Reviewer's report

Title: Prediction of Conversion of Laparoscopic Cholecystectomy to Open Surgery with Artificial Neural Networks

Version: 1 Date: 22 June 2008

Reviewer: Monique Frize

Reviewer's report:

First: Responding to questions in guideline:
1. Is the question posed by the authors well defined? The authors have clearly stated that they wish to predict the conversion of laparoscopic to open surgery, but failed to explain why this would be of interest to surgeons or medical practitioners. Assuming this prediction would not change the outcome, in what sense would it help? Advising the patients prior to surgery? Preparation of the patient? Another point not mentioned in the paper is when is this prediction to be done? Just before surgery, or the day before, or days or weeks before? At the time of the decision that surgery is indicated?

2. Are the methods appropriate and well described? To a certain extent, we can see what is being done. However little detail was provided on the tool (ANN): was the commercial package used as is or were other features added to it? The authors mention that the ANN has two hidden layers, which makes it very complex as a four layer network and extremely difficult to extract the weights at the nodes to identify the variables that mostly influence the outcome. A three layer network (with one hidden layer, an input and an output layer) is often optimal and there are several approaches to extract weights for determining the most important variables. Moreover, when there are few cases in the database (73/793 conversions or 9 percent), which is less than 15 percent, it is difficult to get good results. A method that has been successfully used is a random re-sampling from the conversion cases until the training set contains 20 percent of this population. The test set should remain as it is (9 percent). This improves the sensitivity while not affecting the specificity very much. Regarding the evaluation of the prediction tool, we usually expect to see the results of the ROC curves. A final point: Why is 67 percent acceptable? This has not been discussed or explained.

3. Are the data sound? Considering the comments in question 2, the data collected reflects the methodology of the authors, but improvements could be made on the methods.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition? YES.

5. Are the discussion and conclusions well balanced and adequately supported by the data? The conclusion “appropriate sensitivity and specificity” seems premature considering the limitations of this study. Moreover, some claims in the
discussion are questioned such as on p. 11: “data in the validation group corresponded with that of the training group”. In Table I, two major variables (those that have been found to have a major impact on the outcome) are significantly different: Experience of surgeon and patient admission type. This could bias the results. Another point is that the discussion section is quite mixed, with Table 2 being thrown in there without explaining clearly how this was done (bivariate and multivariate). What about the ANN results for extracting key variables? Here a statistical approach was reported only, while ANN is supposed to be the major approach bring reported. Table 3 shows results for discriminant analysis and ANN, but in the text, little is explained on the discriminant analysis. Some explanation would be useful.

6. Are limitations of the work clearly stated? Other limitations have been mentioned, but not those mentioned above.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished? There is a fairly extensive literature review on the surgery part, but the references are thinner on the current literature on the use of ANNs for medical decision-support.

8. Do the title and abstract accurately convey what has been found? YES

9. Is the writing acceptable? YES well written.

Other comments on more specific aspects:

Top of page 7: some text is repetitive.

P. 8 2nd par. What method is referred to for these results? Next par.: twice the training group results are mentioned. This must be an error and the second should probably say: testing group.

p. 10: 6 lines from bottom: what “other variables”?

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

NO competing interests