified population screening strategies.<sup>49</sup>

In conclusion, this MR study confirms the causative link between *C. trachomatis* infection and overall and high-grade serous EOC. As a key modifiable risk factor for future serous EOC, primary prevention of *C. trachomatis* infection is a crucial public health target and may help reduce burden of EOC much like HPV-vaccination has for cervical cancer. Further studies into the causative role of other STIs and EOC, particularly *M. genitalium*, and *C. trachomatis* and other cancers, namely anorectal, may represent further opportunities to reduce cancer risk.

## Data availability

All data sets used in this study are publicly available at <https://gwas.mrcieu.ac.uk/datasets/> with the following data set identifiers – OCAC ovarian cancer: ieu-a-1120 (overall EOC), ieu-a-1121 (high-grade serous EOC), ieu-a-1122 (low-grade serous EOC), ieu-a-1123 (invasive mucinous EOC), ieu-a-1124 (clear cell EOC), ieu-a-1125 (endometrioid EOC), ieu-a-1230 (serous low malignant potential EOC), ieu-a-1232 (mucinous low malignant potential EOC); UK Biobank + FinnGen ovarian cancer: ebi-a-GCST90018888; UK Biobank *Chlamydia trachomatis* momp D antibody seropositivity level: ebi-a-GCST90018888.

## Acknowledgements

We thank all participants in the OCAC studies and the UK Biobank and FinnGen cohorts.

## Funding

This work was supported by a US National Institutes of Health (NIH) grant to SPK (grant number R01CA259058), UK Research and Innovation Future Leaders Fellowship to SPK [grant number MR/T043202/2], and the UK National Health Service Health Education England East of England Foundation School Specialised Foundation Programme for SLP.

## Author information

### *Authors and Affiliations*

### **Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom**

Sarah L. Perrott

### **Early Cancer Institute, University of Cambridge, Cambridge, United Kingdom**

Sarah L. Perrott & Siddhartha P. Kar

### *Contributions*

SPK conceived and designed the study. SLP analysed the data, prepared figures and tables, and wrote the initial draft of the manuscript. Both authors edited the initial draft of the manuscript and approved the final version.

### *Corresponding Author*

Correspondence to Sarah L. Perrott ([sp2176@cam.ac.uk](mailto:sp2176@cam.ac.uk)) and Siddhartha P. Kar ([sk718@cam.ac.uk](mailto:sk718@cam.ac.uk))

### **Ethics declarations**

#### *Competing interests*

The authors declare no competing interests

#### *Ethics approval and consent*

This study is a secondary analysis of publicly available GWAS summary statistics data sets.

### **Figure Legends**

**Figure 1:** Study design and assumptions of two-sample Mendelian randomisation analysis.

**Figure 2:** Inverse variance weighted Mendelian randomisation analysis results for the association between genetically predicted momp D antibody seropositivity level and overall and histotype-specific epithelial ovarian cancer risk. All odds ratios and confidence intervals are in units of per standard deviation increase in log-transformed antibody mean fluorescence intensity.

**Figure 3:** Mendelian randomisation (MR) sensitivity analysis results for the association between genetically predicted momp D antibody seropositivity level and (A) overall and (B) high-grade serous epithelial ovarian cancer risk. All odds ratios and confidence intervals are in units of per standard deviation increase in log-transformed antibody mean fluorescence intensity.

**Figure 4:** Replication Mendelian randomisation analysis (inverse variance weighted method) and meta-analysis results for the association between genetically predicted momp D antibody seropositivity level and ovarian cancer risk. The replication cohorts were UK Biobank and

FinnGen combined. The meta-analysis combined the estimate from the replication cohorts with that from the Ovarian Cancer Association Consortium (OCAC) case-control studies. All odds ratios and confidence intervals are in units of per standard deviation increase in log-transformed antibody mean fluorescence intensity.

