Assessing the Causal Effects of Environmental Tobacco Smoke Exposure:
A meta-analytic Mendelian randomisation study

Benjamin Woolf\textsuperscript{1,2,3}, Skanda Rajasundaram\textsuperscript{4,5}, Dipender Gill\textsuperscript{6,7 *}, Hannah M. Sallis\textsuperscript{2,8 *}, Marcus R. Munaf\textsuperscript{1,2 *}

\textsuperscript{1} School of Psychological Science, University of Bristol, Bristol, UK
\textsuperscript{2} MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK
\textsuperscript{3} MRC Biostatistics Unit at the University of Cambridge, Cambridge, UK
\textsuperscript{4} Centre for Evidence-Based Medicine, University of Oxford, Oxford, UK
\textsuperscript{5} Faculty of Medicine, Imperial College London, London, UK
\textsuperscript{6} Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom
\textsuperscript{7} Chief Scientific Advisor Office, Research and Early Development, Novo Nordisk, Copenhagen, Denmark
\textsuperscript{8} Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

* These authors contributed equally.

Word count: 4,400

Corresponding Author: Benjamin Woolf, School of Psychological Science, University of Bristol, Bristol, UK. benjamin.woolf@bristol.ac.uk

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Objectives: Cigarette smoking is a major cause of global morbidity and premature mortality. There is evidence that exposure to environmental tobacco smoke (ETS; “second-hand smoking” or “passive smoking”) contributes substantially to ill health. However, it is difficult to establish causality given well-described problems of confounding and selection bias. We applied Mendelian randomisation (MR) to investigate the causal effects of ETS exposure.

Design: We employed six MR estimation approaches: the effects of each parent’s smoking on their offspring, the effects of each parent’s smoking on the other parent, and the effects of self-reported exposure to second-hand smoking both inside and outside of the home.

Setting: Genome-wide association studies (GWASs) with sample sizes ranging from 397,732 to 909,629 individuals of predominantly European ancestry.

Interventions: To mimic exposure to ETS in the first four approaches, we used an individual’s exposure to their relative’s genetic liability to smoke, conditional on the original individual’s genetic liability to smoke. In the final two approaches we used instruments associated with self-reported exposure to ETS.

Outcomes: Lung cancer, chronic obstructive pulmonary disease (COPD), stroke, cardiovascular disease, hypertension, and depression.

Results: Our findings support a causal effect of genetically predicted ETS exposure on lung cancer and COPD (odds per hour exposed in a typical week = 1.41 [95% CI: 1.23 to 1.61] and 1.11 [95% CI: 1.07 to 1.16] respectively). We did not find evidence supporting an effect on other outcomes.

Discussion: ETS may cause both lung cancer and COPD. Although our results do not support an effect of ETS on other outcomes, this might reflect low statistical power to detect smaller effect sizes. These results support public health measures to limit exposure to ETS.
Key message

- Smoking is a major cause of global disease burden.
- Environmental tobacco smoking (ETS; “second-hand smoking” or “passive smoking”) occurs when an individual breathes in another individual’s cigarette smoke.
- Much of the existing literature on ETS has been obtained from observational studies which may be biased by confounding, selection bias, and reverse causation.
- We leveraged a quasi-experimental method, Mendelian randomisation, which can be more robust to these biases to assess the causal effects of ETS exposure.
- Our results imply an effect of ETS exposure on lung cancer and COPD, and therefore support measures to reduce ETS exposure.
- The absence of evidence for effects of ETS on hypertension, stroke, CVD, or depression could reflect low statistical power in our study.
Introduction

Cigarette smoking is associated with numerous adverse health outcomes, ranging from lung cancer to depression (1–7). Over several decades, a range of evidence derived from various methodological approaches has accumulated to demonstrate that the harmful effects of cigarette smoking are indeed causal. Consequently, numerous public health campaigns have attempted to reduce the prevalence of smoking (8). Many of the associations between cigarette smoking and adverse health outcomes have been replicated for environmental tobacco smoke exposure (9–13). Environmental tobacco smoke exposure (ETS, also “passive smoking” or “second-hand smoking”) occurs when an individual breathes in the smoke from another individual’s cigarette smoke.

