ESC European Society of Cardiology

European Journal of Preventive Cardiology (2024) **00**, 1–2
https://doi.org/10.1093/eurjpc/zwae141

LETTER TO THE EDITOR

Immortal time bias in the analysis of time-varying environmental exposures in the UK Biobank

Jacopo Vanoli [1,2]* and Antonio Gasparrini [2]

[1]School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan; and [2]Environment & Health Modelling (EHM) Lab, Department of Public Health Environments and Society, London School of Hygiene and Tropical Medicine, London, UK

Online publish-ahead-of-print 23 April 2024

We are writing to bring to your attention certain concerns regarding the article titled 'Co-exposure to multiple air pollutants, genetic susceptibility, and the risk of myocardial infarction onset: a cohort analysis of the UK Biobank participants' authored by Jiang *et al.*[1]

We believe this represents an important publication that enriches the epidemiological literature. However, after careful examination of the methodology and data presented in the aforementioned paper, it has come to our attention that there seems to be a significant flaw in the design of the study, specifically related to the exposure assignment, which could compromise the validity of the research findings. The authors state in the Methods section that 'We calculated the average concentration level of air pollution for each participant during the period from the four years preceding the enrollment to the dates when MI onset, death, or the end of follow-up, whichever came first'. In this sentence, it is implied that the exposure average assigned depends on the date of exit from the follow-up for a participant, either as a result of the event occurrence, death, or loss to follow-up.

We believe that the aforementioned method of exposure assignment can lead to wrong estimates of the hazard ratio in the Cox proportional hazard model through a mechanism previously described in the literature as 'immortal time bias'.[2] This occurs as subjects assigned as controls for a case in a given risk set, and who did not experience the event at the same time, are assigned exposure values belonging to future periods. In the presence of exposure trends common to the whole or part of the cohort, this can lead to an over- or underestimation of the risk. A practical example can clarify the issue. For instance, a case enrolled in 2008 and experiencing an outcome event in 2012 has an exposure value computed as the average within the period 2004–12 (as the exposure is computed from 4 years before enrolment). However, a corresponding control subject enrolled in 2007 assigned to this case can potentially be followed up until 2022 (the administrative end of follow-up), with an exposure average computed in the period 2003–22. This produces inconsistent exposure summaries between cases (who exit the study at the time of the event) and controls (who might well be in the study until the end of the follow-up). In the specific case of the study, it has been shown that in the last decade, exposure to air pollution has substantially decreased across the UK,[3] indicating that, in general, subjects experiencing the event in later years or never experiencing the event will be assigned, on average, lower exposure values compared to most of the cases, especially those experiencing the event in the early years. This systematic pattern of assignment likely leads to an overestimation of the health risks, as already suggested,[2] and, in general, it is not the correct method for modelling time-varying exposures in the Cox model.[4] These concerns raise doubts about the validity and reliability of the estimates reported in the article by Jiang *et al.*[1] and their conclusions.

Correct analysis of time-varying exposures in Cox proportional hazard models requires a pre-processing step. This involves splitting the follow-up period, which is specified on a time axis such as calendar years or age, into predefined intervals for each study subject. These intervals, also known as person-times (e.g. person-years), define the case–control risk sets for each event. Therefore, the corresponding time-dependent exposure is integrated with the person-times, and the data can proceed to the modelling stage. In the specification of the corresponding Cox time-to-event model, the main timescale chosen for the splitting would be automatically accounted for, while additional time variables can be directly adjusted for either with a linear or flexible functions to control for potential trends. We believe this is the method that should have been applied by Jiang *et al.*[1] in the presence of time-varying pollution data. To our knowledge, most of the published analyses on air pollution health risks using the UK Biobank have relied on time-constant exposure summaries, specifically assigned for specific years at the beginning of the follow-up. While this method can lead to a less precise exposure assignment compared to a time-varying exposure,[5] it is however immune from the immortal time bias and does not require complex modelling extensions as those described above.

Funding

This work was supported by the Medical Research Council-UK (Grant ID: MR/Y003330/1).

Data availability

No new data were generated or analysed in support of this research.

* Corresponding author. Email: jacopo.vanoli@lshtm.ac.uk

References

1. Jiang Z, Zhang S, Gao T, Chen K, Liu Y, Liu Y, *et al.* Co-exposure to multiple air pollutants, genetic susceptibility, and the risk of myocardial infarction onset: a cohort analysis of the UK Biobank participants. *Eur J Prev Cardiol* 2023;**31**:zwad384.
2. Yadav K, Lewis RJ. Immortal time bias in observational studies. *JAMA* 2021;**325**: 686–687.
3. Emissions of air pollutants in the UK – particulate matter (PM10 and PM2.5) [Internet]. Available from: https://www.gov.uk/government/statistics/emissions-of-air-pollutants/emissions-of-air-pollutants-in-the-uk-particulate-matter-pm10-and-pm25
4. Therneau TM, Grambsch PM. Modeling survival data: extending the cox model. In: Dietz K Gail M Krickeberg K Samet J Tsiatis A (ed.), *Statistics for Biology and Health*. New York, NY: Springer; 2000. p68–76. http://link.springer.com/10.1007/978-1-4757-3294-8 (18 Mar 2024).
5. Putter H, van Houwelingen HC. Understanding landmarking and its relation with time-dependent cox regression. *Stat Biosci* 2017;**9**:489–503.