Author’s response to reviews

Title: Effect of CRM team leader training on team performance and leadership behavior in simulated cardiac arrest scenarios: A prospective, randomized, controlled study

Authors:

Ezequiel Fernandez Castelao (e.fernandezcastelao@uni-goettingen.de)
Margarete Boos (mboos@uni-goettingen.de)
Christiane Ringer (chr.ringer@gmail.com)
Christoph Eich (eich@hka.de)
Sebastian G Russo (s.russo@medizin.uni-goettingen.de)

Version: 4 Date: 11 May 2015

Author's response to reviews: see over
To: Editorial Office, BMC Medical Education

April, 30th 2014

From: Ezequiel Fernandez Castelao
e.fernandezcastelao@uni-goettingen.de
Department of Social and Communication Psychology
University of Göttingen, Germany

Dear editors of BMC Medical Education,

Attached is the revised manuscript reporting our latest research contribution, “Effect of CRM team leader training on team performance and leadership behavior in simulated cardiac arrest scenarios: A prospective, randomized, controlled study”.

We think, that this manuscript should be published as the results of our research support and extend the ERC and AHA recommendation of combining clinical competencies with coordination and communication skills of CPR within medical education.

I am one of the corresponding authors and am submitting this manuscript on behalf of all co-authors: Margarete Boos, Christiane Ringer, Christoph Eich and the second corresponding author Sebastian G. Russo.

All authors have made substantial contributions to all of the following:
(a) the conception and design of this study,
(b) our acquisition of data,
(c) our analysis, and
(d) our interpretation of data.

All five co-authors, including of course myself, have been substantially involved in the drafting and revising of the article regarding its intellectual content as well as the final approval of the version attached for submission. A professional English-for-the-sciences proof-reader, M. Neff-Heinrich, has provided English grammar assistance, her support underwritten by our departmental budget.

This manuscript, including related data, figures and tables, has not been published previously and the manuscript is not under consideration for publication elsewhere.

With gratitude for your consideration of our manuscript,

Dr. Ezequiel Fernandez Castelao
**Point-by-point response – Reviewer 1 (DS)**

**DS1:** Results that are clearly statistically (and also conceptually/clinically) insignificant should not be discussed as being different...

**EFC1:** Yes, we agree (lines 238-244). In the abstract we now explicitly mention the non-significant result (lines 46-52).

**DS2:** The abstract should not just contain p values but also measures of effect size with their confidence intervals or other estimates of precision.

**EFC2:** Modified (lines 53–59).

**DS3:** The authors should comply fully with the detail of the CONSORT guideline; they have not, and some examples are mentioned below.

**EFC3:** The study was oriented to the CONSORT guidelines, most of which could be fulfilled except for some points. The requirements that could not be put into effect (i.e., type of randomization) are discussed in the revised manuscript and also addressed as a response to your comment number 4. The CONSORT checklist is attached.

**DS4:** The limitations of quasi-randomisation (bias) should be discussed.

**EFC4:** Done (lines 306-308).

**DS5:** p-values / stat.tests to compare descriptive data at baseline should never be used in either observational or RCT studies (CONSORT & STROBE)

**EFC5:** We adapted Table 2a and 2b in order to adhere to the updated guidelines for reporting parallel group randomized trials (Moher et al., 2010). However, we think that it is important to report (including p-values) that groups were well balanced in terms of age, gender, and subject familiarity (line 210-213).

**DS6:** The steps taken to protect allocation concealment should be mentioned.

**EFC6:** Done (lines 133–135).

**DS7:** Validation and reliability of the ADH tool should be discussed.

**EFC7:** Yes, we see that we missed this is in our original report. We’ve now added that the tool was Delphi-validated (line 200). Regarding reliability, we see the applied procedure described in lines 204–206 as sufficient to ensure stability of the measures for this study. To use the ADH tool for further research, it will be important to establish inter-rater reliability on a strictly quantitative basis.
**DS8:** 10 teams were excluded after randomization. Why? It appears that actually some of them (how many?) had been non-compliant with attending lectures etc instead of not fulfilling eligibility criteria (in which case surely they should have been excluded before randomization??). Therefore, the analysis has not been proper intention-to-treat analysis, therefore it is prone to bias. I would like to see the results compared for all teams as per intention to treat.

