Author's response to reviews

Title: A prospective, blinded evaluation of a video-assisted '4-stage approach' during undergraduate student practical skills training

Authors:

Katrin Schwerdtfeger (kschwerdtfeger@med.uni-goettingen.de)
Saskia Wand (saskia.wand@med.uni-goettingen.de)
Oliver Schmid (o.schmid@t-online.de)
Markus Roessler (m.roessler@med.uni-goettingen.de)
Michael Quintel (mquintel@med.uni-goettingen.de)
Sebastian G Russo (s.russo@medizin.uni-goettingen.de)

Version: 3 Date: 9 March 2014

Author's response to reviews: see over
Dear Sirs,

Thank you very much for your detailed comments which helped to further improve our manuscript. We have revised our manuscript accordingly and hope to meet your expectations.

Please find below our point-by-point answers.

Yours sincerely for all the co-authors

Katrin Schwerdtfeger & Sebastian Russo

Reviewer: Henning Biermann

**Major Compulsory Revisions**

The authors evaluate the effect of training with nine dichotomous items. There is no information on what the used items were. Listing them is necessary and only this allows the reader to judge on adequacy of the chosen items. The majority of the students get all nine points without a significant difference. The global performance is in contrast to that significant better in the study group. It should be discussed if the nine items are appropriate to distinguish between a good and an excellent performance and if there is a bias.

The authors agree with the reviewers that additional information needs to be added in regards to student performance and effect of training with the nine dichotomous items. Moreover, it has been difficult to differentiate between good and very good performances, because of low overall item difficulty level. We argue that, because of the well-structured course content, students in both study groups were able to reach good to very good performance scores. We will discuss this point as a study limitation. The following nine items have been evaluated using a checklist:

5.1. Evaluation of consciousness/ personal introduction to the patient
5.2. Palpation of peripheral pulse
5.3. Verbal statement: “No acute, vital threat”
5.4. Cervical spine immobilization utilizing inline-stabilization and a stiff-neck-collar device
5.5. Auscultation of bilateral lung fields
5.6. measurement of Oxygen saturation (SpO₂)
5.7. Measurement of capillary refill
5.8. Abdominal examination
5.9. Examination of pelvic stability (antero-posterior, lateral)

The authors would like to avoid the inclusion of this list of these items in the paper, which are the structural elements that students are being evaluated on, since the danger exist that future students will focus exclusively on these items for study, achieve only a limited understanding of the topic, but continue to score passing grades in the examination.

The study compares a standard 4-step/stage approach with a media supported approach, where the first two steps have been replaced by a video. It was not absolutely clear to me, but it seems as if there is a difference in step two. While the standard second step contains a demonstration with acoustic explanations from an instructor, the video seems to work with type overlays. That means the students have to read them instead of listening to them. This could address the type of learner (aural-style learner/ visual style learner). This should be briefly discussed.

The reviewer correctly identified that in the second step only typed overlays are used instead of audible commentaries. According to Vester’s theory of learning the visual learner could benefit from this modification. Although there has been increasing critique in Vester’s theory in recent years, we have included this point in the discussion and mention that visual learners may benefit from this change (paper: page 11)

The authors state in the discussion, that “the differences in presentation styles in the control group were also reduced since the instructors began using the video to prepare their own lessons”. Was that during the study period or before? What has been done to reduce the differences in the control group before? Did the instructors have an equal background like a PHTLS / ATLS instructor status? Might those mentioned differences be the cause of the improved global performance evaluation of the study group?

All course instructors did have access to the video clips before the course started and therefore were able to prepare the course accordingly. All instructor participated in precourse tutorials, which included guidance for a structured teaching concept and focused on the 4-SA. Most instructors took part in a tutorial that had an ERC format and not PHTLS or ATLS. Because of this the instructors had received a course script and joined detailed preparatory discussions prior to the module’s start.
The randomization did not include individuals but only the allocation of groups. Have there been any differences in the demographic data of individuals apart from gender? (age, learning type, pre-course experience in the field of trauma / ABCDE approach) A table with baseline group characteristics appears necessary. The distribution of individuals to the groups by the dean’s office appears to be at least doubtful when the title of the paper indicates a randomized evaluation and should be addressed in “limitations”. From my point of view individuals as well as groups have to be randomized for categorizing it as randomized evaluation.

The distribution into groups was done at the dean’s office. After that the groups were randomized to the study or control group. It is difficult to indicate the true level of randomization of individuals and we have therefore chosen to make a change to the title of the paper and eliminated the word ‘randomized’ from the title.

We did not study the differences in demographics between the groups except for gender and added this to the limitations of the study (paper: page 13).

The relevance of the student’s subjective impression of knowledge gained and the free-text comments is not clear. The data compares the period before and during the study, but not the study group with the control group. Within the paragraph “methods” the authors describe a six-point Scale on how to evaluate a pre-hospital trauma patient. Within the paragraph “results” the authors describe a percentage concerning knowledge increase about how to treat a trauma patient. Is this the same data from the same item? Please report the results Pre-Post on the Likert-Scale. How is the percentage of knowledge increase defined? If this is data from the same item then please report the results to the described Likert-Scale. Were the Likert scales really titled “totally agree” vs. “not at all true” or is this due to translation? I would expect them to be “agree/not agree” or “true/not at all true”.

The authors have added language to the methods section to clarify. The medical school (dean’s office) collects data independent of this study to evaluate the subjective knowledge increase of students participating in courses. This data was obtained anonymously before and after the course and did not influence the study as it was not possible to know to which group the responding student belonged.
The data collected includes data from the OSCE stations and demonstrate objective and subjective student evaluations by the instructors of prehospitalized trauma patients.

