Smoking, nicotine and COVID-19: Triangulation of methods and pre-registration are required for robust causal inference

Olga Perski, PhD[1], David Simons, MSc[2], Lion Shahab, PhD[1] & Jamie Brown, PhD[1]

[1] Department of Behavioural Science and Health, University College London, 1-19 Torrington Place, London WC1E 7HB

[2] Centre for Emerging, Endemic and Exotic Diseases, Royal Veterinary College

Corresponding author: Olga Perski, Department of Behavioural Science and Health, University College London, 1-19 Torrington Place, London WC1E 6BT, UK

E-mail: olga.perski@ucl.ac.uk; Phone: +44(0)20 7679 1258

Key words: nicotine; tobacco; COVID-19; SARS-CoV-2; triangulation; causal inference

A recent collection of articles in *Nicotine & Tobacco Research* (1), dedicated to the potential adverse or protective effects of smoking and/or nicotine use on COVID-19 outcomes, highlighted *a priori* biological and behavioural reasons as to why smoking status may be expected to influence COVID-19 infection, disease severity and mortality (2,3). For example, SARS-CoV-2 gains cell entry via the ACE-2 receptor, with elevated ACE-2 receptor levels observed in tobacco smokers – although reduced levels have also been reported (4). At the same time, nicotine is an agonist of the cholinergic anti-inflammatory pathway, which may protect against a hyperinflammatory response to the virus, potentially preventing severe disease/mortality in smokers/nicotine users (5). On the other hand, smoking causes respiratory and cardiovascular diseases that are associated with worse COVID-19 outcomes (6). However, it was concluded that we currently have "more questions than answers", with a need for additional, high-quality research (1–3).

More than a year into the COVID-19 pandemic and several months since the publication of the abovementioned article collection, answers are still few and far between. Our unadjusted meta-analyses of observational studies conducted during the COVID-19 pandemic indicate that current smokers appear to be at reduced risk of COVID-19 infection, and that there is inconclusive evidence of an increased risk of more severe disease among smokers who are infected, while former smokers appear at increased risk of severe disease and mortality from COVID-19 (7). **However, as most studies used observational designs – which may be subject to selection bias (e.g., oversampling of healthcare workers, who have lower/higher rates of smoking compared with the general population, depending on country (8)) and exposure misclassification (particularly when relying on routine electronic health records to ascertain smoking status (9)) –** it is uncertain whether these associations are causal. Here, we argue that triangulation across multiple, complementary methods, populations and time frames, in addition to the pre-registration **(i.e., posting to an independent registry, such as https://osf.io/) of hypotheses, study designs and analytic techniques (which is not yet the norm)** is required for improving our understanding of the potential causal relationship between smoking, nicotine and COVID-19 outcomes.

Although researchers are encouraged to pre-specify assumptions about causal relationships between variables of interest via, for example, directed acyclic graphs (DAGs), this is rarely done in practice (10). In the context of smoking, nicotine and COVID-19 outcomes, there are several plausible causal pathways to consider. Figure 1 presents two of these, the first from smoking/nicotine to SARS-CoV-2 infection (panel a) and the second, once infected, from smoking/nicotine to severe disease/death (panel b). Where research is focused on hypothesis-testing, it is important to use DAGs to demonstrate hypothesised causal relationships (10) and to pre-register these alongside the selected study design and analytic technique(s). Otherwise, we risk judging the plausibility or quality of the science based on the results **(as reflected by widespread publication (11) and citation bias (12))**, rather than on what matters most – the rigour of the methodology (13). As each method/data source has their own benefits and limitations, detecting a consistent effect of a similar direction, magnitude and biological gradient across multiple methods and populations would increase confidence that we have arrived at a true causal effect (14).

<FIGURE 1 HERE>

Many studies have examined the association between smoking status and COVID-19 disease severity/death in hospitalised populations, which may have led to spurious associations between current smoking and disease outcomes due to conditioning on a collider (15). As both smoking and COVID-19 infection may cause hospitalisation, conditioning on hospitalisation by design may produce a spurious association of smoking with COVID-19 disease severity. Even in high-quality observational studies with low levels of missing data on smoking status that use representative or random sampling, the risk of unmeasured confounding is high. In addition, SARS-CoV-2 is a novel infectious disease with time-varying dynamics – i.e., the rate of infection varies over time and across geographic regions, with dynamics dependent on the proportion of susceptible and infected individuals within a population and their contact rates. It is therefore important to examine whether exposure-disease associations observed in cross-sectional or short-term longitudinal studies remain stable across epidemic phases and geographic regions in studies with longer duration and greater measurement frequency. For example, the UK Office for National Statistics COVID-19 Infection Survey monitors SARS-CoV-2 seroprevalence in nationally representative samples on a monthly basis (16). Time series analysis of such data sources, examining whether the association of smoking status and SARS-CoV-2 infection remains stable over the course of the pandemic, would provide valuable information.

