Author's response to reviews

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

**Authors:**

Changiz Gholipour (gholipour@TBZMED.ac.ir)
Mohammad Bassir Abolghasemi-Fakhree (bassirf@tbzmed.ac.ir)
Rosita Alizadeh Shalchi (r_shalchi@yahoo.com)
Mehrshad Abbasi (mehrshad_abbasi@yahoo.com)

**Version:** 3  **Date:** 21 October 2008

**Author's response to reviews:** see over
October 18, 2008

Dear Dr. Norton
The editor-in-chief
BMC Surgery

We would like to list our responses to the reviewers’ comments as follow:

Sincerely,

Changiz Golipour M.D.
Associate Professor, General Surgery.
Sinaea Hospital, Tabriz University of Medical Sciences
Tabriz, Iran
Tel: ++(914)3138212
Fax: ++(411)5422862

Responses to comments of reviewer 1:

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?

The following sentence was added into introduction in lines 66-68: “Pre-operative prediction of a laparoscopic cholecystectomy (LC) can help the patient as well as the surgeon prepare better for the intra-operative risk and the risk of conversion to open cholecystectomy.”

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

---


Sinaea Hospital, Tabriz University of Medical Sciences
Tabriz, Iran Tel: 00(914)3138212 Fax: 00(411)5422862
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.

The ANN was designed employing NeuroSolutions ver. 5 (NeuroDimention Inc.) (lines 118-119). We have not used any add-on to NeuroSolutions program. The configuration of our ANN has been determined by genetic optimization, an option in NeuroSolutions programas, which is a common approach to find the best network fitting of an application.

We worked with ANN as a "black-box" as we intended to design a system that predicts the outcome (conversion) based on some factors (patient characteristics). Determining the importance of these factors by analyzing the ANN was not our goal (although this was done by classification regression analyses). As far as we know, this is the usual approach to ANNs. We have mentioned in lines 153-155: “In contrast, the ANN acts in a way which is called “black box”; that is, the association of variables are not clear and explicable.”

Because we intended to compare the ANN and regression discrimination methods, we are hesitated to use re-sampling from conversions which would over estimate the accuracy of ANN approach over regression discriminant analyses.

The prediction of a dichotomous outcome by the values of a continues variable is best evaluated by ROC curve depictions; however, we have predicted a binary outcome by a binary variable based on which we think the ROC curve depictions would present less data than that we have presented in Table 3. The sensitivity and specificity of ANN (67 and 99, respectively) are superior to those of traditional regression methods (56 and 82, respectively). We have employed this rational throughout the manuscript discussing the appropriateness of the accuracy of ANN. In fact, with the available data in regard to prediction of conversion of LC to open surgery there is no clinically acceptable accuracy threshold for prediction tools, or there is no consensus on the subject. Our manuscript open up or initiate a more precise approach compared to traditional methods in this regard.

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.

We tried to addressed the corresponding comments of question 2 above.

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 1, 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 being 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.

The sentence including the phrase “appropriate sensitivity and specificity” in conclusion section was changed in lines 198-201 into: “Our findings revealed that based on preoperative specifications of the patients, conversion of LC to open surgery is predictable by ANN with an accuracy (sensitivity: 67% ; specificity: 99%) superior to that of traditional Discriminant analyses (sensitivity: 56% ; specificity: 82%).”

We agree that the differences between validation and training groups may bias the results in particular for traditional regression approaches as we have mentioned in lines 157-159: “… the discriminant formula are not trainable while the ANN models can be enhanced according to the information derived from new data and can adopt new conditions, for instance improving surgeons’ skills.” The improving surgeon skill is an inevitable trend in the course of time and the selection of more elective patients in validation group is accepted as a methodological pitfall and is added into weak point section in lines 188-191 as follow: “Forth, the robustness of the results could be impaired by the differences of characteristics of validation and test groups (i.e. surgeon skill and prevalent emergency surgeries) in particular for discriminant regression analyses. Interestingly the ANN approach is especially appropriate for evaluation of data of shifting populations. “

The following sentences were added in lines 100-102 into statistical section for clearing the statistical method employed to generate table 2:” Two set of regression models were designed to assess the association of predicting variables with conversion. In the first set, each predictor variable was entered into the model and then all the variables were inputted altogether.” The second paragraph of the Result section is based on the data of Table 2 intending to provide a clinical inside on the correlates of conversion, a subject which could not be addressed by the ANN approach.

