Author’s response to reviews

Title: Incorporating repeated measurements into prediction models in the critical care setting: a framework, systematic review and meta-analysis

Authors:

Joost Plate (j.d.j.plate@umcutrecht.nl)
Rutger Van de Leur (rutgervandeleur@gmail.com)
Luke Leenen (l.p.h.leenen@umcutrecht.nl)
Falco Hietbrink (f.hietbrink@umcutrecht.nl)
Linda Peelen (l.m.peelen@umcutrecht.nl)
M.J.C. Eijkemans (M.J.C.Eijkemans@umcutrecht.nl)

Version: 1 Date: 13 Aug 2019

Author’s response to reviews:

Dear dr. Krüger and members of the Editorial Board,

Hereby, we are submitting the revised manuscript entitled “Incorporating repeated measurements into prediction models in the critical care setting: a framework, systematic review and meta-analysis “.

We would like to thank you for the consideration of our manuscript. We express our gratitude towards the reviewers, whose effort and insightful comments have been of great value for the improvement of this manuscript.

With the comments made, we hope to have sufficiently improved the manuscript to make it suitable for publication in your journal.

Sincerely,

Joost D.J. Plate, MD, PhD and R. van de Leur, BSc
Detailed point-by-point response to the reviewers’ comments are outlined below

Reviewer: 1

This systematic review and meta-analyses covers a topic of interest in medical research methods. Following Reviewer's reports and my own assessment of the manuscript, I am requesting the manuscript to be revised and re-submitted.

In preparing your revision, please address the comments raised by Reviewer 2. Also consider the following issues:

C1. A careful review of the framework as described in the text and, in particular, of its illustration in Figure 1 is required. I recommend seeking input from expert opinion from a biostatistician or computational statistics researcher familiar with the standard terminology used in the predictive modelling literature. Some definitions of ‘implicit vs implicit modelling’ in the text are equivalent to the more common classification of predictive models into parametric and non-parametric models. The use of the term ‘data-driven’ can also be misleading. In this manuscript, it appears to refer in particular to models in which the modelling of temporal sequences is fully data-driven. Not much attention has been paid to the type of outcome (continuous, binary, time-to-event), which will also influence the choice of model.

R1. We agree that the terms implicit versus explicit are not commonly used and have therefore changed this accordingly. Also, the term data-driven was removed from the framework to reduce any confusion caused with its overlapping use.

Further, although the type of outcome is indeed essential to the choice of a model, we believe that this does not hamper the application of the framework, as the framework shows the modelling steps to handle the repeatedly measured covariates (not per se the outcome). If, for example, features (mean, SD) are extracted and used in an explicit (parametric) model, then – arguably in a next step – the type of outcome should determine the eventual model. The eventual (choice of) model itself is not a direct aim of this framework or research.

This sentence has been added after the analyses of models used (results, 4th paragraph): “It should be noted that the choice of the eventual model is also dependent upon the type of outcome (e.g. binary, time to event or continuous).”

C2. Comparisons amongst models with repeated measured predictors would add value to the manuscript.

R2. We feel that, although this would indeed add value to the manuscript, this is very dependent upon the specific research question and domain, the data at hand and the specifications of the
models under comparison. Also, this may be prone to publication bias as authors will publish the results of well-performing novel algorithms, while novel algorithms which are not good enough are not published. In addition, only few included studies performed such comparisons.

C3. The manuscript contains some grammatically incorrect sentences and expressions. Therefore, I recommend careful editing during the review process.

R3. All authors have carefully re-read and edited the manuscript. We hope to have sufficiently improved this. If this is not the case, we may ask a native English speaker to edit the manuscript.

Reviewer: 2

In this manuscript, the authors proposed a framework and gave an extensive review of approaches used for repeated measurement in the critical care domain. The motivation of this study is clearly stated. And the framework is quite interesting.

C1. My major concern is the disconnect between the proposed framework and the meta-analyses of existed publications. From my perspective, the overview should provide some evidences that the framework, to some point, can help choose the right model better. The superiority can be shown via some case analyses based on the published studies.

R1. As per your suggestion, we added two paragraphs under results – application of the model. In this, we provided two examples of studies which use two-step and end-to-end modelling and show how the framework can be used and interpret to improve the approach to model repeated measurements.

C2. Also, the discussion section of this manuscript is too lengthy for me. The new understandings/findings are not clearly stated. A thorough consideration of the objective of this article is recommended.

R2. The discussion is shortened. The objectives of the article are mentioned in the first paragraph and a consideration can be found in paragraphs 4 and 5. This has been added to paragraph 5: “facilitates in choosing the approach and helps in the comparison between different approaches and models.”
C3. The authors didn't state the statistics very clear. For example, there are no detailed explanation of I square statistics and the c-statistics. Although these statistics are commonly-used, they all should be stated clearly with their usage and their formula for better readership.

R3. The formula for the I2 statistic is added (Methods - Meta-analysis - Synthesis of results).

As for the c-statistic, these sentences were added (Methods – Systematic Review – summary measures): “The c-statistic is similar to the area under the receiver operating curve, which is a graphical illustration of the false positive rate versus the true positive rate at each possible threshold. It can be interpreted as the probability that a random diseased subject is correctly rated with greater suspicion than a random non-diseased subject.ref”

C4. Some of the results, especially the section "Performance of analyses with and without repeated measurement", need better presentation. Here, I strongly suggest the author to read the article in the attachment "Understanding systematic reviews and meta-analysis" by A. K. Akobeng to get a better idea of how to summarize the results in forest plot. In addition, clarifications are need following "I^2, was 87.73%". It would be helpful if the authors can present the results more carefully.

R4. Although indeed the results could be summarized in a forest plot, we chose not to obtain a summary measure for these two reasons (methods, meta-analysis): “First, the predicted outcome differs, which means that the possible achievable c-statistic differs per study. Therefore, for some studied outcomes it will be difficult if not impossible to increase the c-statistic with the incorporation of repeated measurements. Second, the statistical heterogeneity between the studies, measured with the I2 was too high to warrant pooling of the results.”

To clarify this, the following sentence was added in the results section (‘performance of analyses’): “No summary measures were obtained due to reasons outlined in the methods section.”