BMJ Open

Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? A pragmatic, prospective and randomized study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-015977
Article Type:	Protocol
Date Submitted by the Author:	16-Jan-2017
Complete List of Authors:	Freytag, Julia; Charite Universitatsmedizin Berlin, Simulated Patients Program, Office of the Vice Dean for Teaching and Learning Stroben, Fabian; Charite Universitatsmedizin Berlin, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning; Inselspital, University of Bern, Department of Emergency Medicine Hautz, Wolf; Inselspital, University of Bern, Department of Emergency Medicine Eisenmann, Dorothea; Charite Universitatsmedizin Berlin, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning; Charite Universitatsmedizin Berlin, Department of Anesthesiology and Operative Intensive Care Medicine CCM & CVK Kämmer, Juliane; Max-Planck-Institut fur Bildungsforschung, Center for Adaptive Rationality; Charite Universitatsmedizin Berlin, Progess Test Medizin
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Emergency medicine, Communication
Keywords:	MEDICAL EDUCATION & TRAINING, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < ANAESTHETICS

Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? A pragmatic, prospective and randomized study

Version 1.0 / 13.1.2017 15:00h

Julia Freytag, M.Sc.(1)*; Fabian Stroben, MD (2,3)*; Dr. med. Wolf E Hautz, MME (3); Dr. med. Dorothea Eisenmann (2,4); Dr. rer. nat. Juliane E Kämmer (5,6)

*contributed equally

(1) Simulated Patients Program, Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Germany
(2) Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Germany
(3) Department of Emergency Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland
(4) Department of Anesthesiology and Operative Intensive Care Medicine CCM & CVK, Charité Medical School Berlin, Germany
(5) Progess Test Medizin, Charité Medical School Berlin, Germany
(6) Max Planck Institute for Human Development, Center for Adaptive Rationality, Berlin, Germany

Julia Freytag, M.Sc., Research Associate, Simulated Patients Program, Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany
Fabian Stroben, MD, Research Associate, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany & Department of Emergency Medicine, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse 4, 3010 Bern, Switzerland
Dr. med. Wolf E Hautz, MME, Consultant, Department of Emergency Medicine, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse 4, 3010 Bern, Switzerland
Dr. med. Dorothea Eisenmann, Head of Lernzentrum, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany & Department of Anesthesiology and Operative Intensive Care Medicine CCM & CVK, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany
Dr. rer nat. Juliane E Kämmer, Head Research Scientist, Progess Test Medizin, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany & Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany

Correspondence to:
Fabian Stroben, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany
Phone: +49 (0) 30 450 576 403
Fax: +49 (0) 30 450 576 922
fabian.stroben@charite.de

Abstract

Introduction: Medical errors have an incidence of 9% and may lead to worse patient outcome. Training teamwork has the capacity to significantly reduce medical errors and therefore improve patient-outcome. One common framework for training teamwork are the Crisis-Resource-Management trainings, adapted from aviation and usually trained during simulation. Debriefing after simulation is the crucial part for enhancing learning but it remains unclear how to best debrief. Furthermore, teamwork trainings and studies examining education effects on undergraduates are rare. The study aims to evaluate the effects of two teamwork-focused debriefings on team performance of undergraduate medical students.

Methods and analyses: A prospective-experimental study compares a well-established three-phase debriefing method (gather – analyse – summarize; *GAS-model*) to a newly developed debriefing approach consisting of the *GAS-model plus TeamTAG* (Teamwork Techniques Analysis Grid). TeamTAG is a cognitive aid with preselected teamwork principles and behavioural anchors for observable patterns of teamwork, which may help structuring debriefing. Both debriefing methods will be tested during an emergency room simulation comprising six emergency medicine cases faced by 35 last-year medical students in five-member teams. Teams will be randomized into the two debriefing conditions. Team performance during simulation and the number of discussed principles during debriefing will be evaluated. Furthermore, learning opportunities, helpfulness, and feasibility will be rated by participants and instructors. Analyses include descriptive, inferential and explorative statistics.

Ethics and dissemination: The study protocol was approved by the institutional office for data protection and the ethics committee of Charité Medical School. All students participate voluntarily and sign an informed consent after written and oral information about the study. Results will be published in peer-reviewed journals and discussed on scientific meetings. The study was registered with the office of data privacy and the IRB at Charité Medical School under EA2/172/16, because a register for educational studies does not exist.

Strengths and limitations of the study.

- The study design builds on established principles of teaching and assessing teamwork
- One of the first experimental studies of effects of teamwork-focused debriefing on team performance with undergraduate medical students
- Study is embedded in a well-established simulation setting with proven efficacy.
- Pragmatic comparison of two debriefing methods
- Only single centre study.

1. Introduction

Medical errors and adverse events occur with an incidence of about 9% and can harm patients seriously.(1,2) Error rates in settings like emergency rooms are even reported to be twice as high.(3–5) Most medical errors originate from human factors and teamwork(6) or medication errors(7) and about half of all medical errors are considered preventable.(1,7)

Empirical evidence(6,8–11) suggests that improving teamwork may be key to reduce medical error. However, although teamwork and patient safety are prominent objectives in many national outcome frameworks,(12–14) these topics are insufficiently represented in undergraduate education and are rarely assessed, although validated assessment tools exist.(15,16) As a consequence, about 60% of junior doctors in Germany report feeling inadequately prepared for clinical practice(17) and almost half of the residents of a Canadian survey reported feeling overwhelmed leading a resuscitation team.(18) In addition, common interventions targeting the quality of teamwork and human factors such as simulation training and Crisis-Resource-Management (CRM) trainings have demonstrated highly variable effects.(19) In both, simulation and CRM trainings, debriefing is considered to be the crucial part to enhance learning(20) but little is known about how to best debrief. In fact, the highly variable effects of simulation may very well result from differences in debriefing. A feasible and beneficial debriefing method, particularly for undergraduates, may lead to more effective simulation sessions and thus ease the transition into clinical practice for junior doctors, and may ultimately lead to a reduction of medical errors and thus improved patient outcome. This study thus aims to compare the effects of two debriefing methods on team performance and the acquisition of teamwork skills.

Training and Debriefing

Common training methods that address teamwork and human factors are simulations in general and CRM trainings in particular. The concept of CRM was originally derived from safety trainings in aviation and has been adapted to the health care sector, which also represents a high stakes environment.(21) The idea of CRM is to guide individuals and teams in emergency situations (crises) to use all available resources to manage the situation effectively – and thus to prevent critical incidents in the first place. CRM-trainings have been shown be a potent tool to improve teamwork and – as a consequence – patient safety.(22) In our study, elements of CRM set the framework for training and debriefing teamwork during an emergency room simulation.

Debriefing is defined as a bi-directional and interactive way of discussion after simulation to reflect on action and to analyse performance.(20) Within debriefing concepts, feedback is one central process element and often used as a conversational technique especially in participants with little experience in debriefing.(23) Feedback is defined as delivery of information to improve reasoning or behaviour compared to defined performance standards,(23,24) and it is critical to improve learning.(20) How to best integrate feedback into debriefing, which specific aspects to address and how to structure

debriefing to best foster learning is, however, still unknown.(20,25) The goal of this study is thus to compare a well-established debriefing method to a more structured and feedback-focused method to evaluate effects on teamwork, learning opportunities, feasibility and helpfulness for participants (and instructors). The two debriefing methods we will focus at are the following:

(1) **The GAS-method**: This debriefing method consists of the three parts **g**ather, **a**nalyse and **s**ummarize.(26,27) The GAS-method is one of many similar three-steps debriefing structures(23) and is used, for example, in simulation courses of the "American Heart Association" (AHA).(27) During the first phase (*gather*), participants are encouraged to reflect-on-action and establish a shared mental model, which can then be used to discuss the simulation in a learner-centred way (*analyse*). During this process, questions tailored towards specific learning objectives are used to induce learning effects. Finally, the debriefing is summed up and critically reviewed by the team and its instructor (*summarize*).(23,26) Topics that are discussed during the debriefing using this method are mostly self-selected by the team and instructor. Thus, their relevance to enhancing teamwork largely depends on the experience of the instructor.

(2) **The GAS-method plus** a cognitive aid: This newly developed debriefing method uses the GAS-structure and additionally offers the instructors a cognitive aid to structure the debriefing. Cognitive aids are "structured pieces of information designed to enhance cognition and adherence to […] best practices."(28) Cognitive aids have been shown to be beneficial in different areas of medicine.(29–31) Moreover, cognitive aids are useful for debriefing: Using a cognitive aid as instructor may improve the acquisition of behavioural and cognitive outcomes of participants after simulation – especially so with novice instructors.(32) In practice, such aids are often a pocket card, a script or a poster.

We will specifically use a guideline called "TeamTAG" (Teamwork Techniques Analysis Grid) to serve as a cognitive aid to foster observation and feedback relevant to teamwork. In detail, the guideline aims to (a) structure the feedback process during debriefing, and (b) serve as a reminder of what to address during the *analyse*-step of the GAS-method. The TeamTAG lists teamwork-relevant CRM-principles together with behavioural anchors serving as directly observable patterns of teamwork (see below and digital supplementary information).