**EFC8:** The reasons for the exclusion of 11 teams are reported (lines 213-217). As this study was embedded in a teaching module that included all fifth-year students, we had no influence on sample size.
Point-by-point response – Reviewer 2 (MK)

MK1: Thank you very much for the opportunity to review this manuscript!

EFC1 (on behalf of the authors): You are very welcome. Thank you very much for your extraordinary helpful comments!

MK2: The authors have investigated whether training designated leaders in CRM would improve later team performance in a simulated cardiac arrest scenario. I think this is a very well-written paper with an appealing study idea. Improving performance of cardiac arrest treatment is very important. I agree with the authors with respect to the implications of their findings for designing curricula. With respect to the manuscript in its current form I have three main concerns:
1) Leadership theory: given the recent developments of leadership such as functional, dynamically delegated, or shared leadership, I do not yet understand the authors rationale for training only few potential leaders.

EFC2: We have now extended the background part in order to cover the gaps you mention (lines 65-83).

MK3: 2) Validity: a) I noticed that you trained medical students and defined the leader before task performance began. I wonder how this could be applied to clinical practice. From my experience, it is exactly defining and maintaining a leader which is so challenging for multi-professional ad-hoc teams in hierarchically structured hospitals. I think this challenge is not represented in your study design. How do you see this?

EFC3: We agree and think that this is an important point. In this study our aim was to show that teams composed of one CRM-trained team leader and three ALS-trained team members show better performance during simulated CPR scenarios as compared to teams with a non-CRM-trained team leader. Even though it was not our primary target to train students on how to identify the best team leader, we think that ad-hoc teams in clinical practice will more likely be equipped to master this challenge if the involved individuals are not only aware of but also trained in different role requirements.

MK4:  b) The CRM training you performed seemed to consist of presentations rather than experiential learning. I think that just paying attentions to presentations instead trying out new leadership behavior within a team, experiencing how team member react to this behavior and jointly reflecting on leadership behavior afterwards is essential for developing leadership skills and I am concerned that the participants could not learn this. What do you think?

EFC4: We completely agree with this point and that is why the students performed debriefings on simulated CPR after the CRM-TL or ALS-add on training (tutorial 4). At this point it is important to mention that this study was embedded in two independent editions of a two-week mandatory course for fifth-year medical students. From an
educational point of view, the whole course—including lecture, CRM-TL or ALS-add on and four tutorials—is the complete training.

Regarding our 90-minute CRM-TL training, we modified the description of the training in order to clarify the contents. We now point out in the manuscript that the training is more than just paying attention to some sheets (lines 138-151).

MK5: 3) Discussion/conclusion: I think some of your conclusions are too strong and are not based on your findings (particularly with respect to NFT).

EFC5: We modified our interpretation of the NFT results in the Discussion.

MK6: Background: 1. I did not see a definition of leadership and suggest providing one.

EFC6: See our enhanced description in lines 72-76.

MK7: 2. From what you have written it seems to me that you consider leadership as certain communications performed by a pre-selected team member. I think that leadership is a social process that involves interactions among team members. Given recent concepts of leadership, such as functional leadership1 and dynamically delegated and shared leadership in healthcare2 3, what is your take on that?

EFC7: See our enhanced description in lines 72-76.

MK8: 3. I think teaching medical students team leadership is important. However, I think identifying and negotiating leadership ‘in the heat of the moment’ (e.g., a senior surgeon and a senior anesthesiologist managing a trauma – who is the leader?) in ad-hoc inter- and multi-professional teams is a huge problem which seems to me somewhat ignored in your study design by randomly designating a leader beforehand. Usually, nurses are present as well but I did not see them included in your study and I am concerned that this limits the validity of your study. What do you think? How do you see this leadership assignment applicable to clinical context?

EFC8: As mentioned in our response to comment 3, our intervention is not targeted at the negotiation of leadership in emergency action. Nevertheless, we do not want to ignore this topic and added it to our argument to the practical implications (lines 286-294).

