Article details: 2021-0335
Title: Evaluation of the accuracy of the PLCO(m2012) 6-year lung cancer risk prediction model in the CARTaGENE population-based cohort
Authors: Rodolphe Jantzen MD, Nicole Ezer MD MPH, Sophie Camilleri-Broët MD PhD, Martin C. Tammemägi PhD, Philippe Broët MD PhD

Reviewer 1: Dr. Kee Siang Lim
Institution: Kanazawa University
General comments (author response in bold)

The manuscript is well-written. It is a great effort to determine a good screening strategy for lung cancer prediction in Quebec. Here are my comments:
1. Please elaborate for in figure legend for Figure 1.
2. Please amend the Table 2 and Table S2 for better readability.
We thank the reviewer for the comments. We have added more explanation for the Figure 1.
As recommended, we have modified the format of the Tables to improve readability.

Reviewer 2: Dr. Alyson Mahar
Institution: University of Manitoba
General comments (author response in bold)

The authors have described an external validation of the PLCO(m2012) model to estimate lung cancer risk in Quebec population as well as a comparison of multiple screening strategies for lung cancer. This is an interesting although dense manuscript. My primary concerns are around the identification of a lung cancer diagnosis, the amount of follow-up, and the number of events available to complete this validation.

1. The authors cite Tonelli et al. for a validated algorithm to identify lung cancer. However, the Tonelli reference is not a primary external validation of this algorithm for the purpose of identifying lung cancer. The sensitivity and specificity of this outcome measure must be reported and examined with respect to the ability to complete an external validation- how could sensitivity and specificity of the algorithm impacted model performance evaluation? For example, if the algorithm itself misses 20% of cancers, how does that impact the evaluation of model performance? Or if the algorithm overestimates the number of cases in certain age groups, demographics? To me this makes it challenging to disentangle how well the model is actually performing.
We agree with the concerns of the reviewer. The specificities were not indicated as this parameter is generally >90% in studies of chronic diseases among the general population. Moreover, the exact value of the sensitivity was not indicated, while cancers were only separated in two groups: metastatic and non-metastatic cancers.
We agree that the article of Tonelli et al. was a proof of concept for using administrative data (claims) for ascertaining the presence of disease such as lung cancer. As of today, we do not have a more robust algorithm to define incident cancers with the Quebec administrative health databases. The CARTaGENE team is using this algorithm when giving cancer data to researchers. We modified the sentence in the manuscript (page 3) "As in the paper of Tonelli et al. we used administrative data to define…".

2. The authors are reporting on 6 year absolute risk, however 50% of the cohort has follow-up time less than this. The underlying model is a logistic regression which assumes equal follow-up on all participants. How are they accounting for this differential follow-up and for the competing risk of death? The would not have outcome data available, their 6 year risk of lung cancer, for all individuals in validation cohort. In addition, more detail on the underlying model for validation is needed.
We agree that having less than 50% of the cohort with less than 6 years follow-up can impact the results. However, the median is 5.9 years, which correspond to 5 years and 10.8 months. Moreover, the Q1 is 5.7 years, which correspond to 75% of the cohort with at least 5 years and 8.4 months follow-up.
Nevertheless, as there exists some censored individuals, the number of deaths is obtained by the sample size times the Kaplan-Meier estimate of the cumulative lung cancer risk at 6 year.
As mentioned in the Supplementary Material, the model is a logistic model.

3. There were only 205 events in the validation cohort. Did the authors complete a sample size calculation for this? I'm not sure there were enough events to do this reliably.
We thank the reviewer for its comment. As published research works with cohorts having sample size similar to CaG had previously shown calibration results, we did not calculate the minimum sample size in advance. Nevertheless, when using the Riley et al. O/E formula to calculate the minimal sample size (Minimum sample size for external validation of a clinical prediction model with a binary outcome. 2021. Statistics in Medicine), with Ø being equals to 1.5% and SE(ln(O/E)) to 0.1 (assuming O/E is 1, it corresponds to an expected 95% confidence interval of about 0.82 to 1.22) the sample size required is would be:
Thus, the required sample size is less than the total size of our cohort (N=11,652).
Overall, there is a significant amount of output that is not succinctly synthesized for the reader to understand the "so what" of this work or it's importance to policy or screening programs. What are the key take home messages? This would strengthen the manuscript.
We agree with the reviewer that the results were not succinctly synthesized. We have modified the Interpretation section of the manuscript to address this issue (pages 7-9).

Reviewer 3: James Bras
Institution: University of Manitoba
General comments (author response in bold)

Interesting article, and the selection of this cohort provides an interesting opportunity for theoretical comparison of lung cancer screening enrolment strategies on account of the extensive information collected.
In the interpretation for the first objective (6 year prediction accuracy in this cohort), the authors discuss that the model underestimated actual lung cancer incidence due to a previously demonstrated higher age-standardized incidence rate of lung cancer. They posit that this may be due to higher smoking exposure in Quebec than elsewhere in North America. The PLCOm2012 model theoretically should account for smoking exposure/intensity in risk estimation. I was hoping the authors could expand on the reasons why the model would not have sufficiently accounted for the smoking exposure in this population, or if there could be another factor that has influenced this result.

We agree with the reviewer that the PLCO model is taking into account the smoking status and intensity. However, the relative risks have less impact on the risk score than the incidence rate (model constant). Another possibility is the exposition of Quebecers with asbestos, since it is known that the combination of smoking and asbestos exposure greatly increases a person's risk of developing lung cancer. According to studies, the lung cancer risk attributable to asbestos exposure varies from approximately 0.5% to 15%. Finally, it could also be explained by the imputation of smoking related variables (~12%), but as mentioned in the Limitation section of the manuscript, the proportion of missing data was lower than in other large cohort studies.