Author’s response to reviews

Title: INTERVENTIONS TO IMPROVE TEAM EFFECTIVENESS WITHIN HEALTH CARE: A SYSTEMATIC REVIEW OF THE PAST DECADE

Authors:

Martina Buljac-Samardzic (buljac@bmg.eur.nl)

Kirti Doekhie (doekhie@eshmp.eur.nl)

Jeroen van Wijngaarden (vanwijngaarden@eshpm.eur.nl)

Version: 1 Date: 17 Jun 2019

Author’s response to reviews:

Reviewer #1:

Overall the authors have done a remarkable amount of work; searching for all the articles and reviewing all those, really a lot of work.

Nevertheless hereby some comments to improve the article:

Methods:

- Page 7: Why exclude literature reviews? You included already some literature reviews, which were known to you, but were there no other literature reviews? Please clarify this exclusion criterion.

We excluded literature reviews from our result section as we searched for articles that present unique empirical data. We included 23 literature reviews in order to search for additional articles, but did not systematically include literature reviews that resulted from the search. We did not opt for the latter as this would lead to an overwhelming additional task (due to the number and variety of reviews) next to the already broad scope of this review. We do believe that an umbrella review (i.e. a review of previous literature or meta-analyses) per intervention category would be a valuable suggestion for further research. We added the following explanations in the manuscript:

“Studies without (unique) empirical data were excluded, such as literature reviews and editorial letters. Studies regardless of their study design were included, as long as empirical data was presented. Book chapters were excluded as they are not published in peer-reviewed journals.”
“The main limitation of this review is that we cannot claim that we have found every single study per subcategory. This would have required per subcategory an additional systematic review or an umbrella review. As we noted a variety of literature reviews, future research should focus on umbrella reviews in addition to new systematic literature reviews.”

- Please clarify why you excluded "Intervention in order to improve collaboration between teams from different organizations were also eliminated" - in health care these teams can be very important, for example during large scale crises where different health care organisations are involved etc.

The reviewer is right that inter-organisational teams are very important in health care. However we had to focus our search. Especially because otherwise we also needed to include search terms such as ‘integrated care and network’ which would have resulted in an unmanageable number of hits. That is why we choose to focus on teams within organizations. Nevertheless, we do perceive this as one of the limitations and included this as follows:

“Although this review presents all relevant categories to improve team functioning in health care organizations, those categories are limited to team literature and are not based on related research fields such as integrated care and network medicine.”

- Page 8 states that 297 articles were included in the review, however page 9 states a final set of 166. Can you please clarify?

Thank you for pointing out this mistake. We included 297 articles and altered this in the manuscript.

- The text discusses the categorization of about 150 (3x5) articles, however table 2 shows 297 articles. Please elaborate.

We included 297 articles, which we categorized through three iterations. The first and second iteration included both 50 articles (a total of 100 articles). The third iteration included the remaining articles; 197 articles. We adjusted the manuscript by mentioning the numbers per iteration.
Analysis:

- In the analysis, the authors refer to "non-technical skills". Do the authors mean that includes teamwork skills or are there other skills included as well? As this term is being used throughout the article it would be good to provide a definition of what is implied with "non-technical skills" and how it is similar of different from team skills.

Non-technical skills are cognitive and social skills that complement technical skills. Team working is an example of a non-technical skill. Other examples are leadership, situational awareness, decision making, task management, and communication. The first time we mention non-technical skills, we now present a description:

“Interventions focused especially on improving non-technical skills, which refer to cognitive and social skills such as team working, communication, situational awareness, leadership, decision making, and task management [21].”

- The authors made subcategories for certain interventions. However some are not very clear, please provide definitions and explanations what are the differences, including (additional) references.

Thank you for your constructive feedback. We have included your (and the second reviewer’s) feedback as part of the last iteration in defining the categories. We described the three iterations in more detail on page 10. Based on your feedback, we have described the (sub)categories in more detail (page 11 and Table 2 Organization of Results), merged several subcategories (i.e. educational training and general team training, checklist and (de)briefing), re-categorised two subcategories (rounds and (de) briefings) and split one category (i.e. organizational interventions). We also checked all 297 articles and reorganized these under the new categories.

O For example what is the difference between an educational training and a general team building training: doesn't a general team building also have an educational goal (how to work together better) in mind?