## References

1. World Health Organisation. Chlamydia. <https://www.who.int/news-room/fact-sheets/detail/chlamydia#> (2023).
2. Huai, P. *et al.* Prevalence of genital Chlamydia trachomatis infection in the general population: A meta-analysis. *BMC Infect Dis* **20**, 1–8 (2020).
3. Fortner, R. T. *et al.* Sexually transmitted infections and risk of epithelial ovarian cancer: results from the Nurses' Health Studies. *Br J Cancer* **120**, 855–860 (2019).
4. Trabert, B. *et al.* Antibodies against chlamydia trachomatis and ovarian cancer risk in two independent populations. *J Natl Cancer Inst* **111**, 129–136 (2019).
5. Idahl, A. *et al.* Serologic markers of Chlamydia trachomatis and other sexually transmitted infections and subsequent ovarian cancer risk: Results from the EPIC cohort. *Int J Cancer* **147**, 2042–2052 (2020).
6. Yang, D. *et al.* Global, regional, and national burden of ovarian cancer and the attributable risk factors in all 194 countries and territories during 2007–2017: A systematic analysis of the Global Burden of Disease Study 2017. (2021).
7. Jonsson, S., Oda, H., Lundin, E., Olsson, J. & Idahl, A. Chlamydia trachomatis, Chlamydial Heat Shock Protein 60 and Anti-Chlamydial Antibodies in Women with Epithelial Ovarian Tumors Translational Oncology. *Transl Oncol* **11**, 546–551 (2018).
8. McCluggage, W. G. Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis. *Pathology* **43**, 420–432 (2011).
9. Kyo, S., Ishikawa, N., Nakamura, K. & Nakayama, K. The fallopian tube as origin of ovarian cancer: Change of diagnostic and preventive strategies. *Cancer Med* **9**, 421 (2020).
10. Ingerslev, K. *et al.* The potential role of infectious agents and pelvic inflammatory disease in ovarian carcinogenesis. *Infect Agent Cancer* **12**, (2017).
11. Zitvogel, L., Daillère, R., Roberti, M. P., Routy, B. & Kroemer, G. Anticancer effects of the microbiome and its products. *Nature Publishing Group* **15**, 465–478 (2017).
12. Whisner, C. M. & Aktipis, C. A. The Role of the Microbiome in Cancer Initiation and Progression : How Microbes and Cancer Cells Utilize Excess Energy and Promote One Another's Growth. *Curr Nutr Rep* **8**, 42–51 (2019).
13. Jonsson, S., Jonsson, H., Lundin, E., Häggström, C. & Idahl, A. Pelvic inflammatory disease and risk of epithelial ovarian cancer: a national population-based case-control study in Sweden. (2024).
14. Skarga, E., Surcel, H. M., Kaaks, R., Waterboer, T. & Fortner, R. T. Sexually Transmitted Infections and Risk of Epithelial Ovarian Cancer: Results From the Finnish Maternity Cohort. *J Infect Dis* **228**, 1621–1629 (2023).

15. Stewart, L. M. *et al.* Risk of high-grade serous ovarian cancer associated with pelvic inflammatory disease, parity and breast cancer. *Cancer Epidemiol* **55**, 110–116 (2018).
16. Rasmussen, C. B., Jensen, A., Albieri, V., Andersen, K. K. & Kjaer, S. K. Is pelvic inflammatory disease a risk factor for ovarian cancer? *Cancer Epidemiology Biomarkers and Prevention* **26**, 104–109 (2017).
17. Wong, A. *et al.* Could epithelial ovarian cancer be associated with chlamydial infection? *European Journal Gynaecological Oncology* **28**, 117–120 (2007).
18. Parazzini, F. *et al.* Pelvic inflammatory disease and risk of ovarian cancer. *Cancer Epidemiol Biomarkers Prev* **5**, 667–669 (1996).
19. Laban, M. *et al.* Chlamydia trachomatis infection in primary fallopian tube and high-grade serous ovarian cancers: a pilot study. *Int J Womens Health* **11**, 199–205 (2019).
20. Yavorska, O. O. & Burgess, S. MendelianRandomization: An R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol* **46**, 1734–1739 (2017).
21. Butler-Laporte, G. *et al.* Genetic Determinants of Antibody-Mediated Immune Responses to Infectious Diseases Agents: A Genome-Wide and HLA Association Study. *Open Forum Infect Dis* **7**, (2020).
22. Phelan, C. M. *et al.* Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet* **49**, 680–691 (2017).
23. Persyn, E., Redon, R., Bellanger, L. & Dina, C. The impact of a fine-scale population stratification on rare variant association test results. *PLoS One* **13**, (2018).
24. Morris, J. A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet* **51**, 258 (2019).
25. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nature Genetics* **53**, 1415–1424 (2021).
26. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
27. Swerdlow, D. I. *et al.* Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *Int J Epidemiol* **45**, 1600 (2016).
28. Skrivankova, V. W. *et al.* Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *JAMA* **326**, 1614–1621 (2021).
29. Park, H. K. *et al.* Benign gynecologic conditions are associated with ovarian cancer risk in African-American women: a case-control study. *Cancer Causes Control* **29**, 1081–1091 (2018).
30. Rasmussen, C. B. *et al.* Pelvic Inflammatory Disease and the Risk of Ovarian Cancer and Borderline Ovarian Tumors: A Pooled Analysis of 13 Case-Control Studies. *Am J Epidemiol* **185**, 8–20 (2017).
31. Geisler, W. M. Duration of Untreated, Uncomplicated Chlamydia trachomatis Genital Infection and Factors Associated with Chlamydia Resolution: A Review of Human Studies. doi:10.1086/652402.
32. Coussens, L. M. & Werb, Z. Inflammation and cancer. *Nature* **420**, 860 (2002).

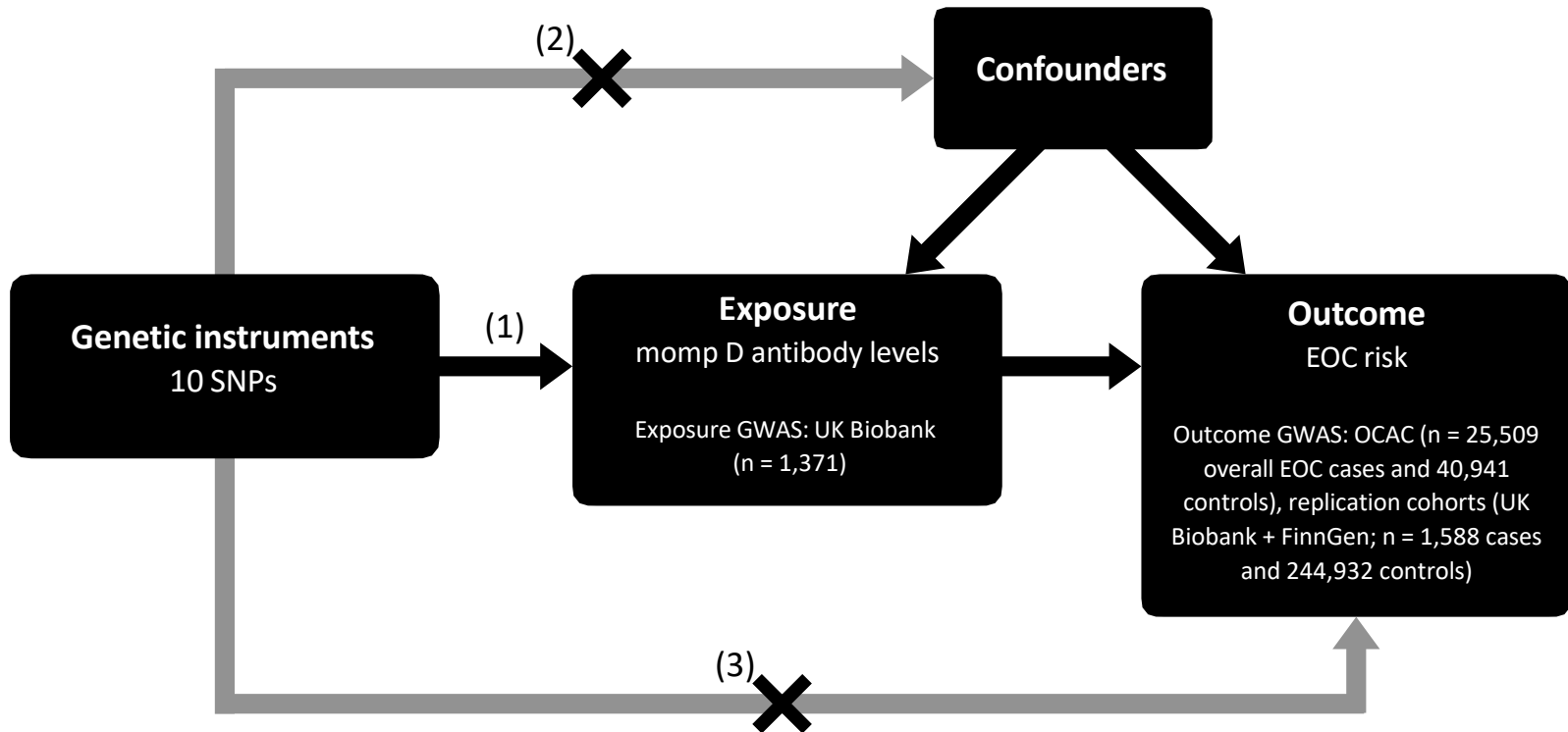
33. Beatty, W. L., Belanger, T. A., Desai, A. A., Morrison, R. P. & Byrne, G. I. Tryptophan depletion as a mechanism of gamma interferon-mediated chlamydial persistence. *Infect Immun* **62**, 3705–3711 (1994).
34. Dean, D. & Powers, V. C. Persistent Chlamydia trachomatis Infections Resist Apoptotic Stimuli. *Infect Immun* **69**, 2442–2447 (2001).
35. Fan, T. *et al.* Inhibition of Apoptosis in Chlamydia-infected Cells: Blockade of Mitochondrial Cytochrome c Release and Caspase Activation. *J. Exp. Med* **187**, 487–496 (1998).
36. Bowden, J., Smith, G. D. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* **44**, 512–525 (2015).
37. Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *European Journal of Human Genetics* **2011 19:7 19**, 807–812 (2011).
38. de Bakker, P. I. W. *et al.* Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum Mol Genet* **17**, R122 (2008).
39. Anyalechi, G. E. *et al.* Chlamydial Pgp3 Seropositivity and Population-Attributable Fraction among Women with Tubal Factor Infertility. *Sex Transm Dis* **49**, 527–533 (2022).
40. Davies, N. M., Holmes, M. V. & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* **362**, 601 (2018).
41. Kreisel, K. M., Weston, E. J., St. Cyr, S. B. & Spicknall, I. H. Estimates of the Prevalence and Incidence of Chlamydia and Gonorrhea Among US Men and Women, 2018. *Sex Transm Dis* **48**, 222 (2021).
