Reviewer's report

Title: Evaluation of Standardized Doctor’s Order Sets as an Educational Tool for Undegraduate Medical Students

Version: 1 Date: 29 August 2012

Reviewer: Irene Ma

Reviewer's report:

Major Compulsory Revisions:

1. Methods: The authors stated that the use of order sets was hospital-based. Specifically, that the COPD order set was implemented only at Hospital B while the CIWA-Ar order set was implemented only at hospital A. Why then were 5 students using CIWA-Ar order sets at TWH and 43 students used CIWA-Ar order sets at SMH? Who were these 5 students who were able to use the order sets? The discrepancy between the described methods and the results make the study difficult to interpret.

2. Outcome: The a priori marking scheme of the questionnaire should be presented as an additional file since the open-ended nature of many of the questions clearly can render the marking to be quite subjective. In particular, how to score writing admission orders that specify dose and frequency would be particularly subject to interpretation. I find it difficult to see how scoring could be standardized. Thus, presentation of the marking scheme would be critical in persuading the audience that the marking scheme is standardized.

3. Choice in outcome: if only the use of AW order entry could be analyzed, why is the primary outcome a composite score of both AW and COPD? Please defend the choice in outcome. With secondary outcome: please clarify if the score consists of admission order writing portion of the examination of both AW and COPD or is it only for AW?

4. Analysis: multivariate analysis uses base hospital as a control variable. Again, please clarify this. If the base hospital is what dictates the exposure to the order writing set, why is this included in the multivariate analysis?

5. Limitations: Additional major limitations need to be expanded upon. For example - a) the fact that exposure to clinical cases is by self-report. Very likely, students who saw and admitted cases of AW or COPD but do not remember so going to be systematically different from students who admitted cases and remembered doing so. b) Power - this study estimated needing 120 participants, but final analysis compared 48 students who used standardized order set vs 20 who did not. c) confounders - the role for confounders cannot be understated in this observational study. d) generalizability.

6. Throughout the manuscript - the conclusion that there was no significant
difference between control and intervention groups should be revised to this study did not detect a significant difference. The inability to detect a difference should not be equated with lack of difference, especially in light of the study's limitations. Likewise - exposure to clinical cases was associated with improvement in knowledge. It should not be stated that exposure to clinical cases improved knowledge.

Minor Essential Revisions:

1. Background: A more thorough review of the literature will help set the stage. Suggest expanding more in either the introduction or the discussion upon the need to balance Only 8 references are cited. Perhaps include other key articles such as Massaro TA Acad Med 1993 and Bates DW's work on CPOE.

2. Methods: The details on the year 3 six-weeks and year 4 three-week rotation should be clarified. Do students undertake 6 weeks in their 3rd year and then rotate again through for 3 weeks in their 4th year or do students rotate through only once. If the former, then the study needs to account for students who have been exposed to both hospitals. Further, workload of trainees for each hospital should be presented. If at one hospital, for example, trainees admit on average 1 patient per evening, compared with another hospital where trainees admit on average 10 patients per evening, then it would be difficult to attribute any learning outcomes to the type of order systems.

3. Methods - The # of admission per hospital per year should be provided in order for comparisons to be made between hospitals in terms of % of admissions per year on COPD and alcohol withdrawals. Presenting only absolute numbers of admissions per year makes it difficult to compare the 2 sites.

4. Outcome measures - page 7 - the authors made note of true or false questions in assessing knowledge. I was unable to identify any true/false questions in the attached appendix.

5. A flow chart is needed in order to present the study population. It would appear that for AW - out of 175 students, 109 were in one hospital, while 66 in the other. of the 109, 74 were exposed to AW patients (35 were not); and of the 74, 50 wrote orders (24 did not), and out of the 50, 43 used standardized order set while 7 did not. In terms of 66 students in the other hospital, 31 were exposed to AW patients (35 were not); and of the 31, 18 wrote orders (13 did not); and of the 18, 5 used standardized order set while 13 did not. Thus the comparisons were 43 in one hospital plus 5 in the other who used order sets compared with 7 in one hospital plus 13 in the other who did not. Thus it will take the readers too long to sort out your comparison groups and where they came from. Also please present the flow chart for the COPD portion as well. This will allow the readers to understand the representativeness of the student population.

6. Page 9 - the authors presented total exam score difference of -0.41 (95% CI -2.44 to 1.62) please clarify if this is truly the exam score difference that is presented or is this the multivariate parameter estimate that is being presented.
Same with the -0.14 presented for the secondary outcome measure.

7. Results, scores' correlation with level of training: please present these results.

8. Table 1 - please refer to hospitals as hospital A and hospital B as is done throughout the manuscript. Please present as a flowchart rather than a table (see #5 above).

9. Table 2- CI of estimates need to be presented

9. Please revise to improve clarity of the variables (what they are being compared to) - eg - year 3 (compared with year 4); etc.

Discretionary Revisions:

1. Discussion, second paragraph: please clarify if this study is evaluating knowledge acquisition or retention. It seems to me that it is evaluating only knowledge acquisition.

2. Face validity raised in the discussion section. This term is generally considered archaic and probably should be avoided.

3. 2nd limitation raised by the authors - the authors can test whether or not scores are trending higher as the study progressed which can then allow inferences on whether or not the impact of order sets may be attenuated.

4. Last line of conclusion - in my version this sentence is incomplete. Not sure if it’s just my version.

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.