For both cigarette smoking and ETS exposure, establishing a causal effect on health outcomes can be undermined by confounding and reverse causation. However, establishing causation for ETS exposure is even more challenging than for cigarette smoking. For example, individuals tend to choose friends of a similar smoking status’ (14), which may induce selection effects in studies of ETS exposure. When studying an outcome caused by cigarette smoking, any observed association with ETS exposure could simply be a function of someone selecting friends with a similar smoking status and the direct effect of cigarette smoking. This type of selection bias has been shown to be difficult to account for in observational analyses of socially transmissible phenotypes (15). In addition, it is difficult to study the effects of exposure to ETS using most quasi-experimental methods because most interventions targeting ETS (e.g., banning smoking in indoor public spaces) may also influence the rate of first-hand smoking (e.g., by reducing the rate that someone might smoke socially). Studies using other methods are therefore warranted.

Mendelian randomisation (MR) leverages the random allocation of genetic variants at conception, analogous to random allocation in a randomised controlled trial (16,17), to reduce bias due to confounding. Furthermore, because germline genetic variants are fixed at conception, MR reduces susceptibility to reverse causality. Thus, under specific assumptions, MR can strengthen causal inference. The three core assumptions are that the genetic variants robustly associate with the exposure of interest, that there are no variant-outcome confounders; and that the variants influence the outcome only via the exposure of interest (18).
Traditional MR analyses can be biased by ‘indirect genetic effects’ introduced because the association between the genetic variants and the child’s phenotype is partially mediated by the parent’s heritable phenotype and the latter’s effect on the child (19). Recently, MR has been extended to ‘within family’ settings to address this and other biases. A typical within family MR study leverages family data to strengthen the validity of the second MR assumption by ensuring that the distribution of genetic variants is truly random (20).

Here, we use a novel application of family data within an MR framework to robustly explore the causal effects between relatives’ ETS exposure on a range of health outcomes. Unlike traditional within family MR studies, which attempt to remove indirect genetic effects, we leverage them to instrument the phenotypic effect of one relative on another. Because the non-inheritance of a genetic variant is also at random, if an individual’s relative inherits a genetic variant which robustly associates with smoking, but this variant was not inherited by the individual, then the individual has in effect been randomised to exposure to ETS via the relative (Figure 1). When the genetic relatedness between relatives is not accounted for and uses both inherited and non-inherited variants as instruments, the resulting MR estimate will be biased by the direct genetic effects of the inherited variants on the relative. Additionally, because biological relatives, like parents and children, do not choose each other, this MR design should avoid potential selection biases (21).

We therefore investigated the causal effects of ETS exposure on a range of health outcomes using an MR study design. Specifically, we use six approaches, analysed with publicly available data – mostly from the UK Biobank. We examined the effect of each parent’s smoking on their offspring’s outcome, the effect that each parent has on the other parent, and the effects of self-reported exposure to second-hand smoking inside and outside of the home. We focused on six outcomes (lung cancer, chronic obstructive pulmonary disease (COPD), stroke, cardiovascular disease, hypertension, and depression) which have previously been identified as being associated with ETS exposure and were measured in the UK Biobank (9–13). After a risk of bias evaluation, approaches which were independent and low risk of bias were meta-analysed to provide a pooled estimate of the effect of ETS on each outcome.
Methods

Overview of MR estimation approaches

We designed 6 approaches that use MR to estimate the effect of exposure to environmental tobacco smoke (ETS). Each approach used publicly available GWAS summary statistics and had six outcomes (lung cancer, COPD, stroke, cardiovascular disease, hypertension, and depression):

First, we explored the effect of maternal smoking on the index individual’s outcomes (row 1 and column 1 of Table 1). This approach was conducted using a multivariable extension of the ‘proxy gene-by-environment’ MR design (22). We used the associations between the index individual’s genotype and their mother’s smoking as instruments. These occur because children inherit half of each parent’s variants but these variants also affect the respective parental phenotype. Since we only used inherited variants, we adjusted for the direct effect that the inherited variants have on the index individual’s smoking, and therefore outcome (Figure 1), with multivariable MR (MVMR) by including the index individual’s smoking as a covariate. Not accounting for the effects that the variants have through the index individual’s smoking would likely result in an overestimation of the true effect (23). The effect estimates of MVMR are interpreted as the direct effect of the exposure conditional on the other covariates. Because of this, MVMR has been previously used to adjust MR estimates for postulated biases, by ensuring that the effect of interest is conditionally independent of a phenotype of concern (24–26). Thus, this estimation approach would provide the effect of maternal smoking on their offspring outcomes independent of the offspring’s liability to smoke. We call this application of MVMR multivariable gene-by-environment MR.