MK9: 4. With respect to new team concepts such as ‘teaming’4 and human resource development I am concerned that training only designated people as leaders will not contribute to fostering teamwork in healthcare. What do you think?

EFC9: Again, we support your point. Our training for designated leaders was not designed to be implemented as a one-shot intervention. On the contrary, it is crucial to constantly apply complementary interventions—team and experience-based approaches, reflecting interventions, skill trainings, etc., in order to establish a
comprehensive understanding of team processes, which in turn can foster teamwork in healthcare. Our intent was to step up to the challenge of differentiating between dream-scenario training vs. what sort of training enhances knowledge and performance in preparation for clinical realities.

MK10: 5. I think that training physicians and nurses would have been helpful is solving the aforementioned problem. What was your reason for training medical students?

EFC10: Generally, we think that it is important to continuously train all involved persons. We chose students because we are convinced that—in terms of the effects of leadership and teamwork on performance—it is crucial to place a stable foundation in an early career stage of future physicians. We think this approach could facilitate the establishment of awareness to this essential topic. Through our training we aimed to point out that teamwork and leadership is more than the transfer of hierarchical structures to the place of action.

Of course this training would be a meaningful supplement if this intervention could be adapted to training future nurses. Ideally every involved person—irrespective of his/her background and experience level—should eventually be integrated into this training. This would be a future step.

MK11: 6. Line 89: How do you define ‘team performance’ with respect to the previous hypotheses?

EFC11: See new lines 72-76.

MK12: Methods: 1. I noticed that you compared team leader training vs. ALS training. I think that comparing team leader training vs. training the whole team would have allowed to you to conclude whether there was a particular benefit of training complete teams or whether training leaders only may suffice in improving performance. What was your reason for this particular design?

EFC12: As our study was embedded in two independent editions of a two-week mandatory course for fifth-year medical students at the University Medical School in Göttingen, we had no influence on sample size—max. 28 teams per edition = 56. Thus, we had to decide on how many groups (conditions) the teams could be randomly allocated. Ideally, we should have compared three conditions: CRM training for the whole team, CRM training for designated team leaders, and a control group. Unfortunately it was not possible to organize this design due to staff shortage and course restrictions at the University Medical School. Regarding the experienced risk of such data shortage, we chose to apply a two-condition design.

Our decision to compare teams led by a CRM-trained student with teams led by an ALS-skills-trained student was based on a previous study from our group. In this study we showed the positive impact of CRM training—for the whole team—on NFT and TLV (Fernandez Castelao et al., 2011). However, the trainings were based on the 2005 ERC guidelines. Thus, in our recent study our aim was to show that CRM training—this time just for the team leader and based on the 2010 guidelines—also has a positive effect on CPR performance outcomes. To confirm our hypothesis we
needed a control condition—the ALS-add on group. This was also necessary as we could not compare the data of the two studies directly.

**MK13:** 2. Training concepts: It looks like the CRM training consisted of presentations and did not include participants actively experiencing new behavior. I doubt that leadership can be meaningfully trained without experience-based learning, that is without actually trying out the new leadership behaviors and reflecting on them afterwards. From how I understand the literature of simulation-based training, experiential learning seems to be one of its fundaments.5-7 How do you see this?

**EFC13:** We agree. The complete course consisted of one lecture, four tutorials, our training and an exam. All tutorials were simulation-based, including our videotaped CPR scenario in the tutorial 4. Additionally, the students had at least one further simulation within the tutorial, which was not taken into account for this study as we expected knowledge bias (see also our response to comment 4).

**MK14:** 3. Outcome measure and assessment: in line 185 you write that “each observable verbalization that could be clearly classified into one of the categories was documented” and that the Kappa was 0.61. That think this Kappa value indicates that the classification process was not fully “clear”, which I think is normal in this kind of research. What do you think?

**EFC14:** Yes, that is correct. Using our method our raters first had to define acts—verbal or non-verbal—and then categorize these acts into one of 44 categories. Depending on video/audio quality, line-of-sight obstruction and speech clearness, it can be complex to make a clear decision about act and category. Anyway, we discarded the term “clearly” in order to not mislead readers (line 224).

