Review of the use of Mendelian Randomization with Smoking and Cancer

Amisha Kumar, Varun Aysola, Fredrick Schumacher*, Sarah Markt* (co-senior authors)
Case Western Reserve University, Department of Population and Quantitative Health Sciences

Abstract

Mendelian Randomization (MR) is a method developed to measure the significance of a correlation between genetic variation and the modifiable exposure of disease. In this paper, we specifically focus on the use of Mendelian Randomization that illuminates risks associated with smoking and various cancers. By conducting a meta-analysis of prior epidemiological and Mendelian Randomization studies, we found a large difference in the overall risk ratio conclusions. We presume that Mendelian Randomization does not represent the effective risk of developing cancer as established epidemiological principles demonstrate that social determinants heavily influence the exposure risks for cancer.

Introduction

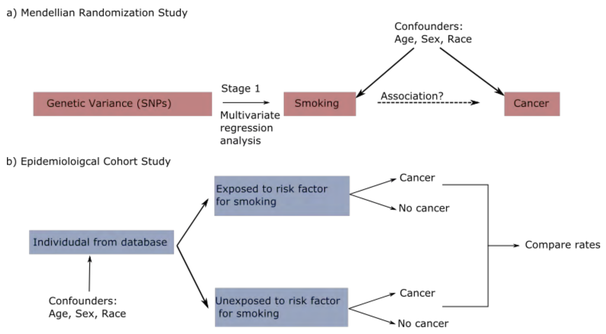


Figure 1. Mendelian Randomization vs. Epidemiological Study

MR is a newer research method that provides information about causal effects between exposure and a health outcome using genetic variation. MR is based on three main assumptions:

1) The genetic variants are associated with the risk factor.
2) No unmeasured confounders for the genetic variant and outcome association.
3) The genetic variant impacts outcome only through the risk facts.

In studying the association between smoking and cancer, MR uses genetic biomarkers for smoking as a proxy measure of smoking, while epidemiological studies use questionnaire based measures of smoking. Prior studies have indicated that individuals' sociodemographic, behavioural, and environmental factors considerably influence whether they choose to smoke, ultimately affecting their risk of developing cancer.

Our Hypothesis: Due to the differences in the methods, we hypothesized that the estimate produced by MR would be lower than those estimates obtained from epidemiological studies.

Methodology

- We found 19 published MR correlation estimates relating the genetic risk of smoking with the development of various cancers: oral, breast, colorectal, rectal, colon, breast, esophageal, bladder, cervical, head neck, ovarian, stomach, and lung cancer. Exposures were differentiated by having ever smoked, smoking regularly, smoking initiation, and lifetime smoking. In the literature we reviewed, the UK Biobank was the main source for the study population.
- Epidemiological studies were found using a collection of data from O'Keefe, et al (2018)
- Odds ratios were derived from both one-stage (obtained after a multivariate regression analysis) and two-stage MR (obtained after a regression analysis of the summary outcomes)

Analysis: We categorized the studies based on the type of exposure and cancer, and conducted meta-analyses utilizing the forest.meta-package in R. Similarly, we conducted a meta-analysis review of epidemiological studies using a collection of data from O'Keefe, et al (2018). We stratified the epidemiology studies by sex. We present both the fixed and random-effects estimates from the meta-analysis.

References

Bowden, J., & Holmes, M. V. (2019). Meta-analysis and Mendelian randomization: A review. *Research synthesis methods*, *10*(4), 486–496. https://doi.org/10.1002/jrsm.1346

Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. *JAMA*. 2017;318(19):1925–1926. doi:10.1001/jama.2017.17219

Fabio Barili, Alessandro Parolari, Pieter A Kappetein, Nick Freemantle, Statistical Primer: heterogeneity, random- or fixed-effects model analyses?, *Interactive CardioVascular and Thoracic Surgery*, Volume 27, Issue 3, September 2018, Pages 317–321, https://doi.org/10.1093/icvts/ivy163

National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. Preventing Tobacco Use Among Youth and Young Adults: A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2012. 4, Social, Environmental, Cognitive, and Genetic Influences on the Use of Tobacco Among Youth. Available from: https://www.ncbi.nlm.nih.gov/books/NBK99236/

O'Keeffe, L. M., Taylor, G., Huxley, R. R., Mitchell, P., Woodward, M., & Peters, S. (2018). Smoking as a risk factor for lung cancer in women and men: a systematic review and meta-analysis. *BMJ open*, *8*(10), e021611. https://doi.org/10.1136/bmjopen-2018-021611

