Dear Editor,

We are writing to comment on the paper “Association between the use of aspirin and risk of lung cancer: results from pooled cohorts and Mendelian randomization analyses” by Jiang et al. (Jiang et al. 2020). Findings from this Mendelian randomization (MR) study suggest that aspirin use decreases the incidence of lung cancer, specifically overall lung cancer and squamous cell carcinoma.

Observational studies have provided some evidence for the use of aspirin as a chemopreventive agent in lung cancer (relative risk 0.93, 95%CI 0.87–1.00) (Qiao et al. 2018). However, few randomized trials have been carried out to answer this question. Using genetic variation as a method of randomization to aspirin use and testing for association with specific cancers within an MR framework is, therefore, an attractive method to appraise causality before long and costly trials are conducted.

The authors used genetic variants from a genome-wide association study (GWAS) on aspirin use conducted by the Neale Lab (UK Biobank—Neale lab) to test whether aspirin intake was causally related to lung cancer incidence. However, there are some potential concerns with this method of instrumenting aspirin use.

One of the first major concerns with using instruments from a GWAS of drug use is disentangling the genetic variants for the drug from those for the drug’s indication. Using the MRC IEU OpenGWAS database (Elsworth et al. 2020), we found that the SNPs that predict aspirin use by this study (rs583104, rs2521501, rs10455872, rs73015016, rs7412, rs1831733, rs117733303) have all been shown to be associated with coronary artery disease (CAD) risk at least at genome-wide significance (van der Harst and Verweij 2018; Nikpay et al. 2015). Therefore, it may be that these SNPs increase the risk of CAD and that these individuals are being instructed to take aspirin as a preventative measure, thereby confounding the SNP association with aspirin use.

Many of the SNPs used to instrument aspirin use are also associated with a large number of other risk factors in the MRC IEU OpenGWAS database (Elsworth et al. 2020). This raises the potential for the violation of two of the MR assumptions: no confounding (independence assumption) and no horizontal pleiotropy (exclusion restriction assumption). Specifically, SNPs rs583104, rs10455872, rs73015016, rs7412 and rs117733303 have previously been associated with levels of low-density lipoprotein (LDL) cholesterol at genome-wide significance (Global Lipids Genetics Consortium, 2013; Prins et al. 2017), with increasing LDL-cholesterol levels leading to increased risk of coronary heart disease (Richardson et al. 2020). If cardiovascular risk factors are causally related to lung cancer incidence, as suggested by a previous MR study where increasing LDL-cholesterol levels was inversely associated with lung cancer incidence (OR 0.90, 95% CI 0.84–0.97 per SD of 38 mg/dl) (Carreras-Torres et al. 2017), then this may introduce confounding into the MR analysis.

Alternatively, associations between the genetic variants being used to instrument aspirin use and other risk factors could indicate violation of the exclusion restriction assumption of MR—namely, that the SNP is only affecting the outcome via the exposure of interest (Lawlor et al. 2008). When SNPs are also associated with other risk factors that may affect disease risk, this is termed horizontal pleiotropy (Burgess and Thompson 2013); however, we do acknowledge that the MR Egger regression was conducted and found little evidence of pleiotropy (Bowden et al. 2015). Furthermore, a weighted median approach was carried out and consistent results with the IVW were observed for overall lung cancer
Another threat to the validity of findings is the potential for selection bias. Since many of the SNPs may also be proxying liability to CAD, this may indicate survival bias, whereby individuals with higher risk of CAD die prematurely and, therefore, do not live long enough to be diagnosed with lung cancer. A frailty analysis could be conducted in this case to re-estimate the causal estimate in the presence of survival bias (Noyce et al. 2017).

A final concern with the results presented are the very large effect sizes (overall lung cancer, OR 0.042, 95% CI 0.003–0.564 and squamous cell lung cancer, OR 0.002, 95% CI 1.21 × 10^{-5} to 0.301) obtained from the MR analysis, which are much larger in magnitude than the corresponding observational estimates presented in the study (overall lung cancer relative risk (RR): 0.95, 95% CI 0.91–0.98, observational RR: 0.80, 95% CI 0.65–0.98, MR OR 0.69, 95% CI: 0.51–0.94; squamous cell lung cancer observational RR: 0.48, 95% CI 0.27–0.87, MR OR 0.48, 95% CI 0.27–0.87).