**MK15:** Results: 1. I suggest providing t-values and confidence intervals in table 2.

**EFC15:** To meet the journal’s demands we subdivided Table 2 into 2a and 2b. Additionally, we modified the contents according to the CONSORT guidelines.

**MK16:** Discussion: 1. How do you explain your findings regarding Hypothesis 4? If it is not leadership communication, what do you think accounts for the differences you found?

**EFC16:** See new lines 273-285.

**MK17:** 2. In line 225 you write “interactive CRM training”. I did not understand the interactive component. Could you clarify, please?

**EFC17:** The primary aim of the CRM-TL training was to improve leadership skills in order to facilitate effective team coordination and communication; which in turn were expected to contribute to reduced NFT rates and increased ADH.
The training consisted of a presentation on the theory of CRM, the transfer of CRM into the context of CPR, and an interactive oral exercise. Behavioral recommendations and short teaching films created by our group were included in the training: the first film simulated a suboptimal CPR process, then the second film simulated a perfect CPR process, including desirable, concise CRM team behavior visually underlined by overlays. At the end of the training, the “perfect” CPR film—without overlays—was shown again. Students were given the task of identifying relevant CRM behavior with the help of the already mentioned behavioral recommendations. Selected scenes were discussed in a plenary session (lines 138-146).

MK18: 3. In line 227 (and in line 246) you write that teams composed of one CRM-trained team leader showed lower NFT and you continue using this conclusion throughout the rest of the manuscript. Given the results you reported on p. 11, the difference between intervention and control group was statistically not significant. Thus, I do not think your conclusions are justified and I am concerned that they mislead the reader. What do you think?

EFC18: In the modified version of our manuscript we explicitly mention a non-significant difference concerning NFT rates. We also modified the Conclusion (lines 321-330) to point out that we found significant differences just with respect to ADH scores and TLV proportions.

MK19: 4. In line 229 you are referring to “higher quality verbal behavior (TLV)”. From how I understand your coding and findings, you analyzed certain communications made by team leaders and found statistically more of them in the intervention group. Thus, I think you found a higher quantity of TLV but I do not see a higher quality of TLV. How do you see it?

EFC19: All four verbalization categories we focused on having a common key function: coordination mechanisms whose rate of occurrence was positively linked to teamwork quality (lines 194-196). Thus, we summarize them as high quality, constructive communications. In the Discussion we modified the phrase (line 242).

MK20: 5. Starting in line 240 you state that “it is not the allotment of the role of a team leader and her/his training in managing the task technically that seems to be crucial for the whole team to perform well but rather explicit (i.e., CRM) training in how to manage the resources of a team in order to perform its task in a planned and well-coordinated way.” As mentioned before, I do not think that you tested this assumption (because you did not compare allotting the team leader role vs. not allotting the team leader role). Therefore, I think your conclusion is not justified by your study findings. What do you think?

EFC20: Thank you for this important indication. We modified this phrase (lines 251-258) by stressing that it is not just the role of a team leader and her/his training in managing the task technically that seems to be crucial for the whole team to perform well but rather explicit (i.e., CRM) training on how to manage the resources of a team
in order to perform its task in a planned and well-coordinated way, a point explicitly supported by Zaccaro’s model of functional team leadership.

**MK21:** 6. Line 327 – 330 – I find this section hard to understand. What is the point you are making?

**EFC21:** The message of this section (now more thoroughly explained in lines 286-297) is that each role—team leader or non-leading team member—has different requirements that are not based on hierarchy but on behaviors of equal importance for the accomplishment of the task. CPR as a task calls for specific behaviors, which can be coordinated by a team leader. Thus, we think that it is important to subdivide role related task requirements in order facilitate a stepwise understanding and clinical performance of CPR roles.

**MK22:** 7. Line 331 – 333, I think conclusion .b is not justified because you did not investigate this. What do you think?