Results

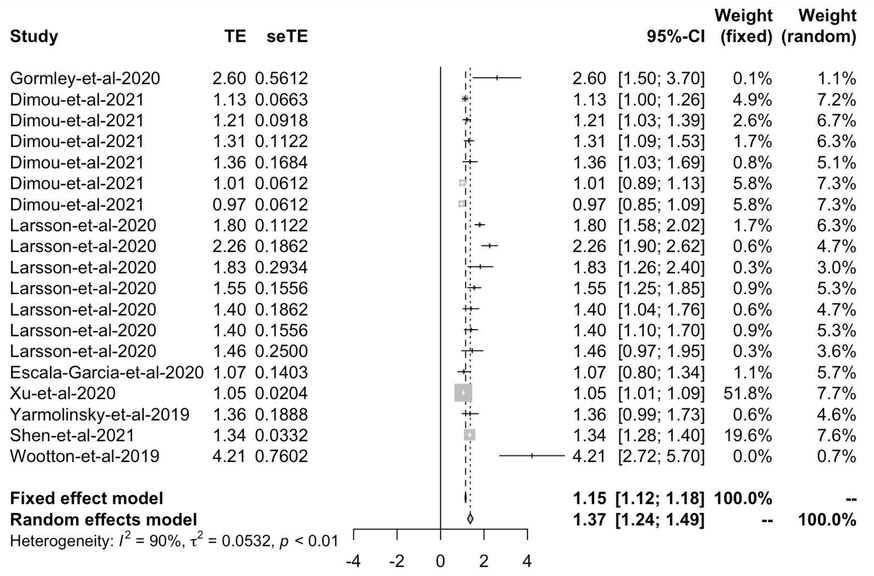
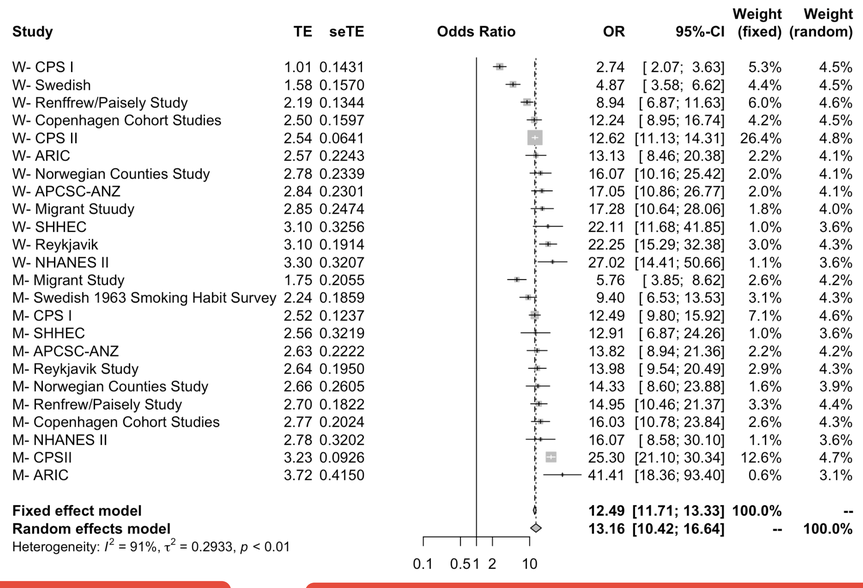


Figure 2. Meta analysis of Mendelian Randomization Odds Ratios

- Forest plot randomizes the weight of each study based on the final odds ratio (labelled as TE) and standard error (labelled as seTE).
- The odds ratios range from 0.97 to 4.21.
- Based on the random effects model built into the forest.meta package, the overall risk ratio for cancer is 1.37 for individuals that are genetically predisposed to smoking.

Figure 3. Meta analysis of Epidemiological Odds Ratios

- Estimates from each database are separated by sex, with 'M' representing male and 'F' representing female followed by the database name.
- The odds ratios range from 2.74 to 41.41.
- The overall risk ratio for cancer is 13.16 for individuals that are predisposed to smoking.
- This is 9.93 times higher than the risk presented by the MR methods.



Discussion

We found that MR produces a vastly different correlation estimate than epidemiological studies. By studying the differences between MR and epidemiological studies, we found that epidemiological studies often account for a broader range of confounding factors that could affect the risk of smoking. For this reason, we presume that in the case of smoking and cancer, epidemiological studies may produce a more effective risk measure.

Future Directions

1. Before assuming a significant association between genetic variance and exposure risk, one should explore established epidemiological correlations
2. There should be guidelines for when MR assumptions can be assumed
3. We hope to see similar reviews of the differences between MR and epidemiological odds ratios with varying exposures and outcomes.