Second, we applied the same logic to explore the effects of paternal smoking on the index individual’s outcomes. This approach (row 1 and column 2 of Table 1) thus included the index individual’s smoking as a covariate in a multivariable gene-by-environment MR analysis of paternal smoking on the index individual’s outcomes. By conditioning on the index smoking, it estimates the direct effect of paternal smoking on the index’s outcome, independent of the index’s smoking.

Third, we leveraged the available GWASs for maternal smoking and paternal outcomes to explore the effect that maternal smoking has on paternal outcomes (row 2 column 1 of Table 1). Here the likely source of bias is assortative mating, where parents tend to partner with people who have a
similar smoking status. A naive association between maternal smoking and paternal outcomes could simply represent the combined effects of paternal first-hand smoking, and smoking mothers choosing partners who smoke (14). Using similar logic to the first two approaches, we controlled for assortative mating by including paternal smoking as a covariate in a multivariable gene-by-environment MR model. The estimates of this approach should be interpreted as the direct effects of maternal smoking on the paternal outcome, independent of paternal smoking.

The fourth approach likewise used a GWAS of paternal smoking and maternal outcomes to explore the effects of paternal smoking on maternal outcomes (row 2 column 2 of Table 1) (27). This estimation approach therefore used multivariable gene-by-environment MR to estimate the direct effect of paternal smoking on the maternal outcome, by conditioning maternal smoking on an MR model of paternal smoking on maternal outcome.

Finally, in the fifth and sixth approaches we explored the effect of the index individual’s reported exposure to other people’s smoking both inside and outside of the home, respectively, on the index individual’s outcomes using traditional univariable MR analyses (row 3 of Table 1). Since these approaches used univariable MR, they estimate the total effect of ETS inside or outside of the home respectively.

Additional details such as information about the data sources (including sample sizes) and statistical methods used in the MR estimation approaches can be found in the Supplementary Methods. In brief, we extracted genome-wide significant variants from exposure GWASs and we used the TwoSampleMR R package to harmonise and clump the genetic data (\(r^2 = 0.001, \text{kb} = 10,000\)) (28). All exposure GWASs and parental outcome GWASs were derived using a sub-sample of genetically unrelated participants in the UK Biobank (UKB), which has been described elsewhere (29). However, the index individual outcome GWASs combined the UKB with independent, but demographically similar, samples of the same trait (UKB + International Lung Cancer Consortium for lung cancer (30,31), UKB + Psychiatric Genetic Consortium and 23andMe for depression (32), UKB + FinnGen for COPD and hypertension (33,34), UKB + CARDioGRAMplusC4D for cardiovascular disease (35), and UKB + the European ancestry sub-sample of MEGASTROKE for stroke (36)). When needed, we tested the same-population assumption using the MRSamePopTest R package (37),
and addressed the Winner’s Curse by using False discovery rate Inverse Quantile Transformation (FIQT) winners curse correction (26,38).

**Sensitivity analyses for the MR estimation approaches**

We ran four sensitivity analyses for the MR estimation approaches. First, as a positive control, we ran a univariate MR analysis of the index individual’s own cigarette smoking on their outcomes. This also demonstrates that there was sufficient statistical power to detect an effect on the outcomes. Second, we used hair colour as a negative control outcome for residual population structure. Natural hair colour is known to vary by population sub-group in the UK but should not be causally related to any of the outcomes or smoking status. Therefore, if the any of the genetic instruments used here (i.e., for the index individual’s smoking, maternal smoking, or paternal smoking) associate with hair colour, then this will most likely be due to residual population structure (39). Third, the UK Biobank genotyped participants (recruited into the UK BiLEVE study) at the extreme and middle of the distributions for smoking and lung function were genotyped using a different genotyping chip as compared with the rest of the sample. This could potentially introduce confounding into genotype-phenotype associations, and previous guidance suggests adjusting for the genotyping chip in analyses (29). However, for smoking or lung function-related phenotypes, adjusting for genotyping chip could additionally introduce collider/selection bias. Although such bias can theoretically be removed by using MVMR (40,41), prior research indicates that the use of MVMR in this context has little overall effect on the final MR estimates (40). Furthermore, adding additional covariates in an MVMR model greatly increases the risk of weak instrument bias and reduces power. Our primary analysis therefore used the genotyping chip adjusted GWAS without additional adjustment for genotyping chip via MVMR. We then performed two sensitivity analyses, one using GWASs that did not adjust for genotyping chip, and a second using MVMR to adjust for genotyping chip. Finally, because our multivariable gene-by-environment MR analyses adjusted for the outcome individual’s smoking using MVMR, we explore the power-bias trade off from using MVMR by replicating the chip adjusted analyses using univariable MR.
MR estimation approach risk of bias assessment

