Author's response to reviews

Title: Characterizing Influenza surveillance systems performance: Application of a Bayesian hierarchical statistical model to Hong Kong surveillance data

Authors:

Ying Zhang (yz62@georgetown.edu)
Ali Arab (aa577@georgetown.edu)
Benjamin J Cowling (bcowling@hku.hk)
Michael A Stoto (stotom@georgetown.edu)

Version: 4
Date: 1 July 2014

Author's response to reviews: see over
July 1, 2014

Dear Ms. Aguirre,

We are pleased to resubmit for publication the revised version of MS: 5888766441173254 “Characterizing Influenza surveillance systems performance: Application of a Bayesian hierarchical statistical model to Hong Kong surveillance data.” The reviewers’ comments were highly insightful and enabled us to greatly improve the quality of our manuscript. Our point-by-point responses to each of the comments of the reviewers are in the following pages.

Revisions in the text are shown using underlining for additions [example], and strikethrough [example] for deletions in the markup version under Additional material files. In accordance with reviewer 3’s suggestion, we have added additional details about the selection of initial values in the text on page 14; we also did additional statistical testing, the results of which are available upon request. We hope that the revisions in the manuscript and our accompanying responses will be sufficient to make our manuscript suitable for publication in BMC Public Health.

We shall look forward to hearing from you at your earliest convenience.

Yours sincerely,

Ying Zhang, PhD and Michael A. Stoto, PhD
Georgetown University
3700 Reservoir Road, NW  Box 571107
Washington, DC  20057-1107
e-mail:   stotom@georgetown.edu
Telephone:  (202) 687-3292
Response to reviewer comments and suggestions

We would like to first thank the reviewers for their careful read and thoughtful comments on the previous draft. We have taken their comments into consideration in preparing our revised draft. Below are our responses (in **Bold and Italic**) to the reviewer’s comments.

**Reviewer 1:**

This paper has several strong aspects, especially the use of many different surveillance systems and data sources; and the availability of serological data (for the pandemic period only), providing estimates of incidence of infection. Unfortunately, it is also a rather intimidatingly long article, which is difficult to read for public health practitioners with limited knowledge of modelling approaches.

*A: We thank the reviewer for these comments, and have addressed the length issue as described below.*

**Major Compulsory Revisions**

**Abstract**

1. The methods section in the abstract provides no information other than that surveillance systems were “identified”. This is in contrast with the very long methods section in the main article.

*A: We have edited the abstract section and added more details to the methods.*

**Background**

2. The background section is too long and a clear and concise (one-sentence) objective or research question is missing.

*A: As the reviewer suggested, we have edited the background sections and clarified the research objectives.*

**Methods**

3. The methods section is very comprehensive consisting of some 10 pages with a large number of tables and figures. Of course, statistical / modelling experts should have the possibility to check the various data sources, modeling assumptions and parameters but I wonder whether this section could not be shortened considerably.

*A: We understand the issue, but due to the novelty of the model framework, we are concerned that cutting down the methods section might affect the readers’ ability to understand the results. However, we have moved some of the methods section to the appendix.*
Discussion

4. It would be good if the authors could start the discussion with a short para saying what their most important finding is.

A: Yes, we have made the change as the reviewer suggested.

5. A large number of references have been used in the background and methods section. In contrast, there are hardly any literature references in the discussion section, while a key role of the discussion would be to place the findings in the light of past work.

A: As the reviewer suggested, we have added more references to the discussion sections.

Conclusion

6. The conclusion section is easy to read, except that more than two pages seem excessively long for concluding remarks.

A: We have shortened the conclusion section as suggested by the reviewer in part by moving the “re-evaluating the usage of information environment data” to the discussion section from the conclusion.

Minor Essential Revisions

Abstract

1. Abbreviations (ILI) should not be used in the abstract, unless explained. The terms %ILI-visit and laboratory

A: We have made the changes as the reviewer suggested.

Background

2. Page 4: I assume that the authors refer to the US CDC. If so, please specify.

A: Yes, it is the US CDC, and the text is changed accordingly.

3. Page 4: It is true that internet-based surveillance systems can provide important supplementary data, but they are certainly not “the baseline standard”.

A: We agree with the reviewer that internet-based surveillance systems should not be used as a baseline standard, however, a number of research studies have, and this comment describes rather than evaluates this work.

