Author’s response to reviews

Title: Feature Selection through Validation and Un-censoring of Endovascular repair Survival Data for Predicting the Risk of Re-intervention

Authors:

Omneya Attallah (o.attallah@aast.edu)
Xianghong Ma (x.ma@aston.ac.uk)
Alan Karthikesalingam (alankarthi@gmail.com)
Peter Holt (peteholt@btinternet.com)
Matthew Thompson (matt.thompson@nhs.net)
Rob Sayers (rs152@leicester.ac.uk)
Matthew Bown (mjb42@leicester.ac.uk)
Eddie Choke (tcchoke@hotmail.com)

Version: 5 Date: 12 Jun 2017

Author’s response to reviews:

Response Letter

Reviewer 2:

• The paper has been revised by a fluent English speaker.
• All writing bugs have been corrected.
• The whole sections 2.7 and 2.8, simulations studies, results of the simulation studies with their tables and figures have been moved into supplemental materials

Reviewer 3: Comment "I am doubtful whether the merits of the proposed approach presented in the paper are sufficient to justify employing such an elaborate method compared to much simpler methods such as the LASSO-penalised Cox model. Given that results in tables 4 and 5 indicate that differences are minor, I would suggest to put more emphasize on the simulation study. Only in this setting the true model is known and one can assess whether competing methods are able to recover the true underlying model better or worse than the proposed approach. Although the authors analysed synthetic data, Monte Carlo simulation is required to answer the question whether the proposed model truly performs better than any of the competing models: conclusions
derived from just a single comparison have very little power. Furthermore, since the true model is known, the coefficients obtained by the individual models can be directly compared to the true model to determine whether a model is suitable. Tables 4 and 5 only list the number of selected features, but do not indicate how these relate to the true model."

First of all I have been asked by one other reviewer to move the simulation study in the supplementary material

- We performed a Monte Carlo Simulation on synthetic data and updated the results in tables 1 and 2 in the supplementary material 2 to reflect this.
- We did not add the coefficients obtained by each model, as our proposed model using the ANN does not predict these coefficients. The output of the ANN is probabilities used to predict the risk of re-interventions. However, other models predict coefficients which are used to generate a risk score which is used to predict the risk of re-interventions. Therefore, it is not helpful to put these coefficients for comparison as they are not available for the ANN model.

Re comment (3): Use of sensitivity as performance measure.

- We removed the sensitivity metric from all tables (table 1 -3) and tables 1 and 2 in the supplementary material 2. I also deleted the paragraphs in the results section discussing this metric.

Re comment (3): Use of p-values.

- We removed the sentences in the results section which was arguing that models with lower with p-values performs better or preferred than others. We just stated that since the p-values are lower than 0.05, this t indicates that the risk groups are separated and distinguished significantly.

Reviewer 3 Comment p. 7, 159-160: The text indicates that the difference between the gradient boosting models in ref. 56,67 compared to ref. 58 are the use of survival trees. However, the main difference is the choice of loss function, which in both cases is optimised via gradient boosting.

- We corrected this sentences (highlighted in blue) and stated that difference is the choice of loss function, which in both cases optimized via gradient boosting.
Reviewer 3 Comment p. 7, 160-162: The statement that random survival forest and cited gradient boosting methods ignore censoring is false. All of these methods specifically address the censoring problem.

- We meant by the methods, those which are mentioned before in the previous paragraph which are 11,28,30,49-52 not the random survival forest and the gradient boosting. We corrected this in line 157.

Reviewer 3 Comment p. 20, 453-454: The authors refer to several papers showing to "have good CI even with high censoring". It is unclear what "good CI" means in this context. Due to censoring the true concordance index cannot be computed exactly and one has to estimate it. Harrell's or Uno's are two estimators. If "good" refers to obtaining an consistent estimate of the CI, the latter is preferred (see Uno et al, Stat Med vol. 30, 2011 for details).

- We removed the sentence have good CI even with high censoring in pa and We stated only some examples of the papers which used Harrel's CI to test the performance of their proposed method. The whole paragraph has been removed to the supplementary material 2.

Reviewer 3 Comment Section 4.2: The LASSO and SCAD penalised models introduce an additional hyper-parameter that controls the number of selected features and impacts models' performance. How was the optimal hyper-parameter determined?

- The optimal hyper parameter was determined using 10 fold cross validation. We added this to line 464.

Reviewer 3 Comment pp. 26-27, 608-632: This paragraph is merely a description of the referenced tables and figure. It is hard to understand what these numbers mean and how they relate to the question whether the proposed method is preferred or not.

- These are the life time tables which is an important technique of survival analysis that deals with time to event data. It can answer the question of the chance of re-experiencing an EVAR operation after experiencing it at the first time. It calculates the cumulative event-free probabilities (freedom from re-intervention). The table tells us, for the study period, how many patients did the re-intervention and how many still at risk after each year and during the duration of the study period. It calculates the cumulative probabilities of re-experiencing an EVAR operation even for patients who did not complete the follow up and left the study or patients who still haven’t experienced the operation yet. As the contributions that they
have made to the event probability are fully accounted for. These will help doctors determine the difference between the probabilities of freedom from aortic complications and the number of patients which are still at risk at each year of the study period. And is the prediction results are separable and differentiable or not.