Reviewer's report

Title: Validation of a Transparent Decision Model to Rate Drug Interactions

Version: 1 Date: 7 March 2012

Reviewer: Daniel C Malone

Reviewer's report:

Review: Validation of a transparent decision model to rate drug interactions

Overall summary:
This paper is a report evaluating a proposed new grading system for drug-drug interactions. The authors report the development of a decision model that categorizes a drug combination into one of 5 categories. The paper is misleading in that this approach is not a true decision model, whereby the cost and benefits are valued, but rather is more accurately labeled a path model or flow diagram.

The authors propose a new grading system using values of A to E, with values of A meaning no interaction and E representing a contraindication. While the grading system is somewhat intuitive, the lack of definitions and/or examples is problematic. What is a “Serious adverse event”? Who defines it? Who determines if medical action is necessary? Why is the focus on outpatient (Category D) – but the validation was conducted using inpatient data? Who determines if the risk-benefit ratio is unfavorable? It is not clear from the paper how these attributes were defined. Left undefined, the rating system is no better than existing systems and has the potential for high-degree of subjectivity bias.

The discussion section of the paper is really results, not a discussion of why this tool would be better than the existing tools available. More importantly, the tool isn’t compared to the Hansten and Horn scale – which was cited in the introduction. This approach doesn’t seem to offer any advantages over other methods to classify drug interactions.

To validate this scale it would have been better to evaluate what happens to patients exposed to interactions at the various levels. Do patients who receive Category E drug pairs have a higher mortality rate? Longer length of stay, increased time in the intensive care unit, be more likely to be hospitalized or go to a urgent care facility? Those metrics are what is needed to improve on the current approaches to rating interactions. This study falls well short of any meaningful evaluation in my opinion.

Below are specific comments the authors may wish to consider.

Specific comments: (Mostly Compulsory Revisions)

1. Introduction – end of 2nd paragraph. “The problem we face today is not the lack of information on DDIs, or the type of classification, but the missing
comprehensibility of each rating.” I agree and disagree with this statement. With respect to disagreement – while there are lots of published studies concerning DDIs, the overwhelming majority are case reports – which is an extremely low level of evidence. We need better evidence, not more evidence. With respect to agreement – the current rating systems are difficult to comprehend because they are not well defined. What does “contraindicated” mean? Under any circumstance the benefit is less than the risk? I would have a difficult time making that distinction – very few drug pairs would never ever be given together.

2. Methods- six questions: As stated above, the authors need to define the terms used to classify the interactions. What is a “serious” adverse event? What is a normal patient population? The lack of definitions for these important questions affects the underlying validity of the approach.

3. It is stated that the sample data for the study was obtained from primary and secondary care and from ward rounds … “to ensure they were clinically relevant”. Thus, by definition, does this mean that no interactions should have been rated as A?

4. Results – it seems to me that the authors started the results by describing problems with the approach – rather than how well the model performed. I would suggest moving the application error to either the methods or later in the results.

5. The level of agreement is not surprising – given that the clinical pharmacologist essentially used the same approach as the decision model. What is more telling is the much lower Kappa value when the results are compared to MMX. While the value is different from zero, the it is not that high either. It is not clear that either standard (clinical pharmacologist or MMX) is a good “gold standard.” Another point of concern is how the analysis was conducted with respect to those drug pairs not rated by MMX. It appears that substantial number of drug pairs were not found in MMX, excluding them artificially inflates the Kappa statistic. More details is needed to explain what was compared.

6. The term “Systematic Error” is used frequently throughout the paper but it unclear to me what is meant by this term. The first case is in the discussion of the Bland-Altman plot. This isn’t systematic error in my opinion – rather “systematic difference” – as used in the first paragraph of this section. Also, what is meant by the phrase “The classification differed by up to three ratings”? Does this mean that one classification was an “A” and other was a “C”? Wouldn’t the term “ratings” be better stated as “classification categories”?

7. It is unclear why disagreements were excluded (second paragraph of Systematic Error” section.

8. Paragraph after figure 3. The first sentence makes no sense. What does the 7% and 9.5% represent? It is unclear. - Minor essential revision

9. Discussion – Systematic Differences – 1st paragraph, 2nd sentence. What is a “higher” case? Do you mean level “E” vs. level “A”? I think this is the direction but it unclear since labels were not numerical, but alphabetical. - Minor essential
The next sentence “Patients requiring complex monitoring....” I didn’t think that patients were evaluated in this study. Do you mean “Interactions that require complex monitoring”? - Minor essential revision

11. The Limitations section fails to recognize subjectivity by the individual raters – and that different raters may result in different findings. This is a major limitation of the study but not addressed. The authors also fail to mention how the lack of evidence for some interactions in MMX was a major limitation – possibly skewing the results.

Tables need more descriptive titles - it is not readily apparent what the numbers represent – as well as the categories. At a minimum – footnotes are needed. Highlights should also be explained in the footnote. Tables should stand on their own.

Figures 2 and 3 need more descriptions – as presented – one cannot readily interpret them.

**Level of interest:** An article of limited interest

**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.

**Declaration of competing interests:**

No conflicts of interest. No stocks, shares, etc. No patents. No financial interests. No non-financial interests.

I declare I have no competing interests.