Author's response to reviews

Title: Forecasting daily attendances at an emergency department using time series analysis

Authors:

Yan Sun (yan_sun@nhg.com.sg)
Bee Hoon Heng (bee_hoon_heng@nhg.com.sg)
Seow Yian Tay (seow_yian_tay@ttsh.com.sg)
Eillyne Seow (eillyne_seow@ttsh.com.sg)

Version: 2 Date: 3 September 2008

Author's response to reviews: see over
Dear Editor-in-chief,

We greatly appreciate the valuable comments and suggestions from the two reviewers on our manuscript of “Forecasting daily attendances at an emergency department using time series analysis”. The reviewers have been of great assistance in pointing out areas that has helped us made the revisions which we have incorporated into the paper.

We have addressed all of the reviewers’ concerns and have revised the paper according to reviewers’ suggestions. The revised paper, and the point-by-point responses are attached below.

Best regards,
Sun Yan, PhD
Reviewer: Mieczyslaw Szyszkowicz

Reviewer's report:

Discretionary Revisions
1. In practice exist two working systems used to predict ED visits on ambient conditions; in Great Britain (for medical service), in Germany (for public).

Yes, we agree with the reviewer. We developed the model to assist resource planning for medical services in the Emergency Department, similar to that in the UK system.

2. Definition of P1-P3 is very rough. (Say if we attempt to predict ED visits for asthma, probably other factors are more relevant.). Also gender and age structure are important. In general, if we add more details, the complexity for predictions starts to grow.

The classification of P1-P3 or the patient acuity category scale (PACS) is currently used in all public sector hospital emergency departments in Singapore for resource planning. Hence we did our modeling at the PAC level.

A new paragraph and a new table have been added in the revised manuscript to define P1-P3 of patient acuity category scale (PACS) by ministry of health (MOH) in Singapore. Please refer to the second paragraph under Methods and to Table 1.

3. As the authors mentioned “The model’s performance is based on historical trends. It is imperative for the forecast to be iterative and updated regularly…” – by a consequence more recent information is more valuable to say “what’s happen tomorrow”.

We agree with the reviewer. We iteratively run the modeling process and update the model every 3 months.

4. Provide a description “how to use such methodology” in practice.

In practice, the emergency department plans its staff roster and deployment on a weekly basis. Therefore, every Saturday we run the model for one week prediction, and the emergency department uses it for following week’s planning. This has been incorporated under Discussion. Thank you.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests: I declare that I have no competing interests.
Reviewer: Jennifer Peel

Reviewer's report:

General Comments
This manuscript covers a potentially interesting topic. However, more detail is needed in the manuscript in order for the reader to be able to understand the methods. Additionally, there needs to be some discussion about what the results mean and how they could be used.

Major Compulsory Revisions
1. Background: If this manuscript is dealing with daily ED attendances, then how could the results be useful for long term financial and strategic planning?

The study has shown the immediate to short-term use of the weekly forecasts in the planning of staff roster and clinical resource planning the following week. For the longer term, as seasonal factors are associated with daily attendances, the daily P3 attendances were higher in the middle of the year. Moreover, P1 and P3 attendances were associated with high PSI readings caused by the seasonal forest fires in neighboring countries. These secular annual forecasts help the department plan staff headcounts and budget allocation a year in advance.

This has been incorporated in the discussion.

2. Background: Explain how these results could be useful for regional health care planning.

The hospital where the study was carried out is a regional hospital, with its catchment of patients geographically determined. The approach proposed and lessons learned from this experience may assist other five regional hospitals and their emergency departments to carry out their own analysis to aid planning and budgeting. Overall, it allows for a basis of macro-planning and allocation of budget by the Ministry of Health, which up to now is based on an average aggregated incremental percentage annual growth.

This has been incorporated in the conclusion.

3. Methods: Provide more information about the P1-P3 levels (e.g., how were they defined and was the definition created a priori?).

A new paragraph and a new table have been added to the revised manuscript to define P1-P3 of patient acuity category scale (PACS) by ministry of health (MOH) in Singapore. Please refer to the second paragraph under Methods and Table 1.
4. Methods: Define month of year: Is this calendar month or month of the study? If calendar month, then the interpretation in the Results section (Univariate analysis) is incorrect.

Table 3 (previously Table 2) has been revised to avoid misunderstanding. Thank you for the feedback.

5. Methods: How were the potential predictor variables chosen? Were there other factors considered or alternative metrics of the variables used (e.g., max or min temperature)? What have the previous forecasting papers shown to be important?

The selection of the potential predictors was based on literature, local observation and availability of data. The results in the literature varied with different healthcare system and different countries. Singapore is a tropical country with only one season throughout the year and the difference in daily temperature is not evident. Hence, we used daily average temperature and found it is not a significant factor for predicting daily ED attendances by time series analysis.

6. Methods: How were the variables entered into the models? For example, as linear terms? Were they same day as the ED visits or lagged (e.g., air pollution from a previous day is often related to today’s ED visits)?

In the Methods section, the modeling and forecasting followed established time series analysis procedures in SPSS. The variables were treated as linear terms and entered into the model by significance tests similar to multiple linear regression models. SPSS automatically finds the best-correlated lag of a predictor for the dependant variable. A new paragraph on lagged correlation has been added in the Results. Thank you for highlighting.