With the introduction of local or national pandemic regulations or restrictions, opportunities for natural experiments arise. For example, a public smoking ban (encompassing outdoor areas) was implemented in Galicia and the Canary Islands in August 2020 (17). In such scenarios, interrupted time series analyses could be conducted to examine whether the introduction of a given regulation led to a reduction in smoking and/or SARS-CoV-2 infection rates in the intervention locality (but not in control localities). The key benefit of natural experiments is that they, despite being observational, allow the examination of variation in the exposure (e.g. smoking or nicotine use) which arises from an external intervention (18). However, non-trivial variation in infection rates may occur across the intervention and control localities due to time-varying disease dynamics, which would be important to account for in the analyses.

Mendelian randomisation (MR) studies are likely to add another important piece to the puzzle. MR approaches harness the 'law of random assortment' (i.e., the random assignment of genes) to reduce confounding in observational epidemiological studies. Genetic markers – typically single nucleotide polymorphisms (SNPs) – that reliably affect a given exposure (e.g., smoking initiation or heaviness of smoking) are identified. The genetic marker must be independent of the outcome and of factors that may confound the exposure-outcome association. Scientists might therefore examine the association between two SNPs (rs16969968 and rs1051730) – located in the CHRNA5-A3-B4 nicotinic acetylcholine receptor gene cluster that are reliably associated with heaviness of smoking –

and COVID-19 outcomes. However, the assumptions underpinning the MR approach may be violated if, for example, the genetic marker affects the outcome via two or more smoking phenotypes (e.g. cigarettes per day and smoking topography), which may have complex relationships (19), or if horizontal pleiotropy is at play (e.g. the genetic marker affects the outcome through a pathway which bypasses the exposure). For example, it is plausible that the SNPs of interest may be associated with both heaviness of smoking and a hyperinflammatory syndrome in response to COVID-19 infection. However, according to a meta-analysis of genome-wide association studies (which are typically used to identify genetic instruments for MR studies), there is currently no evidence that the abovementioned SNPs act as genetic instruments for COVID-19 outcomes (http://app.covid19hg.org/). As sample sizes are small **(typically <500 cases/study)** and different criteria for disease severity may be used across studies, further work in this area is needed and it may be pertinent for MR analyses to focus first on the risk of infection (as opposed to disease severity/death).

Confidence in exposure-disease associations detected in observational studies, and to delineate the effect of nicotine from tobacco use, would be further increased if replicated in clinical trials. Several parallel-arm RCTs (nicotine patch vs. placebo patch) to examine the effect of nicotine on SARS-CoV-2 seroconversion (NCT04583410), disease progression (NCT04598594) and in-hospital mortality (NCT04598594) have been registered on ClinicalTrials.gov. **However,** results have not yet been reported. Another useful method could be to randomly allocate smokers to different cessation aids (e.g. behavioural support, e-cigarettes) and examine rates of laboratory-confirmed infection. Here, one may hypothesise that those allocated to behavioural (as opposed to nicotine-containing) support have increased risk of subsequent infection. Due to the element of randomisation, any time-varying disease dynamics observed during the study period should be balanced across groups, thus having a minimal impact on the estimated average treatment effect.

Vaccine trials (in which previously unexposed individuals are randomised to placebo or an active vaccine) could act as another useful source of information, providing that smoking status is measured. If, for example, a negative association between current smoking and infection were observed in the placebo, but not the vaccine group, results would replicate other observational studies and, depending on the comparative association observed in the vaccine group, may help elucidate where in the pathway smoking is having an effect, and whether any true protective effect persists in the face of a vaccine compared with placebo. Albeit controversial, it may also be useful to examine infection risk by smoking status with controlled human infection studies in which previously unexposed and unvaccinated young adults (who are unlikely to develop severe complications) are deliberately infected with SARS-CoV-2 virus (20). Such a study is currently underway in the UK, in which ~100 volunteers aged 18-30 years will be exposed to SARS-CoV-2 in a controlled setting (https://www.royalfree.nhs.uk/news-media/news/ethics-approval-for-human-challenge-study/).