You are very right; the difference between an educational training and general team training is unclear. Although the authors of these articles themselves refer to interprofessional education (IPE), the difference with articles that don’t is indeed not clear. Consequently, we have merged these two subcategories, adjusted the description and reviewed the 54 articles again. The latter also resulted in the repositioning articles to the subcategory “simulation”, which we therefore renamed into “simulation-based training” referring to “training with simulation as core element”.
What is the difference between a tool which improves and a tool which facilitates teamwork? If the tool only facilitates teamwork: should those articles, given the goal of the study (improvement of teamwork) have been included?

All tools were introduced to improve teamwork. However, the way in which they aim to improve teamwork differs; by structuring, facilitating, or triggering team interaction. By structuring we mean that the tool structures the way teams interact. It sets out specific steps to be taken by team members. By facilitating we mean that the tool creates circumstances/opportunities for team interaction. By triggering we mean that the tool provides information which stimulates teamwork.

We made this clear in the paper by adding the following:

“Tools are instruments that are introduced to improve teamwork by structuring, facilitating or triggering team interaction. Structuring tools standardize partly the process of team interaction. Facilitating tools provide better opportunities for team interaction. Triggering tools provide information to incentivize team interaction.”

As result of the second reviewer’s feedback, we divided the third category “organizational intervention” in two categories; “organizational redesign” and “program”, which represent the previous subcategories:

“The third category, organizational redesign refers to (re)designing structures (through implementing pathways, redesigning schedules, introducing or redesigning roles and responsibilities) that will lead to improved team processes and functioning. The fourth category, a program, refers to a combination of previous interventions (i.e. training, tools, and/or redesign).

In the results the authors show the type of results and the level of evidence. In addition it would be good to clarify the combination of the two: i.e. which have good or bad results with high level of evidence and the other way around.

We understand your point, but are struggling with the underlying assumptions addressing this comment. In table 1 we show per article what the level of evidence is. But we hesitate to draw detailed conclusions based on this in the main text. The outcomes of these studies are very diverse. Different measures are used, with different results in different studies. Therefore a meta-analyses is not possible, but even a less systematic comparison between the studies is tricky. We therefore decided to only state in general terms how many studies with high level evidence showed an intervention to have positive outcomes, without analyzing in detail which outcomes. Another reason for being careful about drawing specific conclusions is that within these high
evidence level studies it was often not clear in detail what the intervention exactly entails and how it was implemented.

- As for quality of evidence (page 11): you mention that action research is increasingly in use: what implications did that have for the quality of evidence?

Following the GRADE principles, using action research would most likely lead to (very) low quality of evidence. At the same time, action research increases insights regarding the process of implementing and tailoring an intervention. These insights are not valued/judged in the GRADE scale as it mainly focuses on the quality of evidence. We therefore added the following comment:

“In recent years there is an increasing number of studies that use an action research approach, which often creates more insight into the processes of implementing and tailoring an intervention than the more frequently used designs (e.g. RCT and pre-post surveys). However, these valuable insights are not fully appreciated within the GRADE scale.“

- Please insert a sentence on page 11 in between the overall findings and then the findings per group of articles, to introduce to the reader that you know will discuss the group of articles.

We inserted the following sentence: “The findings per category will be discussed in more detail in the following paragraphs.”

- Page 12: please clarify the sentences: "The majority of the studies claim an improvement in a number of nontechnical skills that were measures, but some also show that not all non-technical skills measured were improved [45,49,62]. Moreover, there is variation between studies in which non-technical skills were or were not significantly improved": is the second sentence not an overlap of the first? What kind of variation is alluded to here?

We clarified the second sentence as follows: “The majority of the studies claim an improvement in a number of nontechnical skills that were measured, but some also show that not all non-technical skills measured were improved [43,47,66]. Moreover, the skills that did improve differed between the studies.”

- As for "tools" what were the levels of evidence for the other studies?
We clarified this sentence as follows: “The quality of evidence varied from high to very low, and approximately one third presented a high or moderate quality of evidence.”

- As for facilitating tools, one study is very much highlighted, however the level of evidence is not reported- leaving the reader unsure how to judge this study in comparison to the other studies. We added the quality of evidence for this particular study.

- As for a few other methods the analysis of the results is unclear, f.e. under technology. We think that the reviewer means that it is sometimes not clear if the results were positive or negative. In each category we now describe the level of evidence and if the outcomes were positive or negative.

- The discussion provides a summary and comparing the different types of interventions. Only those which provide good outcomes with high evidence were mentioned, it would be good to learn which interventions were the least effective, with high evidence. This comment is strongly related to your comment “In the results…other way around.” Addressing the latter comment, we discussed this issue for the Results and Discussion section.

The discussion could be refined based on the above suggestions.

