Reviewer's report

Title: A Comparative Analysis of Multi-Level Computer-Assisted Decision Making Systems for Traumatic Brain Injuries

Version: 2 Date: 1 May 2008

Reviewer: Rory Wolfe

Reviewer's report:

The authors have made a reasonable job of responding to the first set of review comments. However in making changes to the manuscript they have introduced more errors. Most glaring is the mixing up of the order of sections. The Methods section and Conclusion section are in the wrong places. This is sloppy.

Two major concerns still exist for me.

1) The new description of “Direct MLE” is inadequate. Do the authors mean that all possible variables were entered into the model together and a single likelihood-ratio test was performed? It simply isn’t clear.

2) The linearity assumption in logistic regression has not been explored. Rather, a single crude overall statistical test has been performed. This is insufficient.

There are so many minor errors littered throughout the manuscript that the following is merely a guide and not an exhaustive list:

Page 3: Eckstein [45] should be [12]

Page 3: “Fu [21] states that when more than two outcomes exist, a combination of the loglinear model and CART is more efficient” This may be true however it is no relevance to this paper in which three variables are analysed and all three are binary outcomes, i.e. two outcomes.

Page 3-4: Hasford reference is a very selective choice – many papers exist that disagree with the conclusion that CART is more successful than logistic regression alone.

Page 4: Final sentence of Introduction “trauma survival decision making” doesn’t make sense

Page 6: variables spelt “varibles”

Page 6: The Conclusions paragraph wildly oversells the importance and contribution of this paper. Is the method of rule generation presented in this paper novel? The Conclusion suggests it is. The results in this paper are an illustration of applying these varied methods to three specific trauma questions in a single dataset/institution.

Page 7: A total of 1600 cases but 588+1001 doesn’t add up to 1600!

Page 7: The last sentence of paragraph “Table 2 presents the variables for this
dataset” should be moved to start the next paragraph.

Page 7: The final paragraph contains descriptive information that should be put in the table so that the table can be understood without reference to the text. This comment applies to descriptive tables of other datasets.

Page 8: In description of helicopter data, the text from “The number of days spent in ICU…” to “…on discussion with trauma experts” is unnecessarily long-winded. In the final sentence of that paragraph “classified into two groups” can be omitted.

Page 9: Section on C4.5 – what is “ID3”?

Page 10: SVM section – the sentence “The aim to find the optimal hyperplane…” needs rewording

Page 10: Equation 3 – what is “sgn”?

Page 11: Description of 10-fold cross-validation is unnecessarily long

Page 12: First sentence needs to be reworded

Page 12: Use of term Po(R) doesn’t seem to be necessary

Page 12: Equation 10 – definition of estimate of accuracy should be presented in the presented in the preceding paragraph where accuracy-type measures are discussed.

Page 13: First sentence, especially the words “At this stage…” is a peculiar construction.

Page 13-14: From “Logistic regression provides knowledge…” to that subsection’s end on page 14 is unnecessarily long-winded and is simply restatement of textbook description of logistic regression. A more succinct description should be provided in the paper and a reference to pages of a textbook inserted.

Page 14-15: Discussion – it looks to me like logistic regression results are (broadly speaking) no worse than any other specific method.

Tables: I found the insertion of a paragraph of text peculiar but perhaps this is ok by the journal? I would suggest either clearly making some of this text part of the sub-title for the table or incorporating it in the table. Text such as “this table explains the off-site dataset variables” is unnecessary

Table 1: GCS doesn’t take on values 0 <= FURR <=99 !!

Table 2: Number of decimal places should be more consistent, e.g. for ISS just use one decimal place in all columns for mean and for SD (or even better use no decimal places for this variable).

Table 2: There are 1001 Alive patients but then the numbers in rehab/home 628+213 don’t add to 1001. This needs explanation.

Table 3: “ICU stay > 2 days” would be more helpful than just “ICU>2”

Table 4a: “off-set” should be “off-site”

Table 4a: Wald not Wals
Table 4a and 4b: The number of decimal places is absurd. Please reduce these to the absolute minimum necessary.

Page 24: Under the column “Method” there are two cells with ? in them.

Page 29: Under the column “Test Acc” the text “CART” appears erroneously in one cell. Why abbreviate Accuracy in the heading for this column?

**Level of interest**: An article of limited interest

**Quality of written English**: Needs some language corrections before being published

**Statistical review**: Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests**:

I declare that I have no competing interests