We have used a checklist with nine dichotom items to objectively evaluate student performance and used the Likert scale to subjectively evaluate overall performance of the student by the instructor.

“The topic of the module was the structured examination of a trauma patient following the ABCDE mnemonic protocol [6] as well as integral basic treatment and monitoring (neck immobilisation, monitoring of ECG, pulse oximetry, blood pressure, oxygen administration, insertion of a peripheral venous cannula).” Regarding to the students’ evaluation the course is about examination of a pre-hospital trauma patient. This should be clarified. Why has the ATLS students manual been cited, and not the PHTLS Manual? If the examination is according to the PHTLS concept: Why is ECG part of the primary survey, which normally is not the case? Are there other modifications to the approach?

We cited the ATLS literature since we believe that it is wildly used and that the ABCDE mnemonic protocol was first used in the ATLS system. The ‘A’ and ‘B’ of the ABCDE protocol were first described by Peter Safar, who taught Nancy L. Caroline. We have added a corrected reference based on your comment.

In addition, we choose a modified approach based on the PHTLS recommendations. The use of non-invasive blood pressure as well as a 4-lead ECG is accepted for a trauma patient within the PHTLS concept, if enough human resources are available.

The examiners were only blinded to the allocation of the groups. Has there been a difference in the results of the two examiners, since KF/KS, in contrast to OS, has designed the study and knew about the endpoints? Maybe this should be discussed in “limitations”.

There actually was no difference in knowledge between KS and OS. OS did not design the study, however, did know the study design and hence we expect no difference between examiners. Furthermore, blinding the investigator (regarding the allocation to study or control group) should help to avoid any bias. However, we acknowledge that neither KS nor OS were blinded regarding the overall study intentions.
**Minor Essential Revisions**

There is a low number of citations used for discussion. I would recommend to consider a search for additional citations, since half of the cited papers have already been cited from the authors of reference eight. Maybe there is new data available.

Done

It looks like the course is part of the curriculum. It should be clarified if this is true or if the course is facultative.

You are correct. The course is part of the curriculum. This had to be clarified and has been added to the manuscript.

“This could have been useful in assessing the results of the relative increase in knowledge, and in detecting potential differences in the persistence of the imparted knowledge depending on the teaching method”

Done

Abbreviations of the authors: “KF” is obviously used for Katrin Schwerdtfeger, why not

You are absolutely right. Thank you for that correction.

Methods 2nd paragraph: “patient data” instead of “patient date”

Done

Methods – intervention 2nd /3rd line: either use “14” or “fourteen”

Done

Results – utilization of resources 3rd line: “two” instead of “tow”

Done

**Discretionary Revisions**

ABCDE mnemonic has already been used for the evaluation of casualties in 1979 by Nancy L. Caroline in “Emergency care in the Streets” and since then widely used and therefore not specific for the ATLS manual???
Please refer to the authors comment above.

4-stage approach was adapted for resuscitation training by Bullock as cited but originally published by Peyton as 4-step approach and included the trainees telling what they do, before demonstrating it in the fourth step

Thank you for this reference which has corrected accordingly.

PHLTS, ALS, ETC are not in the list of abbreviations

Done
There is no mention of any power calculations. Was there any calculation attempted? Did the authors consider running a non-inferiority trial?

We did not conduct a formal power analysis before the course began. We suspected, based on previous examination results, that the objective performance would in general be at a high level and hence difficult to differentiate. In addition, there was no data available for the global performance that could have been used to a power analysis. Differences in checklist analysis were small and a very high number of students (<800 Millionen) would have been required for an adequate power (0.8, p < 0.05). The power of the global scoring results was at 0.946, so that we assume that these results are reproducible.

Nevertheless, we understand your comment regarding a NIT. Indeed, we discussed this option prior to start data recording. However, given a consistent course content, we prioritized the OSCE items differently in the past. Additionally, we received student’s feed-back, that despite the same content, teachers prioritized learning goals differently. Hence, we aimed to standardize teaching by implantation of the study video. Thus, the video – provided as a preparation for the teachers (but not shown to the students control group) – was newly implemented. In summary, the needed preconditions for a NIT were not given. Furthermore, the classical 4-SA (control group or potentially used to compare with during a NIT) is still a matter of scientific discussion and research, mainly regarding the need of each individual step to teach single or complex skills.

One of the main conclusions is that the video 4-SA reduces faculty requirements. This is only relevant for topics that require more than one faculty member. It is not generalisable to all topics for teaching. For example, if the trauma station had not involved neck immobilisation, there would not have been a need for two faculty members to demonstrate the ABCDE approach. This needs to be made clear in the manuscript - this particular conclusion is not relevant to all situations.

Thank you for pointing out this important annotation. We revise our conclusion in this point. (See discussion, page 12)

A significant proportion of data was excluded due to the lack of a global scoring - why was this omitted in so many cases? this needs to be clarified in the manuscript. if this
data had been included in the analysis for the checklist, would it have changed the outcome?

This is an important question. The global scoring is unfortunately not an integral part of the scoring process of the OSCE examination, but is added as a separate item. We hypothesize that assessors often did not complete the global scoring because of time restraints or other reasons. We preferred to exclude the checklist analysis from the manuscript when no global scoring was available. There was no statistical difference in the check list analysis, when we included the excluded data (p=0.69).