7. Methods: Explain why the models were used (ARIMA and exponential smoothing).

ARIMA and exponential smoothing are two classes of time series analysis models. ARIMA is the most general and popularly used model. Both classes of models are automatically included in SPSS time series analysis for searching the best-fit model.

Considering that most of the readers of this paper are not familiar with modeling and in order not to confuse the readers, we have revised the manuscript to include only ARIMA model for the forecast. The results did not change significantly.

8. Methods: Define the “test” period and the “validate” period – what do these terms mean?
In the revised paper, the following sentences have been added in the subsection of Study design and methods for clarity:

ARIMA models were iteratively applied to P1, P2, P3 and total patient attendances using data of the first 24 months to train, data of the following 6 months to test, and that of the following 3 months to validate. Models elsewhere are usually trained and their performance evaluated on the test data; finally the model with lowest error is chosen as best-fit model. This strategy, however, leads an optimistic estimation of the performance of the chosen model since the data used for selection is identical with the data used for performance evaluation. Therefore, in this study, we used a third data set for performance evaluation.

9. Methods: Why was the PSI used rather than individual pollution levels? What lag was used for air pollution? Why were the categories of PSI used rather than a continuous PSI?

The National Environmental Agency (NEA) of Singapore adopts the PSI developed by the US Environmental Protection Agency that provides easily understandable information about daily levels of air pollution. A range of 1-50 is considered good, while that 51-100 was moderately unhealthy, and >=100 was unhealthy. The reading on most days in Singapore were within good range. Therefore, we categorized PSI (>50 and <=50) for better statistical power.

The lag was automatically computed by SPSS, which was 2 days for P1 attendances, and 0 days for P3 attendances. A new paragraph has been added in Time series analysis in Results for explanation.

10. The important predictors could vary by season – was this assessed?

Yes, we agree with the reviewer. Although Singapore is a tropical country with only one season throughout the year, in this study, we have included the month of the year as a covariate. All significant predictors identified by the model were controlled (adjusted, corrected) by month as a covariate.

11. Results – Univariate analysis: which model are these results from?

The results of univariate analysis were generated by general linear model in SPSS with only each predictor considered one at a time.

12. Results – Univariate analysis: see #4 above

Please see the response to Question 4.

13. Results – Predictors and prediction evaluation: What is the “simple seasonal model”? 
We apologise for the typo error. It should have been simple smoothing model, which is an exponential smoothing model, also a special case of ARIMA models. However, as earlier explained, for clarity we have revised the manuscript to only include ARIMA model, which did not significantly change the results after excluding exponential smoothing model. Thank you for pointing out.

14. Discussion: How is “good accuracy” defined?

As far as we know, there is no specific definition of “good accuracy” of a model. It usually means that the p-value of the model by Ljung-Box test is not significant (p<0.05) and the MAPE is less than 20%. If the MAPE is less than 5%, the model performance can be regarded as being excellent.

This has been included in the manuscript.

15. Discussion: How would these results be used in the health care setting or for regional health care planning? Do the ED attendance levels vary sufficiently to warrant a change in staffing levels (considering the margin of error of these prediction levels)? Would daily changes in staffing levels be feasible or cost-effective?

During the study period, the daily variations noted were quite significant, with daily P1 attendances ranging from 10 to 72; P2 attendances ranging from 96 to 239; P3 attendances ranging from 138 to 307. The model developed has identified factors associated with these variations in a local setting; which in turn were used to forecast future workload. It was developed mainly to help plan staff roster. We applied the model every week to make a weekly forecast.

16. Discussion: What are the strengths and limitations of the study?

Strengths:
The study was designed to help us understand the reasons causing the variation in daily ED attendances in a local setting and a model was developed to help forecast the daily attendances. It is the first study in Singapore. The approach proposed and lessons learned from this experience may assist other hospitals and emergency departments to carry out their own analysis to aid planning.

Limitations
In the model we only used the predictors that were available. There may be other factors affecting the daily ED attendances, like the availability of other primary care facilities and their workload which may predict ED attendances. Another limitation is the use of average daily temperature. Although the range throughout the day may not be wide, maximum and minimum temperature could be a better predictor.
These have been incorporated into the paper.

17. Figure 4: make the figures larger.

We have made Figure 4 larger in the revised manuscript as suggested. Thank you.

18. Table 2: See #4 above for interpretation of the “month” variable. Provide all p-values (do not put “ns”). What is the reference for the month variable?

Table 3 (previously Table 2) has been revised to avoid misunderstanding. Thank you for the feedback.

19. Table 2: What is the purpose of this table? If the objective of the study is the forecasting model, then why is this important? If the study was not designed to evaluate the relationship of these variables and the daily ED attendances, then this table is not believable and does not contribute anything to the paper.

Table 3 (previously Table 2) presents the results of univariate analysis, while Table 4 presents the results of time series analysis, a multivariate analysis. These are conventional practices.

20. Table 3: Explain the model notation.

Table 4 (previously Table 3) has been revised for clarity and new paragraphs have been added in Time series analysis in Results for explanation. Thank you.

Minor Essential Revisions
21. Define each abbreviation in a table (e.g., P1, P2, P3).

Thank you for the good suggestion. We have added a new table for the definitions of P1, P2, P3 (Table 1).

Discretionary Revisions
22. The title is generic – consider revising.

Level of interest: An article of limited interest
Quality of written English: Acceptable
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.