To our knowledge, there is no existing tool for assessing the risk of bias in MR estimation approaches (or studies more generally). We therefore created a bespoke tool based on the recent systematic review of tools to assess risk of bias in MR studies by Spiga and colleagues (42). Our bespoke tool is described in detail in the Supplementary Methods. Ratings were made initially by BW and then ratified by SR.

Synthesis of MR estimation approaches

A quantitative synthesis of the 6 MR estimation approaches would increase the precision of the causal effect estimate. For such a meta-analysis to be valid, however, we must assume that the design of each estimation approach is independent of one other. In the context of a two-sample MR analysis, this requires that either the instruments are (conditionally) independent of each other, or that the outcome samples are independent of each other. The first is equivalent to a factorial experiment in which people are independently randomised to two separate interventions with the same intended therapeutic effect, e.g., two different drugs that lower blood pressure. The second is analogous to the assumption of no participant overlap made in traditional meta-analyses of randomised controlled trials.

Supplementary Table S1 presents pairwise comparisons of each of the MR estimation approaches and the rationale for the independence, or lack of independence, between them. Of the 15 pairs, independence was questionable for two pairs (between the two estimation approaches using self-reported ETS exposure, and the two estimation approaches with a parent’s smoking and the index individual outcomes). When included in the same meta-analysis, we accounted for within-pair potential correlations by using the average of the treatment effect estimates and variances of each potentially correlated pairs being meta-analysed (43,44). We used leave-one-out sensitivity analyses to explore if the standard error of the primary meta-analysis was robust to potential ‘unmeasured’ correlated pairs. Of the remaining 13 pairs, 2 had both non-overlapping outcome GWASs and independent primary instruments, 7 had non-overlapping outcome GWASs, and 4 had independent primary instruments.
We stratified our meta-analysis by estimation approach risk of bias, and – following guidance in the Cochrane handbook – place primary emphasis on the low risk of bias estimation approaches (45). Results were presented using a forest plot. Our primary estimator for the meta-analyses were fixed effect models (46). The summary measure of the association in meta-analyses was the log odds in the outcome per genetically predicted standard deviation increase in exposure to ETS. One standard deviation of exposure to ETS in the UKB is 5.23 hours in a typical week (mean = 0.97). To aid interpretability, we transform our results into units of hours exposed per typical week, in the amino text.

Sensitivity and additional analyses for the meta-analysis

Heterogeneity was assessed using the $I^2$ and $\tau^2$ statistics, and we used the random-effects model as a secondary estimator. We also used a leave-one-out sensitivity analysis to explore the potential effect of, and robustness to, outliers in the primary analysis. Finally, we control for multiple testing across the 6 outcome using the Benjamini and Hochberg correction (47).

Certainty assessment for the meta-analysis

Uncertainty was evaluated for each outcome using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach for the primary estimate (i.e., fixed-effect meta-analysis among low risk of bias estimation approaches) (48,49). Since our primary estimates are from low risk of bias natural experiments, we start all meta-analyses at high certainty and then downgrade them where appropriate.
Results

Risk of bias in MR estimation approaches

The risk of bias judgments with supporting evidence for each estimation approach and outcome are presented in Supplementary Tables S2a to S2f. Except for cardiovascular disease, which had 3, all outcomes had 5 estimation approaches at low risk of bias. Of these, two were not independent (“Exposure to ETS at home” and “Exposure to ETS outside of the home”). “Paternal smoking on index individual’s outcome” was always at high risk for bias due to extremely conditionally weak instruments. Here we visually summarise the conclusions in Table 2 and Supplementary Table S3.

