The question whether treatment of common chronic diseases—such as oral anti-diabetic medication—may be beneficial in terms of reduced cancer risk and improved prognosis is, of course, an important issue. The identification of such effects may also help us understand the carcinogenic process and eventually contribute to the development of anti-cancer treatment.

There is, however, an inherent problem of interpreting effects when the outcome is influenced by a mixture of host factors and environmental exposures. A special challenge exists when exposure, such as medication of a chronic disease, may vary with time according to changing host factors such as clinical status, which by itself may influence the outcome. Even observational data from a health care system may offer good opportunities for research on this topic, but the quality of the study relies on the quality of the recorded data and the analytic approach.

The present manuscript is based on a sample from recorded data of more than 90 thousand diabetes mellitus patients in a US health care system, whereof some 500 were confirmed to have lung cancer according to the medical records.

Major Compulsory Revisions

1) A major weakness in this manuscript is the lack of description of time --- time of diagnosis, duration of anti-diabetes medication, length of follow-up, and also the quality and completeness of much of the basic information. No description was offered on the criteria used to confirm the lung cancer diagnosis, nor for the distinction between a primary lung tumour and a metastasis.

2) Methods section, Study design, P2, bottom line: The rationale for recording—and later analysing on—anti-diabetic medication in controls ***after the diagnosis of their respective case*** seems obscure. Except for Table 2, there was very little information as to what definition of anti-diabetic medication was used in the analyses.

3) Methods section, Study design, P2, L4 from below: Unknown smoking habits may introduce misclassification, and measures could be taken to avoid the problem, possibly by excluding the relatively small proportion of cases with no
smoking data (2%). Alternatively, analyses could, as a minimum, also be presented or reported with and without these cases included.

4) Results section, Association of Metformin and/or [...] , and Table 2: With matched sets of cases and controls I was surprised not to see the risk of lung cancer being estimated with conditional logistic regression. Despite the fact that smoking habits were a matching factor, residual confounding might distort the results in the present analyses. With conditional logistic regression one could e.g. include pack-years as a continuous variable in the model.

5) Discussion, P1: The apparently contradictive results suggesting reduced risk of lung cancer, but more severe disease stage and worse prognosis, both linked to the same exposure should be discussed more in depth.

6) Discussion, P4 starting “The associations identified in this study ...”: References 14 and 15 (Evans & al 2005; Bowker & al 2006) should be quoted in the background section, as they were cited in the review by Giovanucci & al, 2010 (reference number 9, in the Background section of the present article).

7) Discussion, last paragraph: The discussion of potential bias, misclassification, and links between clinical status, medication and risk factors could be improved. Residual confounding might be a problem, as the smoking variable is quite coarse, and some differences are already described (see comment above).

8) Discussion, last paragraph, L5–6 from bottom: The suggestion that the differences in BMI could be an effect of the cancer itself indicates that the time span for the study is rather narrow, and underlines the need to address better such parameters as time for start with oral anti-diabetic treatment, duration of treatment, length of observation, and the distance (or minimum distance) in time between recorded explanatory variables and diagnosis of lung cancer.

Minor Essential Revisions

1) Methods section, P Statistical Analyses, L4: “... after adjusting for significant factors.” Please clarify the meaning of “significant” (statistical significant or important)

2) Results section, P3 starting “All numbers expressed [...]. This paragraph seems to be footnote for table 2 and should be deleted.

3) Discussion, P2 starting “For an association to be considered ...” needs some more care. The text, I believe, hints to the advice of Sir Bradford Hill, but he preferred the word “guidelines” to criteria, thereby signalling less firmness. A “strong association” is not a prerequisite for a causal association.

4) Discussion, P3 starting “The association between the development ...” suggests that a relative risk of 1.5 is a strong association, but it could easily be argued that this is a weak association compared to the effect of smoking which easily reaches relative risks of 15–20 for the categories recorded in the study.
5) Discussion, P5 starting “The association reported in this study ...” contains a description of plausible biological mechanisms for an effect on cancer risk from anti-diabetic pharmaceuticals. In the present version, where quite some uncertainty remains in the epidemiology, this discussion seems slightly disturbing, as it does not really relate to the contributions from the study. It might work better in a revised version of the manuscript, or be relevant as background information.

6) Discussion, last paragraph, last sentence: In the present version, it could be argued that the study does not offer very much support to the hypothesis of effects on lung cancer risk and prognosis.

7) Table 1 should display the number of cases and controls and the heading should offer more information (diabetes mellitus).

Discretionary Revisions

1) Methods section, Statistical Analyses, L1: A typographical error, Wilcoxon is not with an “a”.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests.