Reviewer's report

Title: Applying a hybrid team-based and lecture-based learning method to neurology clerkship in China

Version: 4
Date: 10 December 2013

Reviewer: Matt Boyd

Reviewer's report:

Manuscript: Applying a hybrid team-based and lecture-based learning method to neurology clerkship in China

Overall

This paper details an interesting result when comparing three methods of delivering a neurology clerkship. A hybrid lecture and team-based approach appears superior to either approach on its own.

There are some major compulsory revisions that would be required before publication of these findings.

1. The research question is suggested in the final paragraph of the introduction, but should be articulated more clearly, and should also appear in the abstract. No formal hypothesis is described. If there was a formal hypothesis this also needs to be stated clearly at the outset.

2. Description of the methods requires more detail. Particularly with respect to Group B, it is not clear where the lectures took place, or how many hours of lectures there were. Did students attend a series of patients or just one? Was the number of hours work that students engaged in with each method comparable? Did group A receive half as many lectures as Group B? And were the topics the same, with reduced content, or were a reduced number of topics covered? Similar questions remain for the team component of Group A.

3. Furthermore, why was a weighting of 40% practice test and 60% theory test chosen? This is relevant because combined scores are reported, and it is not clear that the statistical significance of the group differences for the combined scores would remain under different weighting regimes. Was the weighting determined a priori?

4. Were the practice and theory tests standard and well established in the programme? Or were they developed for the purposes of evaluating this intervention? If so, is there evidence of reliability or validity that can be presented?

5. Do the 127 participants represent the entire 4th year class and if not, how were they selected. How were they randomized to the 3 groups? Was a sample size calculation done?

6. It would be useful to have a copy of the questionnaire that was used in the
end-of-course evaluations, could the actual questions be incorporated into Table 3?

7. In the results it would be useful to report the effect sizes of the significant inter-group differences, as the raw scores, though significantly different, have a narrow range across groups.

8. The final column in Table 2 is not useful. This is because the total score is an amalgam of 40% practice test and 60% theory test, given that there are different relationships between the method of course delivery and scores on practice test, from those in theory test, it seems more appropriate to report just the two individual kinds of test. A cursory glance at the table suggests that there may be a significant difference between LBL and TBL if practice and theory were weighted 50% each. Is this the case?

9. This reviewer found it difficult to interpret the results purporting to demonstrate increased motivation and teamwork spirit of the students (a claim prominently made in the abstract), given that no similar data was reported for the lecture-based course alone. Does TBL-LBL or TBL alone foster these or would all students regardless of course methods report #excellent# scores in these aspects?

10. In general these and other limitations of the study could be addressed more formally. In particular was there any possibility of bias in student self-reports on the evaluation questionnaires? If so, how was this managed?

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

The title could indicate that the research was an experiment comparing three methods of teaching the clerkship.

p. 5 Results should be removed from the introduction.

p. 12 #LBL is suitable for neurology which is known to be less accessible and user friendly than others# # other what?

p. 13, #These factors often result in lower than expected learning outcomes# # what were the expected learning outcomes in this study? Is there data using the same tests from previous classes for comparison?

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: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests: I declare that I have no competing interests