Reviewer’s report

Title: Understanding the importance of key risk factors in predicting chronic bronchitic symptoms using a machine learning approach

Version: 0 Date: 29 Sep 2018

Reviewer: Alexandros Rigas

Reviewer’s report:

Review on the manuscript DMRM-D-18-00230 entitled "Understanding the importance of key risk factors in predicting chronic bronchitic symptoms using a machine learning approach"

Major Corrections

1) Are the data of the risk factors correlated? We know that the performance of the LRM is problematic when there is strong correlation among the risk factors. How does the GBM perform in this case?

2) In the tables which describe the results of the models the measures of accuracy, sensitivity and specificity should be included (the definitions of these measures are given in the reference below).

3) In order to find the optimal threshold from the ROC curve do you use a test set? If a test set is used to find the optimal threshold the results will biased. Instead a validation set from the training set should be used to calculate the optimal threshold.

4) From the group of children not having asthma the children that exhibited sneeze/runny nose symptoms 836 children(25.9%) or itchy eyes 473 children(14.7%) should be studied separately since this group probably represents the children with allergic rhinitis which also might cause BCP symptoms. The results of this analysis should also be compared with the rest of the children.

Minor Corrections

I) It will nice to compare the plots of the ROC curves for the GBM and for the LRM for all the prediction factors and for the top ten prediction factors( for both asthmatic and non-asthmatic children). This can be included in a supplementary file.

II) Most of the data are from questionnaires and that most of the times leads to underestimation of symptoms especially in asthmatic families or failure to identify symptoms like wheezing from non-asthmatic families. It will nice this issue to be discussed by the authors.
Small Corrections in the text

p.8 l.34 The phrase "integrate out them" to become "integrate them out".

p.12 l.48 Is Forno et al. or Forno and Celedon?

The references 12,15 and 17 are not complete.

Reference


This reference is relevant and should be included in the references of the manuscript.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:
1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal