Author’s response to reviews

Title: Diabetes and the Occurrence of Infection in Primary Care: A Matched Cohort Study

Authors:

Waseem Abu-Ashour (wmaa75@mun.ca)
Laurie Twells (ltwells@mun.ca)
James Valcour (james.valcour@med.mun.ca)
John-Michael Gamble (jm.gamble@uwaterloo.ca)

Version: 2 Date: 26 Oct 2017

Author’s response to reviews:

October 26, 2017

Dear Dr. Harris:

Thank you for the opportunity to respond to the comments generated from the review for our revised manuscript (INFD-D-16-01088-R1) entitled “Diabetes and the Occurrence of Infection in Primary Care: A Matched Cohort Study”. All changes to the manuscript are indicated in the marked version using track changes. Please find our responses to each comment below.

Kind Regards,

Dr. John-Michael Gamble
Associate Clinical Professor

Editors Comments:

The authors present an analysis of a match case-control to investigate the effect of diabetes on the occurrence of infection in primary care. The authors have matched the cases to controls in the following way “Controls were selected from patients without diabetes that were alive on the date their matched patient with diabetes entered the cohort”. The authors then state, “We decided not to match on other potential covariates such as age and sex due to limited availability of controls in certain strata and we were also interested explicit testing of effect modification.” and instead use adjusted odds in the regression models to account for potential confounding.
Comment 1: Whilst the methods selected for analysis of the cohort data are statistically sound, the described rationale does not make sense, and should be re-worded. Emphasis should be made on what the authors did (ie. a multivariable conditional case-control regression analysis with confounder adjusted odds ratios), followed by a rationale behind not using a more stringent matching study design. References should be included that justify the approach (such as Pearce 2015 - http://www.bmj.com/content/352/bmj.i969 and Rose and van der Laan 2009 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2827892/ ) as appropriate methods for analysis of case-control data with potential confounding remain a contentious issue and there isn’t really a best practice. More explanation should be given on how the adjusted odds ratios were estimated.

Response 1: As suggested we have reworded the methods section on page 7 to clearly describe the approach taken in our study. The revised text in the inclusion criteria reads: “We used a matched cohort design and adjusted for confounders using a standard multivariable regression approach. Given that matching does not necessarily remove bias and no standard practice exists for which variables to match [16,17], we decided to match on index year only and further adjust for potential confounders using a regression modeling approach. Specifically, we used a multivariable conditional regression model to estimate odds ratios for our primary and secondary outcomes for those with diabetes compared to those without. In addition, we were unable to match on other potential covariates such as age and sex due to limited availability of controls in certain strata. Moreover, we were also interested explicit testing of effect modification which we would not be able to do if age and sex were used as matching variables.”

Comment 2: Should the authors also consider adjusting for smoking in their analysis (as smoking is typically associated with respiratory infections and more smokers were reported in cases)?

Response 2: We did not adjust for smoking in our initial analysis as there is typically a large amount of missing data on smoking status. Indeed, within our study sample, there were 55% of patients with missing smoking status. Nonetheless, we have added smoking status as a categorical variable (indicator variable for missing) in order to adjust for smoking in our analysis. The addition can be found under “2.6 Statistical analysis” on Page 9. Importantly, adding this variable did not change our results. The adjustment of smoking with our other variables can be found in tables 4, 5 Appendix C & D.

Comment 3: Given that so many confounders have been adjusted for, could the authors consider an analysis that would help ascertain where the odds ratio adjustments largely come from (is it consistently driven from one confounder only)?

Response 3: As suggested, we have run several models with fewer selected covariates to explore the impact of various factors on the results. We also ran a backward stepwise regression model as an additional sensitivity analysis in which we forced age and sex as covariates and then removed covariates with a p-value of >0.1 from the model. Remaining covariates included age, sex, smoking status, acid-suppressing drugs, respiratory drugs, respiratory disease, and prior infections. This could be found under “Methods/Sensitivity Analysis” on page 10, under the “Results” section on page 12 and Appendix E.
Comment 4: Was the matching process carried out independently for each analysis (primary and secondary outcome), or were the same matched observations used for each analysis?

Response 4: We used the same matched observations for all analyses. This has been clarified in the methods section on page 9.

Comment 5: For each model, exploration of influential observations should also be considered, for example by using the dfbeta statistic. If there are several matched pairs that have a high dfbeta common to each analysis, it may be worthwhile removing them and/or repeating the randomisation process.

Response 5: We calculated dfbeta’s for a regular logistic model for our primary outcome and then dropped observations with high dfbeta’s (>3/sqrt(n)) and then reran both regular logistic and conditional logistic models. Results were consistent before and after dropping influential observations for both models. We have briefly mentioned this in our method section on page 10.