42. Abraham, S. *et al.* Safety and immunogenicity of the chlamydia vaccine candidate CTH522 adjuvanted with CAF01 liposomes or aluminium hydroxide: a first-in-human, randomised, double-blind, placebo-controlled, phase 1 trial. *Lancet Infect Dis* **19**, 1091–1100 (2019).
43. Collar, A. L., Clarke, T. N., Jamus, A. N. & Frieze, K. M. Ensuring equity with pre-clinical planning for chlamydia vaccines. *npj Vaccines* **2023 8:1 8**, 1–7 (2023).
44. Chumduri, C., Gurumurthy, R. K., Zadora, P. K., Mi, Y. & Meyer, T. F. *Chlamydia* Infection Promotes Host DNA Damage and Proliferation but Impairs the DNA Damage Response. *Cell Host Microbe* **13**, 746–758 (2013).
45. Heijne, J. C. M. *et al.* What explains anorectal chlamydia infection in women? Implications of a mathematical model for test and treatment strategies. *Sex Transm Infect* **93**, 270–275 (2017).
46. Yang, X. *et al.* Chlamydia Trachomatis Infection: Their potential implication in the Etiology of Cervical Cancer. *J Cancer* **12**, 4891 (2021).
47. Shanmughapriya, S., Lency, A., Kavitha, S. & Natarajaseenivasan, K. Prolonged aminoglycoside treatment for pelvic inflammatory disease associated ovarian carcinoma. *Trends in Medical Research* **11**, 49–53 (2016).
48. Brown, B. P. *et al.* Copper intrauterine device increases vaginal concentrations of inflammatory anaerobes and depletes lactobacilli