Editing:

- Abstract: many disturbing abbreviations in it. We clarified the abbreviations in the abstract.

- Please check numbering in abstract: 1=training, 2=tools, 3?= organizational. We adjusted the numbering.

- Lease first time abbreviations in full (f.e. NONTECH tool page 11, SBAR) We checked all abbreviations in the manuscript and presented them in full the first time.
Reviewer #2:

Summary: The authors conduct a systematic review of the literature on interventions in improving team effectiveness. The proposed typology follows their 2008 article and includes types of interventions, types of teams, and quality of evidence. The gap filled is a broader review with prior work being either detailed at the specific intervention or specific team or outcomes targeted. The two overarching goals are to search for empirically based interventions and to assess which ones are evidence based. Having a systematic review in this topic area is important and the authors develop a good framework and deep analysis. In particular the results of the Training section are a good synopsis on training (CRM and TeamSTEPPS) and Tools (SBAR etc). The implication section details major gaps which is very well developed and important for the literature. The "limited innovation" could be refined as the gray literature and other areas likely cover this - consider refining or dropping. Overall the review is well done and important to the literature. Some refinements as listed below are necessary for the paper to be accepted.

We agree; we dropped the term “limited innovation” and rephrased the following sentences and positioned it under “limitations” instead of “implications for future research”:

“In addition, the combination of the publication bias and the exclusion of grey literature has probably resulted in a main focus on standardized interventions and a limited range of alternative approaches which does not necessarily reflect practice.”

Major

- It is unclear if the interventions included are targeting only effectiveness as stated in the abstract or general performance outcomes. There is a incongruency between the stated objectives in the abstract and the search terms employed in the search strategy (p 6), the former being specific, the latter being more general.

Our aim was to present interventions that could potentially improve team performance in general, which will lead to team effectiveness in the end. In our search strategy we focused on a broad range of performance outcomes that are related to team effectiveness and/or will lead to team effectiveness. We use the term “team effectiveness” broader than is common in the OB field, as you already suggested in the comment below. To address this issue:

“Studies were excluded that did not include a team intervention or that included an intervention that did not primarily focus on improving team performance, which is likely to enhance team effectiveness (or other related outcomes).”

“There is a widespread belief that effectiveness of health care teams can be improved by team interventions, as a wide range of studies have shown a positive effect of team interventions on
performance outcomes (e.g. effectiveness, patient safety, efficiency) within diverse health care setting (e.g. Operating Theatre, Intensive Care Unit, Nursing Homes) [3-7]."

“What types of interventions to improve team effectiveness (or related outcomes) in health care have been researched empirically, for which setting, and for which outcomes (the last decade)?”

- If the goal is to broadly study the interventions on team effectiveness, the authors might expand their "team" search term to include other team-based terms, i.e., Microsystems, pods, groups, patient centered medical home, etc.

Expanding our review by including additional search terms would result in including other research areas such as integrated care, patient-centered care, and network medicine. This would result in an unmanageable number of hits and difficulty in synthesizing the results. Nevertheless, we do perceive this as a limitation and included this in the Limitation paragraph:

“Although this review presents all relevant categories to improve team functioning in health care organizations, those categories are limited to team literature and are not based on related research fields such as integrated care and network medicine.”

- In the background section include a definition of key terms, i.e., teams, effectiveness, PRISMA. In particular more detail is needed on "effectiveness" since the overall findings appear to be broader than the strict use of effectiveness in the Org Behavior literature.

We have addressed the term “team effectiveness” in more detail as follow:

In the literature there is a lot of inconsistency about what a team is and what a “real” team is. This resulted in dozens of definitions of a team. We believe that choosing one definition would be challenging, especially in the health care field as this field is characterized by a high variety of teams. As most reviews on teamwork, we reason in line with the included articles. If articles claim that there study setting are health care teams, we support this:

“In addition, how teams were defined was not selection criterion. Due to the variety of teams in the health care field, we find it acceptable if studies claims that the study setting consists of health care teams.”