Results of synthesis

The forest plots for the syntheses of results are presented in Figure 2 with units scaled to the effect of a genetically predicted standard deviation increase exposure to ETS on the log odds of the outcome. Scaling these results by the standard deviation of self-reported exposure to other people’s smoking in the UKB (5.23 hours per typical week), we found in our primary meta-analyses of low risk of bias MR estimation approaches evidence that exposure to ETS results in an increase in the odds of developing lung cancer (odds for every hour exposed in a typical week = 1.41 [95% CI: 1.23 to 1.61, \(I^2 = 0\%\)] and chronic obstructive pulmonary disease [COPD] (odds for every hour exposed a typical week = 1.11 [95% CI: 1.07 to 1.16, \(I^2 = 92\%\)]. Since the prevalence of both lung cancer and COPD is less than 2% in the UKB, the odds ratio estimated above should approximate the risk ratio (50). No association was detected for other outcomes (cardiovascular disease = 1.04 [95% CI: 0.97 to 1.12, \(I^2 = 0\%\)], stroke 1.01 [95% CI: 0.95 to 1.07, \(I^2 = 0\%\)], hypertension = 0.99 [95% CI: 0.92 to 1.06, \(I^2 = 0\%\)], depression = 0.96 [95% CI: 0.81 to 1.03, \(I^2 = 0\%\)]. Further details, including participant flow diagrams are presented in the Supplementary Results.

Results of sensitivity and additional analyses

The full results of the sensitivity analyses for the MR and meta-analyses can be found in the Supplementary Results. In brief, our positive control analysis found evidence of an association
between first-hand smoking and all the outcomes (Supplementary Table S4.1a-f). The leave-one-out analysis indicated that our meta-analyses were broadly robust to outliers (Supplementary Figure S1), but the ‘Maternal smoking on paternal outcomes’ approach did appear to inflate the lung cancer estimates. The crude MR analyses, which did not adjust for the outcome individual’s smoking, had highly inflated beta values when compared to our primary estimates (Supplementary Table S4.4a-f), indicating that the use of multivariable gene-by-environment MR was required to remove bias.

**Certainty of evidence in meta-analysis**

The GRADE evaluation of the certainty of evidence provided by our primary meta-analyses can be found in Supplementary Table S5. We downgrade lung cancer for having a potential outlier, COPD for residual heterogeneity, and depression for potential low power all to moderate certainty. All other outcomes remained at high certainty.
Discussion

Interpretation

In this study, we leveraged within family genetic data in a MR framework to investigate the long-term effects of environmental tobacco smoke (ETS) on six outcomes (lung cancer, COPD, stroke, cardiovascular disease, hypertension, and depression). After accounting for multiple testing, we find evidence of a dose-response effect between second-hand smoking and both lung cancer and COPD. According to the GRADE approach, these should be interpreted with moderate certainty. However, we did not find evidence of an association with cardiovascular disease, stroke, hypertension, or depression.

Our study has several methodological innovations. Firstly, it is the first MR study that we are aware of that sought to explore social/network effects. By leveraging a novel multivariable gene-by-environment MR design, we have created a plausible study design which allows us to explore the causal effect that one relative has on another’s outcomes. Secondly, we designed a bespoke risk of bias tool for our MR estimation approaches to reduce the amount of bias in our meta-analysed effect estimates. Thirdly, our integration of meta-analysis as a core part of our MR study design is unique. MR estimation approaches are generally less powerful than a traditional observational analysis. We were able to mitigate this issue by conducting many independent analyses which we meta-analysed. In addition, since many estimation approaches have independent potential confounders (Table 1) which we attempted to control for using multivariable MR, our meta-analysis is likely to be more robust to residual confounding than if we had conducted a single larger MR study (51).

Our results imply a large effect of ETS on lung-cancer and COPD. A careful interpretation of these estimates is required. Because people are randomised to (not) inherit genetic variants at conception, MR is generally described as estimating the ‘lifetime’ effect of an exposure (52). A literal interpretation of these estimates therefore assumes that participants exposure to ETS at recruitment to the UKB is representative of their typical exposure cross their entire life. However, exposure to ETS is likely to vary across the life course. For example, as children they may have been exposed to less ETS than they might be at the workplace. Although methods are being developed to address these issues (53), they remain controversial (54). Time-varying exposures do not bias MR estimates.
interpreted as the effects of *genetically predicted* levels of the exposure (55) though they may complicate the interpretability of estimates.