The major finding of table 3 in regard to regression approach is stressed in lines 133-135 as:” The models constructed based on the data of validation group – using discrimination analysis – had sensitivity and specificity of 56% and 82%, respectively; the ANN method raised these values to 67% and 99%, correspondingly”. It is noteworthy that the regression approach data presented in Table 3 is posted mainly to provide a basis for comparison with corresponding ANN results.

6. Are limitations of the work clearly stated? Other limitations have been mentioned, but not those mentioned above.
We have extended the Limitation section as noted above in lines 188-191.

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.

The literature is scant in regard to the application of ANN to predict conversion of LC to open surgeries; we have already cited the articles with close correlation to cholecystectomy surgeries. However the ANN has been employed to predict other clinical conditions/outcomes abundantly, we found the corresponding accuracy of ANN to predict other outcomes confusing in term of relevance to our studied subject.

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.

The sentence was corrected as follow in lines 98-100:” The retrospective data of 793 subjects in the training group were used to create a discriminant analysis model. The prospective data of remaining 100 patients of validation group was used to validate the analysis.”

“Bivariate regression model” was added in line 125 and the corresponding analytical method was described in Statistical Method in lines 100-102.
Second “…training group…” was corrected by “…validation group..” in third paragraph of Results section (line:134).

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

In lines 175 to 177 we cleared that the discussion is concerning the findings of table 2 and the “other variables” are those included in multivariate model.

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

We, also, have included the declaration of competing interest in line 209: “The authors declare that they have no competing interests.”
Responses to comments of reviewer 2:

- In order to appreciate the importance of conversion rate, the reader has to know the over-all selection for laparoscopic cholecystectomy. Hence, the number of open cholecystectomy done during the study (training and validation) periods must be given. Length of stay (over-night stay) for completed laparoscopic, converted laparoscopic and open cholecystectomy should be given.

We did not extract the data of open surgeries from archived data files for this study; however, we have already published an article 4 including the corresponding data of the Laparoscopic and open cholecystectomy operations in the same center of this study. We added that citation in lines 74-75 as follow: "We have already published the corresponding comparative data of laparoscopic and open cholecystectomy operations in our center 4." 

- The proportion of emergency cases in the training and validation group differs fundamentally. I guess this is a writing error. If so, it must be corrected. If not, an explanation must be given.

Please refer to the response to question 2 and 6 of the reviewer 1 above.

- Which practical consequences can be drawn from this study? Are conversion risk calculated for the individual patient? What information, if any, is given to the patients after such calculations?

We intended to assess and compare the accuracy of two different analytical approaches to predict conversion rates of LC to open cholecystectomy; however, we may provide the trained ANN program as an additional file for journal website on editor demand and coordination. The regression discriminant function is available for individual risk estimation with accuracies presented in table 3 but is population specific and in contrast to ANN could not be extrapolated for use in other population or settings.

- It is of interest to know if the authors have any experience of small-incision open cholecystectomy, which according best available evidence1 should be considered equal to laparoscopic cholecystectomy concerning recovery and complication.

1. Keus F, deJong JAF, Gooszen HG, van Laarhoven CJHM. Laparoscopic versus small-incision cholecystectomy for patients with symptomatic cholecystolithiasis (Review). The Cochrane Database of Systematic Reviews 2006;Issue 4. Art Nr.: CD006229. DOI:10.1002/14651858.CD006229. Level of interest: An article whose findings are important to those with closely related research interests

We do somewhat similar procedure but we do not have any available systematically collected data to answer the question.

- Quality of written English: Acceptable
- Statistical review: No, the manuscript does not need to be seen by a statistician.

---

Responses to comments of reviewer 3:

1. Is the question posed by the authors well defined? Yes. The intent of this study was to predict the conversion of LC to open surgery using ANN.

2. Are the methods appropriate and well described? The methods are appropriate. The authors define a large retrospective training population and a subsequent prospective training population, although this population is significantly different from the training population in at least two aspects that influence the statistical model. The description of the methods would benefit of clarification of the term “enrolled” for the retrospective section and an easier to read description of the Neural Network construction.