- How was the organization into subcategories conducted (p 9,10). The process is not clear and stated generally, more detail is needed.
“The categorization of our final set of 297 articles is the result of three iterations. First, 50 summarized articles were categorized using the initial categorization: team training (subcategories: CRM based training, simulation training, interprofessional training, and team training); tools; and organizational intervention [8]. Based on this first iteration, the main three categories (i.e., training, tools, and organizational interventions) remained unchanged but the subcategorization was further developed. Training related to the subcategory “CRM based training”, “TeamSTEPPS” was added as a subcategory. The other subcategories (i.e., simulation training, interprofessional training, and team training) remained the same. Tools, the first draft of subcategories, entailed SBAR, checklists, (de)briefing, and task tools. Two subcategories of organizational intervention (i.e., program and (re)design) were created, which was also in line with the content of this category in the original literature review. Second, 50 additional articles were categorized to test and refine the subcategories. Based on this second iteration, the subcategories were clustered, restructured and renamed, but the initial three main categorizations remained unaffected. The five subcategories of training were clustered into principle-based training, method-based training, and general team training. The tools subcategories were clustered into structuring, facilitating, and triggering tools, which also required two new subcategories: rounds and technology. Third, the remaining 197 articles were categorized to test the refined categorization. In addition, the latter categorization was peer reviewed. The third iteration resulted in three alterations. First, we created two main categories based on the two subcategories “organizational redesign” and “programme” (of the third main categorization). Consequently, we rephrased “programme-based training” into “principle-based training”. Second, the subcategories “educational intervention” and “general team training” were merged into “general team training”. Consequently, we rephrased “simulation training” into “simulation-based training”. Third, we repositioned the subcategories “(de)briefing” and “rounds” as structuring tools instead of facilitating tools. Consequently, we merged the subcategories “(de)briefing” and “checklists” into “(de)briefing checklists”. Thereby, the subcategory “technology” became redundant.”

- The third aspect of the categorization (organizational interventions (programs and redesign) are not clear. Is the intent to go from micro to macro, so organizational impacts, are the first two categories not also macro? Providing a clearer definition for this category is necessary.

To address this comment we divided this main category into two main categories instead of two subcategories within one main category. Consequently, we have rephrased the description of them:
Organizational redesign:

“The third category, organizational redesign refers to (re)designing structures (through implementing pathways, redesigning schedules, introducing or redesigning roles and responsibilities) that will lead to improved team processes and functioning.”

“In contrast with the previous two categories, organizational (re)design focuses on intervening in structures, with the aim of consequently improving processes.”

Programme:

“The fourth category encounters a combination of the previous categories. A program most frequently consists of a so-called Human Resource Management bundle that combines learning and educational sessions (e.g. simulation training, congress, colloquium), multiple tools (e.g. rounds, SBAR), and/or coordination intervention (e.g. meetings, standardization). Moreover, a program frequently takes the organizational context into account as it explicitly creates room to maneuver; developing an improvement plan and making choices tailored to the local situation.”

In addition, to check if the articles match the new description, we reviewed the 35 articles again and repositioned 4 articles.

Minor
- Abstract Results Section: Three types of interventions are set up then only 2 are listed.

We have added the Abstract and clearly pointed out the third category.

- Grammar mistakes and awkward sentence structures are common throughout the article.

The editing service of American Journal Experts has edited the revised manuscript.

- If empirical and evidence based interventions are the focus, why exclude students - often studies use students as a way to control the environment and improve reliability and validity?

We excluded articles that base their evidence on student samples because the ultimate evaluation is in a specific health care setting and student settings are often used in the development stage of an intervention. Although we see the added value of these articles, we also wanted to complete
this follow up review in line with the initial review as much as possible. This exclusion criteria was also used in the initial review.

- What is the PRISMA format (p 8).

PRISMA refers to Preferred Reporting Items for Systematic Reviews and Meta-Analyses. It is a minimum set of items for reporting in systematic reviews, which is shown in Figure 1. We clarified this better in the manuscript.

- Consider renaming the 3rd category (interventions) on p 10 since this is the overarching goal of the review and is confusing compared to trainings and tools. Consider adding a reason for why interventions fall into trainings, tools, and 3rd category.

We understand the confusion. Based on the previous comments of both reviewers, we have slightly restructures the categorization. We divided the third category “organizational intervention” in two categories; “organizational redesign” and “program”, which represent the previous subcategories:

“The third category, organizational (re)design, refers to (re)designing structures (through implementing pathways, redesigning schedules, introducing or redesigning roles and responsibilities) that will lead to improved team processes and functioning. The fourth category, a programme, refers to a combination of the previous types of interventions (i.e., training, tools, and/or redesign).”

In order to check if the articles match the new description, we reviewed the 35 articles again and repositioned 4 articles.

- P 17 - at the top the description of what the organizational interventions are is not clear . . . operational level would include structure and process?

As result of restructuring the categorization, we rephrased that paragraph.

- Limitation section uses colloquial language (from scratch, found every single study).

The editing service of American Journal Experts has edited the whole manuscript.