Reviewer's report

Title: Using an onset-anchored Bayesian hierarchical model to improve predictions for amyotrophic lateral sclerosis disease progression

Version: 0  Date: 19 Sep 2017

Reviewer: Edward Bedrick

Reviewer's report:

Referee's report for "Using an onset-anchored Bayesian hierarchical model to improve predictions for amyotrophic lateral sclerosis disease progression" (BMRM-D-17-00238)

I enjoyed reading this paper, and feel that the authors have made a pretty good case justifying the use of the onset-anchor in an important prediction problem. I do have several concerns, as outlined below, mostly related to technical issues surrounding the cross-validation, choice of the point of prediction date and the use of the anchor point. Comments are given in sequential order of the paper (by page and number line)

Specific Comments

1. p 4 line 37: If rates of progression are important to estimate, then does developing a good model for predicting score at year 1 achieve that goal? Not necessarily (a point you later mention indirectly)

2. p5 line 49: It would be useful to provide % for the categories (which is common in a Table 1)

3. p6 line 5: you mention late in the paper that onset time is self reported - perhaps mention this earlier

4. p6 line 30 (and earlier): I didn't find any real rationale for why you wanted to predict at year 1 versus earlier. Perhaps elaborate on this - i.e. make it seem more than just a statistical problem. Also the first rationale for looking at 3 months of follow-up is unclear - why is this challenge relevant here? Just a bit clearer exposition would be useful.

5. Continuation of last point - you do a good job later in the paper discussing how your anchored model does better than the other models when the amount of subject data is less, or more than 3 months, with the advantages diminishing as a function of how many
months of data is available. I think this thread is really important. Another two related issues are (a) does the onset-anchor model do better than the other two choices when you are interested in predicting at 6 months vs one year? i.e. are your conclusions impacted by your choice of time at which the prediction is made? I didn't see this discussed. (b) In your comparison of the onset-anchor model with the other two, it is reasonable to ask whether the improvement is solely due to including one extra time point? Figure 6 suggests this is not the case, but the difference between using the onset-anchor model with 3 months of data versus the linear model with 4 months of data is much smaller than the difference when both are based on 3 months.

6. page 7 line 30: I do not understand what you are doing in your cross-validation (CV) assessment. Reading this gave me the impression that you were looking at the MSE between the actual and predicted year 1 score in the test data set, where I presumed the predicted was simply the mean of the predictive distribution. If so, there is no uncertainty, so your later discussion of the posterior of the MSE was lost on me. After reading further I decided that you were probably looking at the MSE in the predictive distribution of the future data, where the average squared error was centered at the actual observed value, and then this was averaged over all individuals in the test set. Could you elaborate? I know the journal is not very technical, but in a response to this question, could you write down the math so I know exactly what you did (I presume could untangle this from your code, but I would rather not have to do that). One reason for this need to elaborate, besides clarity, is that many readers would think of a prediction as a point prediction. Two other related issues: (a) It initially seemed that you only used one split of the data. Are the results robust to this - i.e. do similar conclusions follow when other splits are used? However, when I jump to page 11 it seems that you are doing multiple splits. I think this all needs to be clarified earlier to exactly what you are doing here and (b) given that the score distribution is relatively heavy tailed (you used a t_3 distribution), you might also consider another measure of error such as the mean absolute deviation which is less sensitive to extremes.

7. page 10 section 2.2.3. A standard Bayesian credo is to "model all uncertainty using probability" so I wonder what might happen if you consider the true day of onset to be the recorded day plus measurement error. Would doing so improve or degrade the performance of your onset-anchored model? I would be concerned if using (as if it could really be measured) the actual day of onset as the anchor led to a deterioration of performance relative to what you have seen. As you are doing things in a Bayesian framework, this analysis is feasible. Regardless, more discussion of this point is warranted.

8. page 10 line 54 onward: You repeatedly bring up the notion of bias. I understand the point you are trying to make, but unbiasedness is primarily a frequentist construct concerned with parameter estimation. For example, Gelman et al in their intro Bayes book largely discount this issue except in very large samples. Furthermore, even if you
considered a frequentist version of the model, and if you accept the notion that bias refers to parameter estimates, then random effects are random variables and subject specific trajectories are not functions of parameters so it is not clear what is meant by unbiasedness. A short elaboration on what you mean might help.

9. page 11 line 40. Do you find problematic how you collapsed longitudinal predictors where the collapsing is over months 1-3? I don't have much of a problem here given that you want to predict 1 year out. However, your model for the score at times prior to 3 months essentially depends on data that has yet to be collected. Perhaps make a note of this?

10. page 16 line 31. I presume you are referring to predictive intervals here - so are you concerned that the 95% intervals only contain the actual values 73% of the time? what does that suggest to you? is this due to the bias that you allude to - in addition as you note further down this page, the model tends to underestimate the disease progression - which I mentioned in comment (1). So if the goal was to estimate rate of disease progression, would the linear model do better in terms of estimating the rate? That is an important discussion point related to the limitations of the model.

11. Figures 4 and 5: I presume these are posterior predictive distributions rather than posterior distributions, as the parameters are averaged out. And these distributions are for the score at year 1 (x-axis label) versus for the estimated score at year 1? I guess this is semantics, but it relates to the fact that most readers will think of a prediction as a point prediction, and will distinguish prediction from estimation, which typically relates to parameters (but that is just me)

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

No

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.

no conflicts

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