Reviewer's report

Title: Long-term survival and healthcare utilization outcomes attributable to sepsis and pneumonia

Version: 1 Date: 14 September 2012

Reviewer: Adrian Barnett

Reviewer's report:

The study used a large sample with striking results in terms of long-term survival with sepsis and pneumonia in those patients that make it out alive. It was good to see that healthcare use was adjusted for survival time. The adjustments for healthcare use close to death were interesting and well argued. Overall it was an interesting paper.

Minor Essential Revisions

My main concern is about subjects were were re-admitted to the ICU after their index admission. The best approach for these patients would be to censor them at their second admission, and then use them in the data again with their second admission with the clock reset to zero. Perhaps there weren't many of these patients, but the numbers in Table 3 show increased healthcare use in patients with an infection.

'The mechanism underlying the larger, persistent long-term mortality effect related to pneumonia relative to sepsis is unclear.' Perhaps this is because of readmission mentioned above? Or perhaps it is due to lasting organ damage.

There is a chance of a healthy survivor effect. Those patients that get an infection but manage to survive may be stronger than those patients who get an infection but die in hospital. This would select an infection group that were 'healthier' than the controls, and so bias results towards the null. Although it is complex to figure this out, as there's also the chance of long-term damage as mentioned above.

Discretionary Revisions

- The 'background' section of the abstract is more like an 'objective' section.
- The abstract background talks about data only in those discharged alive, but then the abstract results talk about the risk of death. I later understood that this risk of death referred to death outside the hospital, but this wasn't clear at the time. The text at the bottom of page 5 and top of page 6 would be a better opening for the abstract.
- Background, are the costs in US dollars?
- "Variation in time at risk is typically handled by including ln(exposure) as a variable in the model and restricting its coefficient to one." This is more
commonly known as fitting exposure time as an offset.

- Two decimal places are not needed for many of the results in Table 1. One or no decimal places would make the table easier to read. Similarly in the text, words like, '869.13 to 937.31 days' would be much easier to read as '869 to 937 days'. I wouldn't say that 0.1 days was clinically important!

- There were quite a lot of acronyms that were unfamiliar to me (e.g. ACC, LTC). I would consider using the longer more informative versions, rather than the acronyms.

- I would consider changing the x-axis in figure 1 to years instead of days.

- Table 2, the difference in survival between the first year and later years is visible in Figure 1. If the x-axis were on the scale of years it would make this change easier to see.

- 'multiple variable Cox' not 'multivariate Cox' (which means multiple dependent variables)

- It would be preferable to have the confidence intervals in Table 2, rather than the SEs.

- Page 12, 'In this sample,' It's not necessary to say this, as it's implied from the methods. The only reason to add qualifiers like this would be if the sample changed.

- Table 3, the zero p-values ('0.000') should probably be changed to "<0.001"

- Table 3 title, a brief explanation of what the statistics in the cells are would be useful.

- Figure 2, Add a horizontal reference line

**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