Reviewer’s report

Title: Validation and updating of risk models based on multinomial logistic regression

Version: 0 Date: 02 May 2016

Reviewer: Kathleen Kerr

Reviewer's report:

This paper shows possible adaptations of updating methods for dichotomous risk models to multinomial risk models. The paper presents a few of observations about these methods and differences between binary and multinomial predictions. The focus of the paper is a single case study. While the paper contains some interesting bits, I found many of the conclusions were not justified.

I did my best to review the paper on substance and not on presentation, but there was a lot of sloppiness in the submitted manuscript that was sometime a small annoyance but often a detriment to understanding the paper. A superficial example is the repeated insertions of "Error! Reference source not found" in the text. A more substantive example is the following: the text on line 178 indicates that Method 5 will be described, but then the notation on line 182 refers to Method 3b. The text on line 191 indicates that Method 6 will be described, by the notation on line 194 refers to Method 5. The text on line 196 indicates Method 7 will be described, but the notation on line 201 refers to Method 6.

Similarly, I found the writing in this paper to be imprecise. Here are some examples. The text refers to "closed testing procedure" but I did not know what "closed" means here. On line 205, it is not clear whether "The procedure" refers to the recently suggested procedure or its extension. On line 207, the paper says that the "overall null hypothesis to be tested … is that the original model has the same fit as an updated alternative." Could the authors instead write the null hypothesis mathematically so we know what is meant by "same fit"?

Closed testing procedure: how should one choose alpha given that alpha will be used for up to 3 tests?

Line 381. "Calibration can often be strongly improved with simple intercept and/or slope adjustments." This claim is made without reference or evidence.

Lines 412-413. A claim is made here without reference or evidence. The first part of the sentence seems to merely say that large sample size leads to more power. The last clause of the sentence seems to pre-suppose that model updating is always good.

On line 287 the papers says logistic recalibration require k-1 intercepts and (k-1)^2 coefficients. This is consistent with what the authors have as Method 2 (lines 151-155) and Table 2.
However, on lines 359-360 the papers says that for k outcome categories there are k-1 calibration intercepts and k-1 calibration slopes, contradicting other material in the paper. It seems the authors make a distinction between (1) calibration intercepts and slopes, as measures of performance; and (2) recalibrated intercepts and slopes, as methods of updating ("logistic recalibration"). It seems that for (1) the authors only consider the "corresponding" LPs, whereas for (2) all LPs are used and the number of parameters to be estimated is then larger. The authors seem switch between the two ideas, confusing the reader.

Regarding logistic recalibration (Method 2, lines 151-155): it isn't obvious why this should be done using the "non-corresponding" linear predictors (I'm using the authors' terminology here). If the answer is in Reference 12, perhaps the authors could share the reason in this paper.

Minor comments

1. "less variables" should be "fewer variables"; "less parameters" should be "fewer parameters"
2. line 267, "were" should be "where"
3. line 277, "is" should be "of"
4. line 296, "visible" recalibration - please be more precise

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

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.

I declare that I have no competing interests.

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.