In conclusion, observing a consistent pattern of results across studies with complementary methods, populations and time frames is required for arriving at a potential true effect of smoking/nicotine on COVID-19 outcomes. We hence strongly recommend nicotine and tobacco researchers to use DAGs to demonstrate hypothesised causal relationships and to pre-register these alongside selected study designs and analytic technique(s), to ensure that studies are judged based on the rigour of the methodology rather than their results.

Declaration of Interests

OP and DS have no conflicts of interest to declare. LS has received a research grant and honoraria for a talk and travel expenses from manufacturers of smoking cessation medications (Pfizer and Johnson & Johnson). JB has received unrestricted research funding to study smoking cessation from manufacturers of smoking cessation medications **(Pfizer and Johnson & Johnson).** All authors declare no financial links with tobacco companies or e-cigarette manufacturers or their representatives.

Acknowledgements

OP receives salary support from Cancer Research UK (C1417/A22962). OP, LS and JB are members of SPECTRUM, a UK Prevention Research Partnership Consortium (MR/S037519/1). UKPRP is an initiative funded by the UK Research and Innovation Councils, the Department of Health and Social Care (England) and the UK devolved administrations, and leading health research charities.

References

1. Edwards R, Munafò M. COVID-19 and Tobacco: More Questions Than Answers. Nicotine Tob Res. 2020;22(9):1644–5.

2. Berlin I, Thomas D, Faou A Le, Cornuz J. COVID-19 and Smoking. Nicotine Tob Res. 2020;1650–2.

3. Majmundar A, Allem J-P, Boley Cruz T, Unger JB. Public Health Concerns and Unsubstantiated Claims at the Intersection of Vaping and COVID-19. Nicotine Tob Res. 2020;1667–8.

4. Oakes JM, Fuchs RM, Gardner JD, Lazartigues E, Yue X. Nicotine and the renin-angiotensin system. Am J Physiol-Regul Integr Comp Physiol. 2018;315(5):R895–906.

5. Farsalinos K, Niaura R, Le Houezec J, Barbouni A, Tsatsakis A, Kouretas D, et al. Editorial: Nicotine and SARS-CoV-2: COVID-19 may be a disease of the nicotinic cholinergic system. Toxicol Rep. 2020;7:658–63.

6. Clift AK, Coupland CAC, Keogh RH, Diaz-Ordaz K, Williamson E, Harrison EM, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. BMJ. 2020;371:m3731.

7. Simons D, Shahab L, Brown J, Perski O. The association of smoking status with SARS-CoV-2 infection, hospitalization and mortality from COVID-19: a living rapid evidence review with Bayesian meta-analyses (version 7). Addiction. 2021;116(6):1319–68.

8. Nilan K, McKeever TM, McNeill A, Raw M, Murray RL. Prevalence of tobacco use in healthcare workers: A systematic review and meta-analysis. PLoS ONE. 2019 Jul 25;14(7):e0220168.

9. Polubriaginof F, Salmasian H, Albert DA, Vawdrey DK. Challenges with Collecting Smoking Status in Electronic Health Records. AMIA Annu Symp Proc. 2018 Apr 16;2017:1392–400.

10. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2020;1–13.

11. Munafò M, Neill J. Null is beautiful: On the importance of publishing null results. J Psychopharmacol (Oxf). 2016 Jul 1;30(7):585–585.

12. Duyx B, Urlings MJE, Swaen GMH, Bouter LM, Zeegers MP. Scientific citations favor positive results: a systematic review and meta-analysis. J Clin Epidemiol. 2017 Aug;88:92–101.

13. Nosek BA, Ebersole CR, Dehaven AC, Mellor DT. The preregistration revolution. Proc Natl Acad Sci U S A. 2018;115(11):2600–6.

14. Munaf MR, Davey Smith G. Repeating experiments is not enough. Nature. 2018;553:399–401.

15. Griffith GJ, Morris TT, Tudball MJ, Herbert A, Mancano G, Pike L, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. Nat Commun. 2020;11(5749):1–12.

16. Office for National Statistics. Coronavirus (COVID-19) Infection Survey pilot: England, 21 May 2020. 2020.

17. Coronavirus: Spanish regions ban smoking over Covid-19 risk [Internet]. BBC News. 2020. Available from: https://www.bbc.co.uk/news/world-europe-53763267

18. Vocht F De, Katikireddi SV, McQuire C, Tilling K, Hickman M, Craig P. Conceptualising natural and quasi experiments in public health. BMC Med Res Methodol. 2021;21(32):1–8.

