Author's response to reviews

Title: Students Benefit from Developing their Own Emergency Medicine OSCE Stations A Comparative Study Using the Matched-Pair Method

Authors:

Wolfgang Heinke (wolfgang.heinke@medizin.uni-leipzig.de)
Daisy Rotzoll (daisy.rotzoll@medizin.uni-leipzig.de)
Gunther Hempel (gunther.hempel@medizin.uni-leipzig.de)
Michaela Zupanic (Michaela.Zupanic@uni-wh.de)
Patrick Stumpp (patrick.stumpp@medizin.uni-leipzig.de)
Udo X Kaisers (udo.kaisers@medizin.uni-leipzig.de)
Martin R Fischer (Martin.Fischer@med.uni-muenchen.de)

Version: 2 Date: 22 March 2013

Author's response to reviews: see over
Do Students Benefit from Developing their Own Emergency Medicine OSCE Stations? A comparative study using the matched-pairs method

Wolfgang Heinke, Daisy Rotzoll, Gunther Hempel, Michaela Zupanic, Patrick Stumpp, Udo X Kaisers and Martin R Fischer

Dear Mr Aldcroft,

We would like to thank you and the reviewers for your valuable comments and advice on our manuscript. Please find enclosed the now revised manuscript together with our responses to the reviewers’ comments and the action we have taken.

Our results showed first of all that students should be involved in the preparation of practical tests (OSCEs). We proved that this increases learning success and are convinced that these findings will be of great interest to the readers of BMC Medical Education. The manuscript will contribute to the broader use of OSCE exams and to the inclusion of students in their preparation. In addition, the manuscript also contains suggestions for further research.

As you have recommended, we have had the language and grammar edited by a native English scientific translator.

We hope very much that our manuscript will soon be published.

Yours sincerely

Wolfgang Heinke

Enc.

Responses to the reviewers’ comments
Comments of Reviewer 1 – Answers and action

As a single study with very few participants it inevitably of course has problems in interpretation and design, mainly because of issues of generalizability.

*We agree with the reviewer on this point and have discussed it as a limitation of the study. However, since it is the first study to address this research question, we think our results are of interest to the readers of BMC Medical Education.*

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

*We have followed the reviewer’s proposal and changed the title of the manuscript to “Students Benefit from Developing their Own Emergency Medicine OSCE Stations – A Comparative Study Using the Matched-Pair Method” to underline the main finding of our 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”.

*We would like to thank the reviewer for these comments. We have modified the abstract accordingly.*

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.

*We agree with the reviewer. We have deleted the last paragraph of the background section and provided a research question as suggested.*

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)

*We have inserted the informative section “Emergency medicine on the curriculum in Leipzig”.*

13 students 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 they 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.

We have tried to make it more clear in the revised manuscript that all 46 participants took the regular OSCE after the four-week practical course in emergency medicine. In addition, the 13 students of the intervention group took “a self-OSCE” before the practical training started. During this “self-OSCE”, participants of the intervention group completed an OSCE comprising their self-developed OSCE station. This is described in the revised version of the manuscript as follows: (“The self-OSCE was held under the supervision of the responsible teaching staff (including some of the authors: W.H., G.H., D.R. P.S.) one week before the practical parts of students’ regular emergency training started. The participants of the intervention group had therefore not been trained in practical emergency medicine by the responsible teaching staff when they tested their knowledge in the ‘self-OSCE’. At this point in the curriculum, participants had merely had the opportunity to attend emergency medicine lectures but had not received practical training.”)

In parallel to the “self-OSCE”, i.e. on the same day, and also before practical training started, 20 volunteers took the 10 OSCE stations designed by the responsible staff as described in the method section to test the suitability of our OSCE (= test OSCE group): “This group served as control to exclude the advantage of the intervention group, who became familiarized with the OSCE format by trying out their stations in a ‘mock’ OSCE. Therefore, we compared the intervention group to a group of students who completed the emergency OSCE before regular emergency training started (see above). This group was formed for two reasons. Firstly, we wanted to test the suitability of our OSCE. Secondly, we wanted to demonstrate the efficacy of our four weeks’ practical training in emergency medicine. Therefore, 20 voluntary students of the 7th semester took the emergency OSCE one week before regular emergency training started (i.e. at the same time as the self-OSCE took place). In order to avoid a familiarity bias with the OSCE format, we compared the OSCE scores of the intervention group to the OSCE scores of these 20 students (= test OSCE group).”

