Author's response to reviews

Title: Is it possible to identify cases of coronary artery bypass graft postoperative surgical site infection accurately from claim data? A multi-model comparison study over 2005-2008

Authors:

Tsung-Hsien Yu (ericthyu@ntu.edu.tw)
Yu-Chang Hou (houyc0115@mail2000.com.tw)
Kuan-Chia Lin (kuanchia@ntunhs.edu.tw)
Kuo-Piao Chung (kpchung@ntu.edu.tw)

Version: 4 Date: 28 March 2014

Author's response to reviews: see over
Dear Associate Editor-in-Chief, Dr. van Mourik, and Dr. Petherick:

Thank you for your helpful comments and encouragement, which have helped us tremendously in revising the manuscript. As you will see, we have made substantial modifications (see italicised and bold sentences with underlining) to improve this manuscript. The following are our point-to-point responses to your comments and recommendations.

Yours sincerely,

Kuo-Piao Chung, PhD
Professor, Institute of Health Policy and Management
College of Public Health, National Taiwan University

Reviewer: Maaike SM van Mourik

The authors have made several changes to their manuscript to improve readability and quality of reporting – several of my questions, however, remain and I have listed them below.

Major compulsory revisions:

1. The authors use the terms ‘traditional surveillance’ in a way that may cause unnecessary confusion. I fear that many readers will think of traditional surveillance as surveillance by manual chart review, hence I would recommend to state in the introduction that all approaches (ICD-9CM model and the alternative approaches) will be compared to surveillance by infection control personnel (your gold or reference standard) and that the relative performance of the ICD-9CM approach will be compared to other methods and refrain from referring to ICD-9 based surveillance as traditional surveillance. I feel that the authors also make the confusion between the traditional approach and the ICD9 approach on p.11

Authors’ reply:

Thanks for your suggestion, the term “traditional surveillance” might lead to confusion. We modified it in this revision. We also adopted your suggestion to modify the statements of our study purpose.

2. The statement on page nine referring to the interpretation of the logistic regression model is, to my knowledge, not entirely accurate. The area under the curve is determined by all the points on the ROC curve, not just a single point. I would therefore state the area under the curve (with confidence intervals) in the results
section for both model development and validation, and refer to selection of the cut-off point as based on maximizing sensitivity and specificity (or any other criteria the authors may have used).

Authors’ reply:

Thanks for your comment, indeed, AUC is not the only indicator, we also adopt the Youden index to determine the optimal cutoff point. However, this information were not included in previous version. We added it into this revision. Furthermore, the value of AUC (with confidence intervals) of development and validation models were also provided in this revision. (page 10, 12-13)

3. Furthermore, please explicitly state the variables included in the logistic regression model and (reasons for) any variable selection. Also it is unclear to me how the use of more than 7DDD of cefazolin and antibiotics could have been made continuous. Was the number of DDD included instead of the binary indicator? Also it seems that some of the antibiotic use variables will overlap, e.g. DDD of cefazolin and all antibiotics, how was this handled? Please clarify the variable specifications in more detail, perhaps in a supplementary table.

Authors’ reply:

Firstly, all criteria were based on previous studies and consultations with infectious disease specialists. However, such criterions varied across different countries. For example, most surgeons in Taiwan prescribe 7 days of prophylactic antibiotic for CABG surgery, which is longer than the usual practice in western countries. Therefore, we cannot adopt the criteria from previous studies directly without modification. We invited several surgeons to modify the criteria. We provided such information in this revision. (page 9)

Secondly, we want to thank you for your comment, the multicollinearity issue might exist in logistic regression model. Therefore, model 4 was redone in this revision. We adopted the stepwise selection procedure to select highly influential variables. The results showed that length of stay and type of antibiotics were left in the final model. The new results of development and validation models were also provided in this revision. Fortunately, the results were similar with the previous version, and the results of model 4 have been updated.

4. Results, p11: please state in the first paragraph the number of SSIs identified by infection control.

Authors’ reply:

In previous version, we put this information in the first paragraph of data source. In this revision, we provided the information according to your suggestion. (p.12)
5. As I also suggested in my previous review, I would be very interested to see what the positive predictive value would the logistic regression model would have been when selecting different cut-off points. For example selecting the cut-off where sensitivity is equal to model 5 (87.5), with the best possible PPV. This will allow for a fair head-to-head comparison of the methods; this is currently not the case and thus with the way the information is currently presented I would not support all the conclusions drawn by the authors.

Authors’ reply:

Thanks for your suggestion. We added a paragraph in discussion section to demonstrate the results of head-to-head comparison between logistic regression model and decision tree model (page 17). The results showed that when the sensitivity of model 4 is equal to 87.5, the specificity, positive predictive value, negative predictive value and accuracy were 97.78% (971/993), 48.84% (21/43), 99.69% (971/974), and 97.5% (992/1017), respectively. The positive predictive value of model 5 was still better than model 4.

Minor essential revisions
1. Please change ‘multivariate model’ into multivariable model’ as this is a more accurate description.

Authors’ reply:

Thanks for your suggestion; we modified the term in this revision.

2. I would suggest paraphrasing the citations to references [12] instead of directly quoting.

Authors’ reply:

Thanks for your suggestion; we modified it in this revision.

3. Given that manual surveillance is not perfect I would rephrase ‘gold standard’ into ‘reference standard’. This is a general recommendation for diagnostic accuracy studies.

Authors’ reply:

Thanks for your suggestion; we modified it in this revision.

4. Please clarify: model 5 is the model with or without pruning?
Authors’ reply:

It is without pruning, we added a note on page 13.

5. Expanding the figure legends for the CART models may help inexperienced readers understand the use of the model.
Authors’ reply:

Thanks for your suggestion; we provided the instruction of the CART model in page 10-11. It would be helpful for inexperienced readers.

6. The authors should present the estimates (including the intercept and standard errors) of the multivariable approach to allow for reproducibility of the results, perhaps in a supplemental data file.
Authors’ reply:

Thanks for your suggestion; we added a new table (Table 2) to present the results of the multivariable approach.

7. Wording issue p 15: the authors refer to the PPV as the ‘worst performance indicator’. This sentence is open for two interpretations a) PPV is not a useful measure of performance or b) some models have a very low PPV. I think the authors mean the latter but I would rephrase the statement to avoid confusion.
Authors’ reply:

Thanks for your suggestion; we rephrased this sentence in this revision. (Page 16)