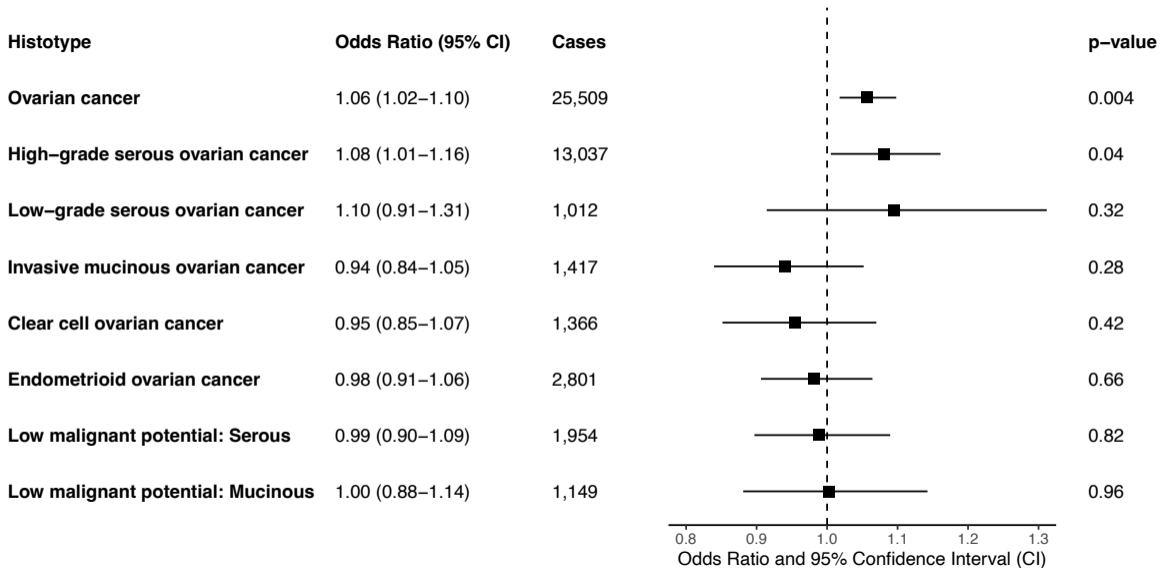


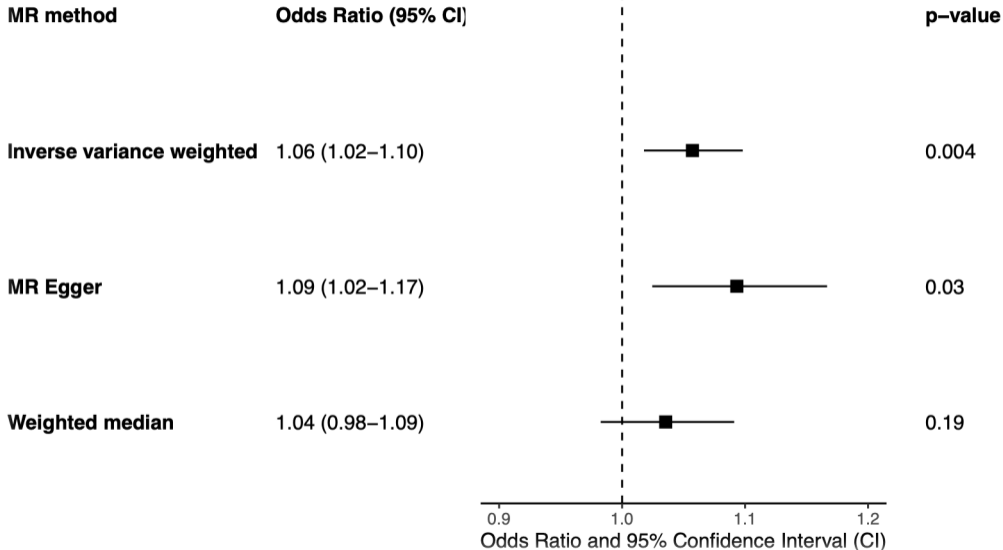
compared to hormonal options in a randomized trial. *Nature Communications* 2023 14:1 **14**, 1–13 (2023).

49. Huntley, C. *et al.* Utility of polygenic risk scores in UK cancer screening: a modelling analysis. *Lancet Oncology* **24**, 658–668 (2023).

- (1) Relevance assumption: Genetic instruments (i.e., SNPs) are associated with the exposure.
- (2) Independence assumption: These SNPs are not associated with known or unknown confounding factors that affect the exposure-outcome relationship.
- (3) Exclusion restriction assumption: The impact of SNPs on the outcome occurs solely through the designated exposure and not through an alternative pathway (i.e., not via horizontal pleiotropy).







**MR method**

**Odds Ratio (95% CI)**

**p-value**

**Inverse variance weighted**

1.08 (1.01–1.16)

0.04

**MR Egger**

1.15 (1.02–1.30)

0.05

**Weighted median**

1.04 (0.97–1.12)

0.23

Odds Ratio and 95% Confidence Interval (CI)

MR method	Odds Ratio (95% CI)	p-value
Inverse variance weighted	1.08 (1.01–1.16)	0.04
MR Egger	1.15 (1.02–1.30)	0.05
Weighted median	1.04 (0.97–1.12)	0.23