One can alternatively treat MR as a test of the null hypothesis (56). Because these do not necessarily require interpretable units, this approach circumvents the above issue and is generally more methodologically robust (26). Specifically, MR remains a valid test of the null hypothesis even when there are time-varying exposures (52). It has been similarly argued that meta-analyses of randomised controlled trial should be used to "determine whether or not some type of treatment - tested in a wide range of trials - produces any effect" , rather than “provide exact quantitative estimates” (57). Thus, while our point estimates may not be intuitively interpretable, our study should still provide valid tests of the existence of dose-response relationships. This approach requires having sufficient power to detect effects, which can be difficult to determine in a 2SMR setting, and especially when using MVMR. Because of this we have instead used first-hand smoking as a positive control.

*Comparison with existing research*

Our results lend further credence to the measures introduced by the UK government to reduce or ban smoking in public places by providing support for causal harmful effect of ETS (58). These measures were controversial when introduced (59). In certain respects, our results are consistent with previous studies. Prior studies support an association between ETS and lung cancer (10,60). For example, a weighted average of the two studies included in the study by Hori and colleagues, provides a comparable dose-response effect to the one we observed here (60). Moreover, Dietrich and colleagues observe a null effect ETS exposure on blood pressure (61). Indeed, the Global Burden of Disease estimated that ETS accounts for around 1.3 million deaths annually, which makes it the 13th leading non-aggregated cause of death (62).

However, it has been claimed that previous estimates of ETS exposure and its attributable risk are either spurious or overestimated (63,64), and in certain respects our findings support these claims. Previous meta-analyses of observational studies comparing people’s reported ETS exposure to outcome risk, found a significant dose response relationship between time exposed to ETS and risk of
depression, stroke, and cardiovascular disease (12,65–70). Barnoya and colleagues, for example, report a similar per unit time effect on CVD related phenotypes for first-hand smoking and exposure to ETS (71). Previous meta-analyses also found evidence of a dose-response effect for COPD (72,73), but estimates of this effect are much greater in magnitude than a naive interpretation of those observed here (73,74). It is possible that the absence of an association with CVD observed in the present study reflects a greater robustness to residual confounding, selection bias, and reverse causation. To this extent, our findings may support claims that previous observational estimates of ETS exposure and its attributable risk are overestimated (63,64). However, none of these studies are policy evaluations, and direct comparisons with these previous estimates of the effect of ETS are difficult due to the common practice of treating ETS exposure as a dichotomous variable (9,75–77). Moreover, given the difficulty in performing reliable power calculations for multivariable-MR studies, we cannot exclude the possibility of an effect of environmental tobacco smoking on stroke, cardiovascular disease, hypertension, and depression, albeit of a magnitude much smaller than that of first-hand smoking.

Strengths and limitations

The use of MR provides our analysis with several strengths. Firstly, because genes are fixed at conception, genetically predicted smoking cannot be impacted by the amount of smoking in one’s environment. Thus, our MR estimates should be more robust to reverse causation than an observational study. Secondly, by meta-analysing independent and low risk of bias MR estimation approaches, we analyse between 78,000 to 247,000 cases, and generally over a million controls, for each of our outcomes. We therefore provide large-scale evidence, robust to the influence of reverse causation and confounding, on an important question that could not feasibly or indeed ethically be investigated using a randomised clinical trial. Thirdly, even though parental smoking variants are likely to be pleiotropic, and therefore cause other parental phenotypes, it is unclear which parental phenotypes other than smoking could cause COPD and lung cancer in their children. We believe that this, and the consistency of pleiotropy robust estimators used in our analysis imply that residual pleiotropy is unlikely to be a major source of bias.
This study also has certain limitations. The first limitation pertains to the generalisability of our results. Since our study mostly uses data from the UKB, which has a relatively middle to older aged British population (mean date of birth = 1952), our effect estimates may not be applicable to younger generations or non-European populations. However, we believe that it is reasonable to expect the harmful effect of ETS exposure on the risk of lung cancer and COPD to generalise across age, ethnicity, and sex (57). Second, the use of summary-level data precludes exploring the presence nonlinearities or of effect modifiers, e.g., sex-specific effects, time-varying effects, and whether the strength of effect differs by duration of exposure to each relative (and thus if co-habiting relatives exert a stronger effect than non-cohabiting ones). However, MR should still produce a valid estimate of the average causal effect (78). In the supplementary discussion we discuss some of the supplementary results and why any theoretical limitations due to measurement error and sample overlap are unlikely to result in a meaningful bias to our result.

