Commentary

Leverage of genetic variants proxying smoking intensity to explore the broad health consequences of smoking

Dipender Gill[a,b,c,d,e,*]

[a] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom
[b] Centre for Pharmacology and Therapeutics, Department of Medicine, Hammersmith Campus, Imperial College London, London, United Kingdom
[c] Novo Nordisk Research Centre Oxford, Old Road Campus, Oxford, United Kingdom
[d] Clinical Pharmacology and Therapeutics Section, Institute of Medical and Biomedical Education and Institute for Infection and Immunity, St George's, University of London, London, United Kingdom
[e] Clinical Pharmacology Group, Pharmacy and Medicines Directorate, St George's University Hospitals NHS Foundation Trust, London, United Kingdom

ARTICLE INFO

The recent explosion in large-scale genetic association study data and genotyped biobanks offers an unprecedented opportunity to leverage natural genetic variation for inferring causal effects using the Mendelian randomization paradigm [1, 2]. In this approach, genetic variants are used to proxy modification of an exposure and study its effect on an outcome [1]. The random allocation of genetic variants means that such associations are relatively devoid of the confounding and reverse causation bias that can hinder causal inference in observational research [1]. The application of this approach to a range of clinical outcomes across the phenome allows for efficient investigation of the broad health implications of a genetically proxied exposure [3, 4].

In the paper by King and colleagues, the authors perform Mendelian randomization across the phenome (i.e. phenome-wide Mendelian randomization analysis) to investigate the broad clinical implications of smoking [5]. They find the expected associations of genetically predicted smoking intensity with respiratory, cardiovascular and cancer outcomes, and also identify more novel associations including links with acute renal failure and septicaemia [5]. In total, the authors generate genetic evidence for detrimental effects of smoking on 28 disease outcomes. The findings provide important insight, both adding to existing work, and also offering useful advances. By considering smoking intensity as the exposure, the findings support the notion that efforts to reduce the number of cigarettes smoked per day will likely still be of benefit where complete smoking cessation is not achievable. The association of higher genetically proxied smoking intensity with increased risk of a broad array of adverse clinical outcomes but no evidence of any beneficial effects adds unwavering supports to the detrimental effects of smoking on health.

The innovative methodology employed by King and colleagues builds on previous work using genetic variants related to smoking heaviness in smokers as genetic proxies for studying the effect of smoking intensity [6]. This further allows for a negative control sample in those that have never smoked. Through hypothesis-free untargeted analysis, the phenome-wide association study approach additionally allows for wider investigation than in previous studies, such as those considering cardiovascular outcomes specifically [7].

In their findings, King and colleagues make useful insight towards prioritising further research efforts [5]. Of note however, the Mendelian randomization approach has limitations and should not be used to infer the effect of a clinical or public health intervention that reduces smoking intensity. Importantly, the Mendelian randomization estimates may be biased by pleiotropic effects of the variants on the considered outcome through pathways unrelated to smoking intensity. Despite the best efforts of the authors [5], it is never possible to completely exclude the possibility of bias related to such pleiotropic effects [8]. For example, the genetic determinants of smoking are closely related to those of alcohol consumption, and as such it is unclear whether the identified Mendelian randomization association of smoking intensity with alcoholism represents evidence supporting causal effect or simply shared genetic aetiology. Other considerations are that Mendelian randomization estimates typically relate to the cumulative lifetime effect of varying an exposure (such as smoking intensity), while in practice an individual might differentially vary their smoking intensity throughout the life course. While such analyses may therefore be better served towards identifying causal relationships rather than estimating causal effects, there are two further caveats. Firstly, the consideration of a large number of outcomes forces a correction for multiple testing, which may in turn increase risk of false negative findings. Secondly, while such analyses may provide evidence to support a causal effect of smoking intensity on particular outcomes, they cannot in isolation offer mechanistic insight.

With the continued growth in availability of both summary-level genetic data and individual-level genetic data linked with electronic health care records [9], there is increasing opportunity to efficiently study the broad health implications of different exposures using the Mendelian randomization paradigm within a phenome-wide association study context [2]. Smoking is a leading cause of morbidity and mortality worldwide, and the current effort by King and colleagues represents an innovative approach for exploring the breadth of this [5]. Further work is now required to triangulate findings with other sources of evidence [10], provide mechanistic insight, and explore the effects of clinical and public health interventions that aim to reduce smoking.

Declaration of Competing Interest

DG is employed part-time by Novo Nordisk, outside of this work.

Funding

References

[1] Davey Smith G, Ebrahim S. Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003;32(1):1–22.

[2] Burgess S, Davey Smith G. How humans can contribute to Mendelian randomization analyses. Int J Epidemiol 2019;48(3):661–4.

[3] Gill D, Benyamin B, Moore LSP, Monori G, Zhou A, Koskeridis F, et al. Associations of genetically determined iron status across the phenome: a mendelian randomization study. PLoS Med 2019;16(6):e1002833.

[4] Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. Bioinformatics 2010;26(9):1205–10.

[5] King C, et al. Mendelian randomization case-control PheWAS in UK Biobank shows evidence of causality for smoking intensity in 28 distinct clinical conditions. EClinicalMedicine 2020;26:100488.

[6] Millard LAC, Davies NM, Timpson NJ, Tilling K, Flach PA, Smith GD. MR-PheWAS: hypothesis prioritization among potential causal effects of body mass index on many outcomes, using Mendelian randomization. Sci Rep 2015;5.

[7] Larsson SC, Mason AM, Back M, Klarin D, Damrauer SM, Million Veteran P, et al. Genetic predisposition to smoking in relation to 14 cardiovascular diseases. Eur Heart J 2020.

[8] Burgess S, Butterworth A, Malarstig A, Thompson SG. Use of Mendelian randomisation to assess potential benefit of clinical intervention. BMJ 2012;345.

[9] Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018;562(7726):203–9.

[10] Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. Int J Epidemiol 2016;45(6):1866–86.