Methods

4. Page 8: “sentinel surveillance monitoring” is not a good term
A: We took the word “sentinel” out from the sentence as suggested.

5. Page 8: use the standard, internationally accepted terminology of “Influenza A(H1N1)pdm09 virus infection” rather than “pH1N1”, at least the first time.

A: We have changed the term according to the reviewer’s suggestion.

Results

6. Figure 1 is not very informative and could be left out

A: We understand there is some redundancy in the text and the figure; however, figure 1 seems to help our readers to develop an intuitive understanding of why the model was structured the way it is. Since readers of the published article will not have to go to the end of the paper to view it, they can skip it if not helpful without affecting the flow.

7. While “completeness” or under ascertainment with underreporting are often used in surveillance systems, the term “excess” is not so common and that authors should explain this term in plain text, not only with formulas.

A: The term “excess” is a short-hand for a concept in this specific model, which has quotation mark throughout the whole manuscript including the legends. We assume this term will not be used outside the context of this paper. Also, its meaning is explained on page 14.

8. I would suggest that the authors put some effort in minimizing the use of abbreviations and acronyms.

A: Please refer to the answer to comment (9).

Discussion

9. The discussion section could have been kept free from unintelligible abbreviations / acronyms such as “GP”, “RHE”, and “P&I-HA(0-15yr)”.

A: We have spelled out some acronyms such as GP and RHE in the manuscript as suggested. However, some acronym such as “P&I-HA(0-15yr)” refers to a specific subset of data. For such terms we feel that it will be easier for the reader to keep track of them if we keep it abbreviated and consistent all through the paper. We also highlighted the acronyms in table 1 by bolding them so that they are more visible for the reader.

10. Page 25: I do not understand the statement “The more specific the case definition is, the more likely it is influenced by the information environment”. Is it not just the other way around?

A: There is some confusion in this statement, and we have made the change accordingly. The term case definition usually refers to the criteria for identifying cases (i.e. clinical definition, epidemiological definition, etc.). In the original text, we were referring to the terms/codes used
in diagnosis, case reporting and such. The more specific the term/code is – “flu” as compared to “pneumonia and influenza”, for instance, the more likely the decision of using that term is influenced by what people hear from the news. When physicians were “primed” by the information environment of pH1N1, they were more likely to use the specific term when encountering an ILI patient.

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

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.

Reviewer 2:

One point seems to me as a major revision

1. Is the question posed by the authors well defined?

The question posed is clear but the scope of the question should be better defined. Influenza surveillance is a wide topic and can seek after different goals: early warning, unexpected emerging phenomena, impact, interventions effect, vaccine strain adequation... What kind of surveillance aspects is concerned in this paper must be clearly defined.

A: We agree that the purpose of the study needs to be clarified so have added the following statement to the purposes of the study: “This methodological paper is for public health surveillance specialists to better understand and improve the performance of data systems. “

Some minor changes may improve the quality of the paper and be useful before publication.

2. Are the methods appropriate and well described?

- Study period sounds not very clear to me.

A: The data series are available in different time spans, as shown in table 1. We focused on two time periods: (1) pandemic period – June 15th 2009 to November 22nd 2009 (2) non-pandemic period – January 1st 2007 to April 19th 2009. In addition to the week number, actual dates are added as suggested.

- Modelling methodology is hard to understand if not involved in this topic, but the explanation accompanying the modelling are very clear and useful.

A: We thank the reviewer for this comment.
3. Are the data sound?

Yes but I must admit that the modelling aspects are very detailed and a little bit hard to understand if not specialized in this question.

Table 1 in very important as all the abbreviations are not introduced in the text: should you improve that?

A: As the reviewer pointed out, we do have many acronyms throughout the paper. To simplify things for the reader we listed all of the acronyms in a consistent way in one table. We assume that this table will appear at a point in the published paper that makes it convenient for the reader to understand and refer to.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?: More or less.

Some points should be transferred in others sections:

- from purpose of the study to methods or results

- from data and methods to results section.