Conclusion

We meta-analysed a series of MR estimation approaches to estimate the causal effect of ETS exposure on lung cancer, COPD, stroke, cardiovascular disease, hypertension, and depression. Our results support the existence of a causal effect between ETS and both lung cancer and COPD. This strengthens the evidence supporting the use of public-health interventions to reduce people’s exposure to ETS.
Other information

Funding

BW is funded by an Economic and Social Research Council (ESRC) South West Doctoral Training Partnership (SWDTP) 1+3 PhD Studentship Award (ES/P000630/1). BW, HMS and MRM are all members of the Medical Research Council (MRC) Integrative Epidemiology Unit at the University of Bristol (MC_UU_00011/7).

Author contributions

BW and designed and implemented the study under the supervision of DG, HMS, and MRM. All authors contributed to the writing of the manuscript.

Data and data sharing

The data that was used in the primary analyses of this study is publicly available from the MRC-IEU OpenGWAS platform. The R code used in the manuscript, and GWASs created for this project, are available from https://doi.org/10.17605/OSF.IO/BKYXT. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol (http://www.bris.ac.uk/acrc/).

This project was conducted using UK Biobank application no. 15825. UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and the Northwest Regional Development Agency. It has also had funding from the Welsh Government, British Heart Foundation, Cancer Research UK and Diabetes UK. UK Biobank is supported by the National Health Service (NHS). UK Biobank is open to bona fide researchers anywhere in the world.

Ethics Approval statement

UKB received ethics approval from the North West Multi-Centre Research Ethics Committee (REC reference 11/NW/0382). All participants provided written informed consent to participate in the study. Data from the UKB are fully anonymised.
Conflicts of interests

DG is employed part-time by Novo Nordisk. All other authors declare no conflicts of interest.
References


30. Kimberley Burrows PH. Genome-wide Association Study of Cancer Risk in UK Biobank [Internet]. data.bris. 2021 [cited 2022 Jun 28]. Available from: https://data.bris.ac.uk/data/dataset/aed0u12w0ede20olb0m77p4b9


51. IEU Public Lecture: Paul Rosenbaum [Internet]. 2022 [cited 2022 Oct 5]. Available from: https://www.youtube.com/watch?v=m7xIRNNrsrC


Figure 1. Diagram illustrating randomisation to non-inherited.

A and B represent two alleles that robustly predict first-hand smoking and are inherited by both parents and directly affect parental phenotype. Only A forms part of the offspring genotype and has a direct effect on the offspring’s phenotype. The offspring randomly inherits A and randomly does not inherit B from parents. Both A and B exert an effect via the parental phenotype in the form of passive smoking. B therefore predicts second-hand smoking exposure in the offspring manner robust to the influences of confounding and reverse causality.
**Figure 2:** Forest plots for new MR estimation approaches, stratified by risk of bias.

### A) Lung cancer

<table>
<thead>
<tr>
<th>Design</th>
<th>log odds per SD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROB = Low</td>
<td>Maternal smoking on index individual</td>
</tr>
<tr>
<td></td>
<td>Maternal smoking on father</td>
</tr>
<tr>
<td></td>
<td>Paternal smoking on mother</td>
</tr>
<tr>
<td></td>
<td>Exposure to ETS</td>
</tr>
<tr>
<td></td>
<td>Common effect model</td>
</tr>
<tr>
<td></td>
<td>Random effects model</td>
</tr>
<tr>
<td>ROB = High</td>
<td>Paternal smoking on index individual</td>
</tr>
<tr>
<td></td>
<td>Paternal smoking on mother</td>
</tr>
<tr>
<td></td>
<td>Paternal smoking on father</td>
</tr>
</tbody>
</table>

### B) COPD

<table>
<thead>
<tr>
<th>Design</th>
<th>log odds per SD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROB = Low</td>
<td>Maternal smoking on index individual</td>
</tr>
<tr>
<td></td>
<td>Maternal smoking on father</td>
</tr>
<tr>
<td></td>
<td>Paternal smoking on mother</td>
</tr>
<tr>
<td></td>
<td>Exposure to ETS</td>
</tr>
<tr>
<td></td>
<td>Common effect model</td>
</tr>
<tr>
<td></td>
<td>Random effects model</td>
</tr>
<tr>
<td>ROB = High</td>
<td>Paternal smoking on index individual</td>
</tr>
</tbody>
</table>

