Reviewer's report

Title: Do Students Benefit from Developing their Own Emergency Medicine OSCE Stations? A comparative study using the matched-pairs method

Version: 1 Date: 9 December 2012

Reviewer: Sigrid Harendza

Reviewer's report:

This article addresses the interesting topic to have students design OSCE stations as a teaching format and to test whether they achieve higher scores in an OSCE exam. As a single study with very few participants it inevitably of course has problems in interpretation and design, mainly because of issues of generalizability. Its importance though is probably best seen in terms of it being a paradigmatic example of an interesting teaching format (designing OSCE stations and testing them) which will need a stronger focus in terms of the research question and the conclusions drawn.

Major compulsory revisions:

1. The title should be changed from a question to a statement which includes the main finding of this study.

2. In the abstract a few terms should be adapted. The term “learning performance” is misleading; there is either “learning” or a “learning method” or “performance” (as in demonstrating something). In the method section of the abstract it says that students were preparing for the OSCE. If I understood the outline of the project correctly, all participating students had just finished an MC emergency medicine exam and had not had the practical course in emergency medicine at the time of the study, so why should they be preparing for the OSCE which would have to be performed after that course? The term “exam results” is misleading because it could also mean the MC results and should be changed to “OSCE results”.

3. The background section should not end with a summary of the results but rather provide a solid research question as a connecting passage to the methods section. Please delete the last paragraph of the background section and provide a research question instead.

4. Please restructure the methods section because some aspects of the design of the study remain unclear. Please also provide some information about the curriculum in Leipzig, especially how and where emergency knowledge and training are provided and how and where the exams (MC test and OSCE relate to the teaching).

As far as I understand there is some sort of emergency medicine teaching in the 7th semester which is followed by an MC test. From the students who passed the
test 13 volunteered to participate in the intervention group, 13 students were matched to those and another 20 were randomly chosen. These 46 students took the new OSCE before the participated in the four week (?) practical course in emergency medicine where then the OSCE would be provided at the end to test practical emergency medicine skills for all students.

Several questions arise if I understand the design correctly:
• How were these 46 students able to participate in this OSCE after the course as part of their normal assessment when they already new the stations (which would be an advantage compared with the other students)?
• What does it mean that the OSCE was carried out as a “compensatory examination” (page 9). What does it compensate? Or is it compulsory? Please clarify.
• Why is the OSCE described in such great detail and no data are provided about the outcome of the first cohort?
• Why is the OSCE set up in two circuits? This increases the risk of rater differences.
• “Thorough training” of raters is mentioned. Please provide details about how rater reliability was assessed. Please also provide details how standard setting was performed for the stations with global rating.
• Please also provide details on the constructive alignment learning objectives of the emergency medicine course design and the OSCE assessment.
• Please also provide some more details on the actual intervention. On page 6 it says that the students had two workshops. Did they design the stations and were actually being able to test them as raters and participants? How did they get feedback on the design of their stations?

5. There are a few data missing to actually being able to value the results of the study:
• Please provide the maximum amount of points that could be reached in the MC exam, so the reader can estimate how well the students performed on this exam.
• Since the maximum amount of points in the OSCE is 250 and students are able to reach 223 points (which is almost 90%) without having taken the emergency medicine practical course and 233 points (which is a bit over 93%) after the have taken two workshops in designing an OSCE station it makes me wonder how many points students reached in this OSCE after they had taken the emergency medicine course. Please provide these data for comparison.
• Please explain the asterisk in the legend of figure 1.
• In this legend, it also says p>0.05 but I think I should be p<0.05.

6. Depending on the additional data requested the discussion might take a different turn. If it turns out that the whole cohort of students will reach a similar average point number (i.e. 233) as the intervention group one could argue that a two day workshop of designing an emergency medicine OSCE station provides a
similar amount of knowledge and skills gain as the entire emergency medicine
course and that this workshop might be a good replacement for the course which
will free up many resources.

I would not recommend using the argument (page 13) that students can replace
faculty and therefore enhance the use of OSCEs. It could rather be said that if
older students (e.g. during their practice year) help preparing OSCE stations that
this might help to retain their skills and knowledge in emergency medicine (this
might be an interesting follow up project to pursue).

The authors seem to be surprised that students and faculty developed OSCE
stations that were fairly similar. This is what I would expect if a course for which
this type of assessment was designed has learning objectives. Maybe the
authors could not so much express their surprise but rather comment on the
constructive alignment.

7. The number of participants certainly also is a limitation of the study
(particularly having only 13 students in the intervention group) and should be
mentioned; this is balanced by the matched-pairs approach. The relevance of the
overall increase in points (from 223 to 233) should be also discussed here, even
though the Cohen’s d was high. It should also be compared with the overall
results of the student cohort as mentioned above.

8. The summary should be enhanced with conclusions drawn from this project.

Level of interest: An article of limited interest

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.