We agree that the differences between validation and training groups may bias the results in particular for traditional regression approaches as we have mentioned in lines 156-158: “… the discriminant formula are not trainable while the ANN models can be enhanced according to the information derived from new data and can adopt new conditions, for instance improving surgeons’ skills. “ The improving surgeon skill is an inevitable trend in the course of time and the selection of more elective patients in validation group is accepted as a methodological pitfall and is added into weak point section in lines 188-191 as follow: “Forth, the robustness of the results could be impaired by the differences of characteristics of validation and test groups (i.e. surgeon skill and prevalent emergency surgeries) in particular for discriminant regression analyses. Interestingly the ANN approach is especially appropriate for evaluation of data of shifting populations. “

The definition of participants of training group was cleared in lines 79-81 as: “A total of 793 consecutive patients (639 females and 154 males) operated between 21th March 1997 and 20th March 2004 were considered as the training group.”

1. Are the data sound? Univariate and multivariate analysis were employed and compared with the results for the ANN. There is potential for significant interaction among several of the variables bilirubin, Alk phos, CBD size, CBD exploration, experience of the surgeon) listed as reason for conversion, which is not addressed.

To assess the correlates of conversion, we employed the bivariate models; and only one multivariate model including all predicting variables was designed. We agree that there are some certain co-linearities between the laboratory data and sonographic examination findings. Nevertheless, the results of an ANN are not influenced by the co-linearity of the predicting variables. However, we do not intend to conclude on correlates of conversion in details, on reviewer demand, we may expand Table 2 and corresponding texts with more multivariate models and different levels of adjustments considering the interacting covariables.

2. No information is provided on the standard management of common bile stones in the practice of the authors. CBD stones were the single significant determinant of conversion in the multivariate analysis however.
The description of the routine practice of the CBD stones is provided in a previous publication\textsuperscript{4}. We added the citation in lines 74-75 as follow: “We have already published the corresponding comparative data of laparoscopic and open cholecystectomy operations in our center\textsuperscript{4}.”

3. The authors predict 5/9 conversions with a discrimination model and 6/9 conversions with the Artificial neural networks correctly. The occurrences of conversions in the validation group may be too small to show that the ANN is indeed significantly more capable in predicting the conversion than the easier to use method.

Yes, we agree that the difference of sensitivities are not significant due to few conversion cases in validation setting; however, the much higher number of false positive conversion cases has made the positive predictive value of discriminant method (24\%) significantly lower than that of ANN(86\%). Based on the positive result of a predictive tool for conversion with high PPV, the surgeon may decide to perform an open cholecystectomy, or be prepared for conversion during LC.

As an aside: Results for bleeding time which was listed in the method section are not given. The corresponding data was added into table 1 (line:293).

The trend of conversions over the time period of the study (1997-2004 and 2006) would be interesting to know, as in many institutions equipment issues and staff learning curves have affected the conversion rate.

Interestingly, conversely the prevalence of conversion of LC to open cholecystectomy increased over the training time (graph bellow), whoever, the conversion rate decreased in validation group (2006). This mainly is the result of inclusion of young surgeons into practice experiencing during the training period and less couscous selection of patient while supergenes were inhabited with LC. We will include the following graph and corresponding data and discussion into the manuscript if the reviewer justify, however, we think that this is not directly correlated to the main topic of the article which is prediction of conversion not details of correlates of the conversion.
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 discussion concludes that the conversion of LC to open surgery is “fairly” predictable ….by ANN. While the technologically more complex method appears to have less false positives, the utility of the ANN beyond the discrimination method (or even beyond the bedside assessment of the experienced surgeon?) in preparing the patient and the team for a conversion is not clearly supported by the data. It would be interesting to know if the authors changed their practice based on their findings.

Frankly and grossly, we did not change the protocols of our practice. The findings of this study should be revalidated in larger samples in our center, be replicated by other investigators, and be confirmed by
international scholars to be applicable. To answer whether the findings have affected personal impression of practicing surgeons involved in the study, we may run a before-after comparative study over the time.

6. Are limitations of the work clearly stated? Several of the limitations of the study and the Artificial Neural Network method have been clearly stated in the manuscript.
7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished? Yes
8. Do the title and abstract accurately convey what has been found? Yes
9. Is the writing acceptable? Yes
The above would all fall under minor essential revisions.
What next?: Accept after minor essential revisions
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: 'I declare that I have no competing interests'

Sinaea Hospital, Tabriz University of Medical Sciences
Tabriz, Iran Tel: 00(914)3138212 Fax: 00(411)5422862