### C) Cardiovascular disease

<table>
<thead>
<tr>
<th>Design</th>
<th>log odds per SD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROB = Low</td>
<td>Maternal smoking on index individual</td>
</tr>
<tr>
<td></td>
<td>Maternal smoking on father</td>
</tr>
<tr>
<td></td>
<td>Paternal smoking on mother</td>
</tr>
<tr>
<td></td>
<td>Exposure to ETS outside of the home</td>
</tr>
<tr>
<td></td>
<td>Common effect model</td>
</tr>
<tr>
<td></td>
<td>Random effects model</td>
</tr>
<tr>
<td>ROB = High</td>
<td>Paternal smoking on index individual</td>
</tr>
</tbody>
</table>

### D) Depression

<table>
<thead>
<tr>
<th>Design</th>
<th>log odds per SD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROB = Low</td>
<td>Maternal smoking on index individual</td>
</tr>
<tr>
<td></td>
<td>Maternal smoking on father</td>
</tr>
<tr>
<td></td>
<td>Paternal smoking on mother</td>
</tr>
<tr>
<td></td>
<td>Exposure to ETS</td>
</tr>
<tr>
<td></td>
<td>Common effect model</td>
</tr>
<tr>
<td></td>
<td>Random effects model</td>
</tr>
<tr>
<td>ROB = High</td>
<td>Paternal smoking on index individual</td>
</tr>
</tbody>
</table>

### E) Hypertension

<table>
<thead>
<tr>
<th>Design</th>
<th>log odds per SD [95% CI]</th>
</tr>
</thead>
</table>

### F) Stroke

<table>
<thead>
<tr>
<th>Design</th>
<th>log odds per SD [95% CI]</th>
</tr>
</thead>
</table>

All results have units of the odds of the outcome per genetically predicted standard deviation exposure to smoking. COPD = Chronic obstructive pulmonary disease, ROB = Risk of bias.
Table 1: Description and Directed Acyclic Graph of the 6 MR estimation approaches conducted.

<table>
<thead>
<tr>
<th>MR Estimation Approach</th>
<th>Description of the approaches in the row</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Maternal smoking on index individual’s outcome</td>
<td>The parent’s genetic liability to smoke can affect the index’s outcomes via the parent’s phenotype or via the index’s inherited SNPs. We included the index’s liability to smoke as a covariate in an MVMR model with parent’s smoking as the primary exposure to account for this.</td>
</tr>
<tr>
<td>2) Paternal smoking on index individual’s outcome</td>
<td>Although parents should be unrelated to each other, they may select their spouse in part because of shared smoking status. This would cause a correlation between the paternal genetic liabilities to smoke. We thus include both parent’s liability to smoke in an MVMR model.</td>
</tr>
<tr>
<td>3) Maternal smoking on paternal outcomes</td>
<td>Genetic associations with ETS could occur via evocative or active gene-environment correlation (e.g., active smoking causing people nearby to socially smoke or to select friends who smoke). We ran univariate MR analyses with the index individual’s self-reported exposure to ETS inside and outside home as the exposures.</td>
</tr>
<tr>
<td>4) Paternal smoking on maternal outcomes</td>
<td></td>
</tr>
<tr>
<td>5) Exposure to ETS at home</td>
<td></td>
</tr>
<tr>
<td>6) Exposure to ETS outside of the home</td>
<td></td>
</tr>
</tbody>
</table>
Note: We ignore the presence of potential pleiotropic, and phenotypic confounding, effects in the DAGs to emphasise possible biases that could occur due to the genetic covariance between people. The names in brackets are used to refer to the estimation approaches from here on. ETS = environmental tobacco smoke; GRS = genetic risk.
Table 2: Risk of bias summary figure illustrating the overall judgement about each MR estimation approach.

<table>
<thead>
<tr>
<th>MR estimation approach</th>
<th>Overall risk of bias judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lung cancer</td>
</tr>
<tr>
<td>Maternal smoking on index individual’s outcome</td>
<td>![ ]</td>
</tr>
<tr>
<td>Paternal smoking on index individual’s outcome</td>
<td>![X]</td>
</tr>
<tr>
<td>Maternal smoking on paternal outcomes</td>
<td>![ ]</td>
</tr>
<tr>
<td>Paternal smoking on paternal outcomes</td>
<td>![ ]</td>
</tr>
<tr>
<td>Exposure to ETS at home</td>
<td>![ ]</td>
</tr>
<tr>
<td>Exposure to ETS outside of the home</td>
<td>![ ]</td>
</tr>
</tbody>
</table>

COPD = Chronic obstructive pulmonary disease.