**EFC22:** We agree and dismissed this part of the Conclusion.

**MK23:** 8. Line 362, I am curious about your reason for suggesting teaching CPR leadership separately.

**EFC23:** The three main reasons are (1) the cognitive economics such training offers by focusing on training only the team leader on leadership vs. also training the non-leading team members on leadership (lines 259-272), (2) the positive results regarding ADH scores and TLV proportions, showing that training a single team member to lead the team had an effect (see results), and (3) there is no compelling need to train individuals only as a team—i.e., to establish a team mental model—if in clinical practice teams are rarely predefined but rather build ad-hoc (external validity).

**MK24:** 9. What are the theoretical and practical implications of your study? I look forward to the revision!

**EFC24:** For practical implications see lines 286-294 and for theoretical implications see line 294-297.
Point-by-point response – Reviewer 3 (SN)

Reviewer’s report: This is an interesting article, and clearly a great deal of work and consideration has gone into the delivery of the program. The use of a randomised controlled trial methodology for an educational intervention for the workforce is also relatively novel, and also presents some challenges. I believe that the article makes a potentially useful contribution, however have some methodological concerns. I also feel that the main point of the article (ie single leader vs whole team training) gets lost in the other details and complexity of the article. I think that trying to keep this emphasis will simplify the paper and increase its value and readability. One of the challenges of performing an RCT is that the method is now so mechanised, that it is easy to spot any deviations from the correct method. It is helpful that the authors have included the CONSORT guidelines. The key finding, that training only the leader can have a large impact on team outcomes, is important.

SN1: Minor essential revisions: 1. The aim of the study needs to be included in the abstract

EFC1: First of all, thank you very much for reviewing our paper. We think your comments helped us to stay focused and will contribute to enhancing its comprehensibility. The aim of the study is now included in the abstract (lines 32-34)

SN2: The use of the word training is singular (the word trainings is used throughout the text and should be corrected)

EFC2: Corrected.

SN3: There are numerous acronyms in the document which become a bit overwhelming - it would be good if these could be reduced, and a glossary introduced for the acronyms that are retained

EFC3: Yes, that was also our first impression after reading the complete paper. But afterwards we decided to retain them in order to reduce the manuscript’s length. However, in the modified version of our paper, we added an abbreviation list at the end (line 332) and rephrased some parts in order to reduce the use of acronyms.

SN4: Do not use an acronym in the title (ie CRM). In fact, the title could be simplified to more clearly represent the comparison between the two groups.

EFC4: I see your point, but we think that CRM is a widely known acronym—especially regarding leadership and teamwork in the medical field—so that it should understandable for almost all readers. However, if you think that it is absolutely necessary to discard CRM as an acronym, we will reconsider this point.
it would be valuable to mention the resource implications of the key findings (ie training only the leader is cheaper)

Practical implications of our study involve training team leaders separately from non-leading team members for the cognitive economics of such an approach in the clinical setting; CPR team members are then focused only on their own subtasks. However, to acquire teamwork skills, an education on the skills for a leading and a non-leading role in the CPR team is essential for team functioning and should be provided to every student, especially in the early stage of their medical education. Thus, we cannot produce an argument regarding cost reduction.

There is no sample size calculation within the text. This does need to be calculated so the reader can determine whether or not the lack of statistical significance is actually a sample size problem (Type II error). This should also be mentioned in the limitations.

As our study was embedded in a mandatory course for fifth-year medical students at the Georg-August University Medical School in Göttingen, we could not influence sample size. All available students represented our sample. This is now explicitly mentioned (lines 305-317) in the limitations paragraph.

The randomisation process needs to be explained in more detail - alphabetical order is not randomisation... how did they randomise participants based on alphabetical order?

See lines 119–135 for the description of the quasi-randomization process and also some limitations regarding the randomization process (lines 305-317).

The outcomes should be described in terms of primary and secondary outcomes (and these included in the description of the hypotheses).

We defined all outcomes as primary outcomes as we see them as equally important.

The results are contextualized largely within the CRM literature. To do the findings justice, it would be valuable to incorporate wider literature on leadership training and performance, and try to draw conclusions or at least meaning for and from that body of work.

Very important suggestion, thank you! See lines 314-318.