A: As noted above in response to Reviewer #1, we have moved a large section of the former introduction to the discussion section. In addition, the Simulation and Model Fitting sections includes results from various sensitivity analysis and model performance measures. Since those results are the immediate steps in fitting the model, we separated them out from those of the final model, which have more practical meaning and implications for public health practice and research. Also, in response to another reviewer’s suggestions, we have moved some contents to Appendix I.

Some parts of the text are redundant and should be simplified.

A: We agree that there is certain level of redundancy in the manuscript, but early feedback from readers suggested that it is very easy to get lost in the technical details, especially when the model structure looks so complicated. Reminding the reader of the overall picture in various point seemed to improve understanding.

5. Are the discussion and conclusions well balanced and adequately supported by the data?

- Yes, but should perhaps be completed in the light of the goal of surveillance chosen for this work (see paragraph 1).

A: We have made the changes accordingly.
- Some points given in the results (especially in abstract results) should be interesting in the conclusion.

A: We have added the findings from the abstract as the reviewer recommended.

- And what may have change by and since the pandemic experience in influenza surveillance should be discussed as it could modify the conclusion of this paper.

A: During the first wave of the pandemic (between April 2009 and April 2010) there were a few additional sources of surveillance data, as described in our manuscript. There were no structural changes to influenza surveillance as a consequence of the pandemic.

6. Are limitations of the work clearly stated?: Yes

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

Level of interest: An article of outstanding merit and interest in its field

Quality of written English: Acceptable

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

Reviewer 3:

I would like to state my expertise is in statistics and I do not have enough knowledge to judge the Public health content. They authors apply Bayesian hierarchical models, which offer many advantages in practical applications. They use OpenBug to fit their model.

Page 7 last line delete “and thus one that is likely to hold true in the future.”

A: We have taken out the sentence as suggested.

Why has DIC been used for variable selection? It has been shown not work very in nonlinear problems. Other Bayesian variable methods should have been employed at least for comparison.

A: We agree that there is some controversy over using DIC as a measure for Bayesian model fitting. We chose DIC because it is a default built-in measure in Openbugs, while others (e.g. AIC, BIC) are not available in the software. Aware of the limitations of using DIC as the only measure, we also tried other methods such as Singular Value Decomposition (SVD) and error-based measure (RMSE) in post-check, to ensure the variables are reasonably selected. In
particular, we used Singular Value Decomposition of the predictors to examine if there is any obvious collinearity or redundancy among the predictors. The max and the min eigenvalues for the predictor matrix show a reasonable spread; there are no singular values. The SVD results are supportive for our variables selection. The RMSE results also agree with the DIC results. Those results are available upon request.

The fact that the author had a problem in setting initial value for the MCMC algorithm could be index of problems of prior specification or identifiability. Looking at their results I suspect the latter. I would perform more extensive sensitivity analysis. Given the nature of their data, stronger shrinkage prior might have been more appropriate.

A: We believe that the problems that arose with some initial values had to do with the fact that randomly generated initial values from OpenBUGS sometimes lead to prior distributions that are not epidemiologically reasonable. As explained in the revised text, we addressed this by trying three different sets of manually chosen initial values and found that the results are robust to different model specifications.

In OpenBUGS, we used initial values close to zero (see Appendix II) as large initial values resulted in very slow initiation of the algorithms. This seems to be a problem specific to OpenBUGS (and the default sampling algorithms it uses). We also used WinBUGS 14 and did not have any issues with selecting any desired initial values. OpenBUGS is an open source version of WinBUGS and although it provides many more sampling algorithms, it may not use the most efficient ones by default. In particular, WinBUGS used "updaterGLM.LogUpdater" and OpenBUGS used "log-linear rejection updater" for some of the alpha's and beta's and "conjugate normal updater" for the rest. The choice of the updater may be in some cases forced by the user but most often the software will make the choice based on an “optimality” criterion.

In order to make sure the algorithm is searching through the whole parameter space, we tried larger values in WinBUGS 14 – about 100 times or more as large, which not only ran smoothly but also gave the same results. When examining the traceplots, we did see the algorithm is exploring the parameter space, and quickly reached convergence with the other two sets of initial values (results available upon request). All these results suggest that our algorithm is robust.

Further investigation of the model is necessary before the paper can be published (as described above).

A: We have done various investigations as mentioned above and are confident about the variable selection given the SVD results and such; the results are robust to different model specifications. Those results are available upon request.
Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report