In other words, the participants of the intervention group and the test OSCE group completed two OSCEs for the purposes of the study: one OSCE before the practical training in emergency medicine started (“self-OSCE” or “test OSCE”) and one OSCE at the end of the practical emergency training (“regular OSCE”). In contrast, students of the control group participated merely in the regular OSCE at the end of the emergency curriculum of the seventh semester.

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 knew the stations (which would be an advantage compared with the other students)?

Again, we would like to make it clear that, as already explained, 13 participants did not take the “test OSCE” or the “self-OSCE” (control group). These 13 participants served as a control group to test our hypothesis that developing OSCE stations would result in higher scores during the exam. Therefore, we agree with the reviewer that it appears possible that the members of the intervention group had an advantage during the regular OSCE due to their familiarity with the OSCE format.
In order to exclude this bias, we compared not only the scores of the intervention group and the control group obtained during the regular OSCE, but also the scores of the intervention group and the test OSCE group. This comparison is intended to rule out any advantages for the intervention group because both groups (intervention group and test OSCE group) were familiar with the OSCE format. We have described our procedure in the method section of the revised manuscript as follows:

“This group served as control to exclude the advantage of the intervention group, who became familiarized with the OSCE format by trying out their stations in a ‘mock’ OS. Therefore, we compared the intervention group to a group of students who completed the emergency OSCE before regular emergency training started. This group was formed for two reasons. Firstly, we wanted to test the suitability of our OSCE. Secondly, we wanted to demonstrate the efficacy of our four weeks’ practical training in emergency medicine. Therefore, 20 voluntary students of the 7th semester took the emergency OSCE one week before regular emergency training started (i.e. at the same time as the self-OSCE took place). In order to avoid a familiarity bias with the OSCE format, we compared the OSCE scores of the intervention group to the OSCE scores of these 20 students (= test OSCE group)."

• 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.

As requested by the reviewer, have we added the following statement to the text: „Examinees did not have to pass every station – instead, their exam grade depended on their total score.” In addition, we have added a citation to the text which describes the differences between a compensatory and a non-compensatory OSCE format as follows: “In a compensatory approach, examinees must simply achieve a certain total score in the OSCE: an unsatisfactory score at one OSCE station can be compensated for by a better performance at some other station.” [15] Nickendei C, Jünger J: OSCE – hands on instructions for the implementation of an objective structured clinical examination. GMS Z Med Ausbild 2006, 23:1-8.

• Why is the OSCE described in such great detail and no data are provided about the outcome of the first cohort?

The OSCE is described in detail because we believe this description is necessary for the average reader. We think many readers do not know what an OSCE actually is. This detailed description therefore helps the reader understand the whole project. In addition, this description is helpful for other medical teachers developing OSCEs in emergency medicine.

With respect to the outcome of the first cohort, we do not understand what the reviewer means. If the reviewer means that we should have reported the exam results of the remaining students, we would like to point out that, in our view, the remaining students’ results are irrelevant to the research question. Including them would have made the paper more complex and harder to understand for the reader.

• Why is the OSCE set up in two circuits? This increases the risk of rater differences.

The OSCE was divided into two circuits for practical reasons. We had to test 321 students. This was logistically only possible by examining 20 students per hour. We do not think that this affected our results.

• “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.
This began with a 45-minute theoretical training session on OSCE for all raters one week before the OSCE. On the day before the OSCE, once the OSCE course had been completely set up, the raters were trained at their stations. A responsible lecturer demonstrated the performance expected of examinees at each station. The raters were also shown how to use the checklists. The raters were given a final briefing 30 minutes before the start of the OSCE. This training was intended to minimize raters’ impact on the test results. In addition, after each round of tests, the points awarded at each station were reviewed by the first author of this study (W.H.). If any anomalies were found (e.g. above-average or exceptionally low scoring), the raters were given additional training.

A “standard setting” was not specified for individual stations but only for the whole OSCE. Since the OSCE was being held in Leipzig for the first time, a pass score was set arbitrarily at 60% of the maximum score possible. We have described this procedure in the revised manuscript in more detail: "The OSCE was carried out as a compensatory examination (i.e. examinees did not have to pass every station – instead, their exam grade depended on their total score) and the pass rate was set at 60% [15]."

Please also provide details on the constructive alignment learning objectives of the emergency medicine course design and the OSCE assessment.

