Author’s response to reviews

Title: The impact of the lookback period and definition of confirmatory events on the identification of incident cancer cases in administrative data

Authors:

Jonas Czwikla (czwikla@uni-bremen.de)
Kathrin Jobski (kathrin.jobski@uni-oldenburg.de)
Tania Schink (schink@leibniz-bips.de)

Version: 2 Date: 19 Jul 2017

Author’s response to reviews:

Manuscript “The impact of the lookback period and definition of confirmatory events on the identification of incident cancer cases in administrative data” by J. Czwikla, K. Jobski and T. Schink (BMRM-D-17-00141R1)

Dear Prof. Haneuse,

Thank you for giving us the opportunity to submit a second revision of our manuscript. We also thank you and Mr. Goldsbury very much for your constructive comments regarding our conclusions which were very helpful to further improve the manuscript. We hope that our revisions will be sufficient to make our manuscript acceptable for publication in BMC Medical Research Methodology.

The second revised version of the manuscript and our corresponding point-by-point response were reviewed and approved by all authors.

All authors declare that they have no conflict of interest.

Informed consent was not required by law, since the study was based on pseudonymous data.

Yours sincerely, on behalf of all authors,

Jonas Czwikla, M.A.

University of Bremen
Editor Comments:

The authors have done an excellent job in responding to my comments.

Inline with one of the comments from the reviewer, I would like the authors to clarify the message in the Conclusions. In particular, it isn't clear if the sentence starting "Further validations studies ..." is a suggestion for the broader field (perhaps in the hope of getting generalizable conclusions?) or if it is a suggestion specific to researchers who but are thinking about using administrative data. That is, are the authors saying "If you are thinking about using administrative data then it would be important to first get a handle on the impact of lookback and confirmation periods in your data?".

Authors’ response: Thank you for this comment. We think that it is important for validation studies to show how several choices of the lookback and confirmation periods and the definition of confirmatory events impact the results, e.g. whether the predictive values differ considerably. Furthermore, we believe that administrative data can be used to estimate cancer incidences, considering the appropriate caveats as the reviewer David Goldsberry wrote. In these studies, it is important to examine how several choices of the lookback and confirmation periods and the definition of confirmatory events impact the results, i.e. to assess how robust the results are to these choices. We rephrased the abstract and the “Conclusions” section accordingly as:

Abstract (page 2, lines 47-50):

“Further validation studies as well as studies using administrative data to estimate cancer incidences should consider several choices of the lookback and confirmation periods and the definition of confirmatory events to show how these parameters impact the validity and robustness of their results.”

“Conclusions” section (page 15, lines 370-375):

“In the light of our findings, we advise to carefully consider which algorithm to use and to clearly describe how incident cases were identified. Further validation studies as well as studies using administrative data to estimate cancer incidences should consider several choices of the lookback and confirmation periods and the definition of confirmatory events to show how these parameters impact the validity and robustness of their results.”
Reviewer reports:

David Goldsbury (Reviewer 1): The authors have made several useful clarifications and informative additions, such as highlighting the many false positive cases introduced when not using confirmatory events.

I have two minor comments. The conclusions of the abstract and manuscript say that "Further validation studies should consider several choices... to show how these parameters impact the validity". There is no clear statement that the results can actually be used to estimate cancer incidence and rates using these data (with the appropriate caveats), and sounds like the results can only be used to guide further validation work. Perhaps the conclusions could say that "Future studies should consider several choices... to show how these parameters impact the results" or something similar.

Authors’ response: Thank you for pointing out that our conclusion is a bit too defensive. According to the comment from the editor and your suggestion we changed the abstract and the “Conclusions” section as follows:

Abstract (page 2, lines 47-50):

“Further validation studies as well as studies using administrative data to estimate cancer incidences should consider several choices of the lookback and confirmation periods and the definition of confirmatory events to show how these parameters impact the validity and robustness of their results.”

“Conclusions” section (page 15, lines 370-375):

“In the light of our findings, we advise to carefully consider which algorithm to use and to clearly describe how incident cases were identified. Further validation studies as well as studies using administrative data to estimate cancer incidences should consider several choices of the lookback and confirmation periods and the definition of confirmatory events to show how these parameters impact the validity and robustness of their results.”

The other point is that you could highlight the relatively small impact of exclusions by indicating the percent of people excluded (<1%?).

Authors’ response: Thank you for this comment. As suggested we added the percentage of exclusions to the “Results” section (page 8, lines 195-199):
“31,240 (0.5%) insured persons with at least 7 years of continuous insurance were not included in the study population because of missing or invalid information on sex (n = 18), year of birth (n = 472) or place of residence (n = 30,750). Further 18,884 (0.3%) insured persons were not included since they resided outside of Germany.”