One alternative approach to study drug effects using MR is to identify SNPs that mimic a drug’s mechanism of action by investigating SNPs in the genes of the targeted protein (Gill et al. 2019). For example, statins inhibit the enzyme 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (HMGCR) resulting in reduced levels of LDL-cholesterol (Ference et al. 2015). Based on this understanding, SNPs that are in or around (within 100 kb) the HMGCR gene and that are associated with LDL-cholesterol have proven a useful method to instrument exposure to statins in MR studies (Ference et al. 2015, 2016). In the case of drugs such as aspirin that have multiple targets, it may be useful to conduct proteomic analysis to identify proteins targeted by the drug and instrument the effect of changes in mRNA/protein levels on cancer risk (Nounu et al. 2020). Instrumenting levels of mRNA/protein expression provides a continuous exposure, compared to aspirin use which is a binary variable and, therefore, results in complications when conducting and interpreting MR analyses (Burgess and Labrecque 2018).

Whilst we acknowledge that conducting MR studies of drug use and cancer incidence would provide much needed answers for clinical intervention, careful consideration in MR study design is needed when instrumenting drug use in MR to avoid the potential pitfalls highlighted. Methods and guidelines are now available for informing best practice in MR (Burgess et al. 2020; Davey Smith et al. 2019) so that appropriate inference can be made.

Acknowledgements The authors thank Dr Philip C. Haycock for his feedback considering the manuscript.
Author contributions AN provided the initial concept, performed the statistical analysis and drafted the manuscript. VW helped with reviewing and editing the manuscript. RCR helped with supervising as well as reviewing and editing the manuscript.

Funding This work was funded by Cancer Research UK (C18281/A19169) and the Medical Research Council Integrative Epidemiology Unit, which is supported by the Medical Research Council and the University of Bristol [MC_UU_00011/1 and MC_UU_00011/4].

Code availability The code to reproduce this analysis can be found at (https://github.com/an0435/aspirin_lung_cancer_MR/).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Availability of data and material Data are freely available on the online platform for MR Base (http://app.mrbase.org/) as well as through the TwoSampleMR R package (github.com/MRCIEU/TwoSampleMR).

References

Bowden J, Davey Smith G, Burgess S (2015) Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. Int J Epidemiol 44:512–525

Burgess S, Labrecque JA (2018) Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. Eur J Epidemiol 33:947–952

Burgess S, Smith GD, Davies NM, Dudbridge F, Gill D, Glynour MM, Hartwig FP, Holmes MV, Minelli C, Relton CL, Theodoratou E (2020) Guidelines for performing Mendelian randomization investigations. Wellcome Open Res 4:186

Burgess S, Thompson SG (2013) Use of allele scores as instrumental variables for Mendelian randomization. Int J Epidemiol 42:1134–1144

Carreras-Torres R, Johannson M, Haycock PC, Wade KH, Relton CL, Martin RM, Davey Smith G, Albanes D, Aldrich MC, Andrew A, Arnold SM, Bickeboller H, Bojesen SE, Brunström H, Manjer J, Brüske I, Caporaso NE, Chen C, Christiani DC, Christian WJ et al (2017) Obesity, metabolic factors and risk of different histological types of lung cancer: a Mendelian randomization study. PLoS ONE 12:e0177875

Davey Smith G, Davies NM, Dimou N, Egger M, Gallo V, Golub R, Higgins JP, Langenbein C, Loder EW, Richards JB, Richmond R, Skrivanova VW, Swanson S, Timpson NJ, Tybjærg-Hansen A, VanderWeele TJ, Woolf BA, YardomSK (2019) STROBE-MR: Guidelines for strengthening the reporting of the reporting of Mendelian randomization studies. PeerJ. https://peerj.com/preprints/27957

Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, Bates P, Palmer T, Haberland V, Davey Smith G, Zheng J, Haycock P, Gaunt TR, Hemani G (2020) The MRC IEU OpenGWAS data infrastructure. bioRxiv. https://doi.org/10.1101/2020.08.10.244293v1

Elsworth B, Mitchell R, Raistrick C, Paternoster L, Hemani G & Gaunt T (2019) MRC IEU UK Biobank GWAS pipeline, version 2, 18/01/2019. https://data.biobank.ac.uk/datasets/pmoat8cx0o5u5p26ynfaekeig/MRC IEU UK Biobank GWAS pipeline version 2.pdf. Accessed 16 Nov 2020

Ference BA, Majeed F, Penumetcha R, Flack JM, Brook RD (2015) Effect of naturally random allocation to lower density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or Both: A 2 × 2 factorial mendelian randomization study. J Am Coll Cardiol 65:1552–1561

Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, Voros S, Giugliano RP, Davey Smith G, Fazio S, Sabatine MS (2016) Variation in PCSK9 and HMGCR and risk of cardiovascular disease and diabetes. N Engl J Med 375:2144–2153

Gill D, Georgakis MK, Koskeridis F, Jiang L, Feng Q, Wei W-Q, Theodoratou E, Elliott P, Denny JC, Malick R, Evangelou E, Dehghan A, Dichgans M, Tzoulaki I (2019) Use of genetic variants related to antihypertensive drugs to inform on efficacy and side effects. Circulation 140:270–279

Global Lipids Genetics Consortium (2013) Discovery and refinement of loci associated with lipid levels. Nat Genet 45:1274–1283

van der Harst P, Verweij N (2018) Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ Res 122:433–443

Jiang Y, Su Z, Li C, Wang R, Wen Y, Liang H, He J, Liang W (2020) Association between use of aspirin and risk of lung cancer: results from pooled cohorts and mendelian randomization analyses. J Cancer Res Clin Oncol. https://doi.org/10.1007/s00432-020-03394-5

Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G (2008) Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 27:1133–1163

Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjommes A, Chasman DI, Chen S, Ford I et al (2015) A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47:1121–1130

Noua A, Greenhough A, Hoesom KJ, Richmond RC, Zheng J, Weinstein SJ, Albanes D, Baron JA, Hopper JL, Figueiredo JC, Newcomb PA, Lindor NM, Casey G, Platz EA, Marchand L Le, Ulrich CM, Li CI, Dujinbhoen FJ van, Gaur A, Campbell PT, et al (2020) A combined proteomics and Mendelian randomization approach to investigate the effects of aspirin-targeted proteins on colorectal cancer. https://cebip.aacrjournals.org/content/early/2020/12/1055-9965.EPI-20-1176

Noyce AJ, Kia DA, Hemani G, Nicolas A, Price TR, Pablo-Fernandez E De, Haycock PC, Lewis PA, Foltynie T, Smith GD, International Parkinson Disease Genomics Consortium, Schrag A, Lees AJ, Hardy J, Singleton A, Nalls MA, Pearce N, Lawlor DA, Wood NW (2017) Estimating the causal influence of body mass index on risk of Parkinson disease: a Mendelian randomisation study. PLoS Med. 14: e1002314

Prins BP, Kuchenbaeker KB, Bao Y, Smart M, Zabaneh D, Fatemifar G, Luan J, Wareham NJ, Scott RA, Perry JRB, Langenberg C, Benzeval M, Kumari M, Zeggini E (2017) Genome-wide association of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. Sci Rep 7:11008

Qiao Y, Yang T, Gan Y, Li W, Wang C, Gong Y, Lu Z (2018) Associations between aspirin use and the risk of cancers: a meta-analysis of observational studies. BioMed Cent Cancer 18:1–57

Richardsen TG, Sanderson E, Palmer TM, Ala-Korpela M, Ference BA, Davey Smith G, Holmes MV (2020) Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: a multivariable Mendelian randomization approach. PLOS Med. 17:e1003062

Sedgwick P (2014) Relative risks versus odds ratios. BMJ 348:g1407

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.