As is apparent from the revised manuscript, emergency medicine training includes a substantial practical section (see the new paragraph: Emergency medicine on the curriculum in Leipzig). As described in the introduction and discussion of the manuscript, an OSCE can therefore be regarded as a suitable examination form. This is now comprehensible to readers thanks to the additional section on the emergency medicine curriculum in Leipzig and the reviewer’s point is reflected in the revised manuscript.

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?

As requested by the reviewer, we have provided more details on the intervention: "The development process of the OSCE stations was supported by two authors (W.H., P.S.) of this study. Students received feedback on the OSCE stations they had developed: on medical aspects from the relevant teaching staff and on educational aspects from lecturers with a Master of Medical Education (W.H., P.S.). All members of the intervention group used the eight OSCE stations they had developed to test each other’s knowledge of emergency medicine in a ‘self-OSCE’. For this purpose, on the day of the self-OSCE the students were given additional training in order to take on the role of the rater. The self-OSCE was held under the supervision of the responsible teaching staff (including some of the authors: W.H., G.H., D.R. P.S.) one week before the practical parts of students’ regular emergency training started."

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.

We have provided the requested information in the legend of Table 3.
• 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.

The reviewer has misunderstood the study. The scores given refer to the regular OSCE, which was held after the completion of practical and theoretical training. This should now be clear from the revised version of the manuscript. In our view, the difference in scores indicates that by developing their own OSCE stations, the emergency medicine expertise among the students in the intervention group was strengthened.

• 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.

Thank you for this. We have added an explanation of the asterisk to the legend and corrected the p value.

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.

In response to the reviewer’s comments, the revised manuscript should be easier to understand for the reader. As described in the manuscript on a number of occasions, the intervention group carried out two OSCEs: a self-OSCE, before which the students developed their own stations, and the regular OSCE. The self-OSCE took place before practical training in emergency medicine, whereas the regular OSCE was held after practical and theoretical training. This means (as we have discussed in detail in our discussion) that the students in the intervention group benefited from the additional study of practical aspects of emergency medicine in preparation for the exam. We therefore see no reason to amend the discussion.

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).

We have followed the reviewer’s suggestion and added the following statement to the discussion: “OSCE stations could be developed not just by students preparing for the regular emergency medicine exam but also by more senior students (e.g. during their practice year) in order to refresh their skills and knowledge in emergency medicine..”
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.

*Thank you for this valuable suggestion. We have added the following passage to the text:*

“The similarity between OSCE stations reflects the ‘constructive alignment’ of the emergency medicine curriculum. Learners construct their own learning through relevant learning activities. The teacher’s job is to create a learning environment that supports the learning activities appropriate to achieving the desired learning outcomes. The key is that all components in the teaching system – the curriculum and its intended outcomes, the teaching methods used and the assessment tasks – are geared to each other [20].”

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.

*We agree with the reviewer on this point and have therefore mentioned it as a limitation of our study:*

“Nevertheless, for these reasons – and also because the number of participants in the intervention group was relatively low – follow-up studies need to be carried out to confirm the findings, preferably in other areas of medicine and with more experienced test-takers than in our study, because the practice effect on OSCE novices is probably greater.”

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.

*We don’t quite understand what the reviewer means. Comparison with the entire cohort (entire semester) is not permissible for statistical reasons. In addition, we believe that another cohort by way of comparison would make the manuscript more complicated. If the reviewer is interested, they should read up on comparison with the overall cohort in Wolfgang Heinke’s thesis for his Master of Medical Education (Do Students Benefit from Developing their Own Emergency Medicine OSCE Stations? Thesis submitted by PD Dr. med. habil. Wolfgang Heinke at the University of Heidelberg, MME-D, 5th intake 2009–2011)*

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

*We have followed this suggestion by the reviewer: “This finding is a good argument for involving interested students in the development of their assessment.”*
I am confused as to the exact numbers of students; the protocol describes 13 in the intervention and control groups and 20 in the OSCE group, but then discusses 3 males and 13 females (total of 16 students) in the test OSCE group.

Thank you for reading the paper so attentively! This was a typing error. There were in fact 3 males and 17 females. We have corrected the abstract and the methods section.

I greatly enjoyed this paper; the concept is well presented and it contains useful information, with both specific and broad applications. It is given in easily comprehended language and I did not find any significant syntax or grammar errors. I believe that this is worthy of publication - with the revision described above - and that further evaluation of faculty versus student development of stations (as the authors have suggested) is an area worth exploring.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

Thank you very much indeed for these useful and encouraging comments.