Author's response to reviews

Title: External validation of a Cox prognostic model: principles and methods

Authors:

Patrick Royston (patrick.royston@gmail.com)
Douglas G Altman (doug.altman@cs.m.ox.ac.uk)

Version: 4 Date: 11 February 2013

Author's response to reviews: see over
11 February 2013

Dear Editor,

We note that we have now satisfied the criticisms of reviewers 1 and 3. Please find below a point by point response to reviewer 2’s most recent comments

Please consider the revised manuscript for publication.

Thank you
Patrick Royston and Douglas G. Altman.
Title: External validation of a Cox prognostic model: principles and methods

Version: 3

Date: 29 January 2013

Reviewer: Frank Harrell

Reviewer's report:
The authors have put a lot of good work into this manuscript. Some of the points of contention will always be legitimate differences of opinion. There are three remaining points about which the authors have not made any progress in convincing me of their validity/optimality.

We believe we have now dealt with these three points. See our responses below.

Major Compulsory Revisions
First and foremost is the continuing use of risk groups. In many ways, Figure 2 points out the possible fallacy in clinging to them. We see differences between red and black curves, but we see no information to confirm the necessary assumption that the distributions of risks within the three groups are identical when derivation and validation datasets are compared. This is necessary for one to be able to fully interpret these curves. A not-fully-conditional analysis is exposed to such mixture of distributions problems, not to mention resulting in a display and validation that is low resolution.

We thank the reviewer for this important clarification. We have responded by increasing the number of risk groups from 3 to 4. At the same time, to create the risk groups we have used Cox (1957) cutpoints, which are designed to minimize the loss of information due to grouping a continuous variable. To accommodate the change in the risk groups, the wording in various places throughout the manuscript has been altered accordingly, as of course have several figures and tables. In addition, to check the distribution of the PI between the derivation and validation datasets in each risk group, we have added a new figure (5). This shows that the corresponding cumulative distribution functions in each risk group are indeed very similar, as required for a correct interpretation of the calibration plots.

The continuous approach to which I was referring was the later versions of the val.surv and calibrate functions in the R rms package. The continuous true survival curves are estimated by adaptive linear spline hazard regression, which allows for the possibility of non-proportional hazards.

Thank you. We have added a reference to the R rms package.

Finally, the approximation of PI by a Gaussian distribution is rarely adequate. Attempts to use the Gaussian distribution should be abandoned or supplemented by a competing method that does not make this assumption, while alerting reasons to problems with the Gaussian assumption.

We realized that in fact our methods do not depend in any way on an assumption that the PI has a Gaussian distribution. We have deleted the relevant sentences. Putative Gaussian distribution of the PI is no longer mentioned in the manuscript.