Author’s response to reviews

Title: Prediction of Intracranial Findings on CT-scans by Alternative Modelling Techniques

Authors:

Tjeerd van der Ploeg (tvdploeg@quicknet.nl)
Marion Smits (marion.smits@erasmusmc.nl)
Diederik W Dippel (d.dippel@erasmusmc.nl)
Myriam Hunink (m.hunink@erasmusmc.nl)
Ewout W Steyerberg (e.steyerberg@erasmusmc.nl)

Version: 2 Date: 19 August 2011

Author’s response to reviews: see over
Dear Editor:

We would like to thank you for reviewing our paper, MS: 3156340865403205 - Prediction of Intracranial Findings on CT-scans by Alternative Modelling Techniques, and giving us the opportunity to revise it.

As you requested, we added a ‘competing interests’ section to our manuscript.

We will address the comments of the reviewers point-by-point. The changes made in the manuscript were marked in yellow, in the manuscript as well as in this cover letter.

REVIEWER 1: Ivan Ng

Remark:
Perhaps from a practical standpoint, the authors could discuss the significant salient features of the study and its impact on ascertaining variables of critical import which may be of value in streamlining current existing guidelines.

Response:
Thank you for this valuable advice. We inserted your suggestion into our discussion as follows:

The outcomes of this study suggest that the use of alternative modelling techniques may also have practical value in ascertaining variables of critical import and in streamlining current existing guidelines. Smits et al. used 14 variables for their modelling based on expert opinion and previous studies. We started out with these same 14 variables to be able to compare the model of Smits et al. with modelling based on alternative modelling techniques. However, the CHAID model only used 10 out of these 14 variables. The variables PTA, Change, EMV-13 and Seizure were not used, which suggests that these variables may be of lower importance for the outcome. However, the CHAID model performed poorly in comparison with logistic regression modelling. For most of the evaluated models, the variables of critical import were: Fracture skull (v69), Cause (cause3) and Age (age10). Based on our study, any guidelines should certainly contain these variables.

REVIEWER 2: Paolo Emilio Puddu

Remark 1:
There is a need to stress more and more that out of 14 variables more than 80% are not continuous (Table 2) and this may possibly favour logistic regression.

Response:
Thank you for your valuable remarks. We adapted our manuscript accordingly:

The superior performance of the logistic regression modelling might be explained by the high number of categorical variables (10 out of 14), which might favour logistic
regression modelling. The somewhat disappointing performance of tree models like CHAID and CART may be more realistic, because these models are well suited for dealing with categorical and continuous variables, although the latter are categorized by these models.

**Remark 2:**
The low performance of SVM (also considering the optimism, quite large, subtracted from the apparent AUC-value for each model on the original sample to obtain optimism-corrected estimates of model performance) surprises me a lot. What SVM software was used? There are other softwares with greater performance?

**Response:**
We share your surprise with respect to the optimism of the SVM model. In another current study we also found that SVM has a very high optimism compared to the other evaluated models. For the modelling we used Clementine software, which nowadays is part of SPSS (PASW) and is considered one of the best packages for modelling. However, it would be worthwhile to investigate if other software packages use other SVM-algorithms leading to other performances.

We inserted the following statement into our discussion:

For the modelling we used Clementine software. It would be worthwhile to investigate if other software packages use other SVM-algorithms lead to other performances.

**Remark 3:**
Logistic regression coefficients need be provided, especially considering that the potential for a formula (see Appendix 2) is a significant distinction of LR;

**Response:**
A table with the regression coefficients of the logistic regression model was inserted into the manuscript. (see Table 4 and Formula 1)

**Remark 4:**
Since the Authors stated “These findings may be seen as confirming the validity of the previously developed CHIP prediction rule” they need be aware of (and include a statement on) the fact that they have used an internal NOT an external validation of that rule.

**Response:**
Thank you for pointing this out. We inserted the following statement into our discussion:

However, it should be noted that these results are an internal validation of the developed CHIP-rule and that external validation is still required.

We hope that these adaptations are in line with your expectations and look forward to your response to the re-submission of our manuscript.

On behalf of all the authors,

Yours sincerely