19. VanderWeele TJ, Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in Mendelian randomization. Epidemiology. 2014;25(3):427–35.

20. Deming ME, Michael NL, Robb M, Cohen MS, Neuzil KM. Accelerating Development of SARS-CoV-2 Vaccines — The Role for Controlled Human Infection Models. N Engl J Med. 2020;e63:1–4.

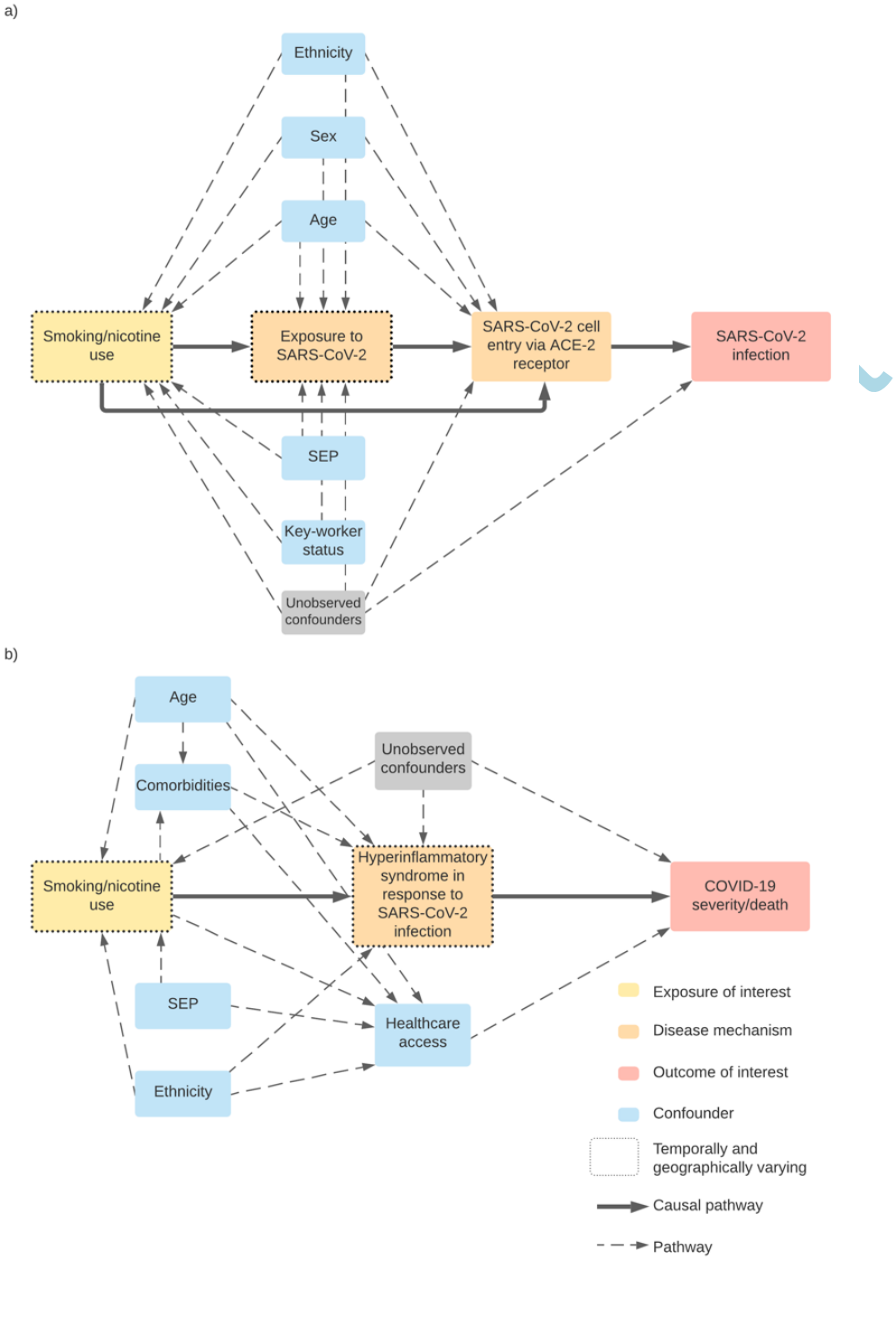

Figure 1. Proposed causal pathways a) from smoking/nicotine use to SARS-CoV-2 infection; and b) from SARS-CoV-2 infection to severe disease/death. **SEP = socioeconomic position.**