Hypotheses

First, we assume that debriefing based on the *GAS-method plus TeamTAG* is a more effective tool than the common *GAS-method* alone, and will lead to the discussion of more teamwork-relevant principles. Debriefing using the *GAS-method plus TeamTAG* should thus result in more learning opportunities for teams and ultimately in improved team performance. This hypothesis is based on the fact that the TeamTAG is concise and guides observation and feedback with practical examples. Using these examples during observation may help focus the observers' attention(33) and result in the team

discussing more teamwork-relevant CRM-principles. In undergraduate education, instructors are often novices and vary considerably regarding their experience in debriefing. Because novices were shown to benefit more from structured debriefing scripts than more experienced instructors,(32) we consider our environment (see methods section) ideal to detect differences between both debriefing methods if they exist.

H1a: We expect that the participants who receive debriefing based on the *GAS-method plus TeamTAG* show a greater improvement in team performance than those discussing the simulation according to the common *GAS-method* alone.

H1b: We expect that the participants who receive debriefing based on the *GAS-method plus TeamTAG* report a higher number of CRM-principles discussed.

Second, we expect that teams receiving debriefing based on the *GAS-method plus TeamTAG* perceive teamwork skills as more important, which should increase their sensitivity towards a culture of safety and the likelihood for inducing changes in behaviour.(34,35) Moreover, perceiving the content of the debriefing as more important should lead to a higher overall satisfaction with and perception of helpfulness of the debriefing.

H2a: We expect that the participants who receive debriefing based on the *GAS-method plus TeamTAG* report a higher level of perceived importance of teamwork principles than those discussing the simulation according to the common *GAS-method*.

H2b: We expect that the participants who receive debriefing based on the *GAS-method plus TeamTAG* report a higher satisfaction with and helpfulness of the debriefing they received.

Third, we focus on the satisfaction of the instructors as a measurement of feasibility and efficiency. We expect higher satisfaction when using the *GAS-method plus TeamTAG* as it may facilitate a more structured feedback and gives a better opportunity for instructors to address the learning objectives of their participants.

H3: We expect that instructors using the *GAS-method plus TeamTAG* report a higher feasibility and efficiency of their debriefing.

2. Methods and analysis

This investigation is designed as a prospective-experimental superiority study with intervention and control group, receiving either debriefing during a simulation training based on the *GAS-method plus TeamTAG* or based on the *GAS-method* alone, respectively. It will be executed during an Emergency Department (ED)-simulation at Charité Medical School, Berlin, Germany, on 14th January, 2017. The simulation has been implemented at the local skills lab since 2013 on a peer-led basis. Main goal of this extensive, 8-hour simulation-training during the night is to give students the opportunity to experience being the person in charge of a patient's healthcare. This event takes place once a year, with about 35 students in their final year of medical studies participating voluntarily. Participants will be recruited via newsletter and advertising posters. The students act in randomly assigned teams of five and self-select into different roles (team leader, team member, observer), which they switch during the night. Simulated patients and high-fidelity simulators are used to create realistic case simulations; simulated radiologic and laboratory-service is provided. One of the main goals of the event is to improve students' confidence in working with medical emergencies in an ED over the course of the night.(36) The simulation was awarded with a project prize by the German Association for Medical Education (GMA) in 2016.

Each student team has to work on six simulated cases. Each case is staffed with a case instructor, who is responsible for the simulation and provides technical help. Each student team is accompanied by a group instructor who guides the participants during the night. After every case, feedback is provided by simulated patients, observing participants and case instructors as a multi-source feedback. As part of our study, in 2017, participants will additionally receive a teamwork-based debriefing by the group instructors after every case in one of two conditions (*GAS-method* vs. *GAS-method plus TeamTAG*).

Development of the TeamTAG as cognitive aid

As a basis for this study, the TeamTAG guideline was developed with the goal to have a feasible and time-efficient feedback instrument that supports teaching the basic teamwork skills to participants. Two investigators (JF and FS) developed the TeamTAG guideline that presents six common CRM-principles,(21,37) each accompanied by the description of behavioural anchors. The six chosen CRM-principles are: (1) anticipate and plan ahead, (2) set priorities dynamically, (3) call for help early, (4) exercise leadership and followership, (5) communicate effectively, and (6) re-evaluate repeatedly. The TeamTAG can be found as a digital supplementary information. The CRM principles and its behavioural anchors were chosen by the following criteria: a) simulation setting, b) presumed skills of participants, c) experience of instructors and d) observability. The tool was reviewed and adjusted by an experienced group of anaesthesiologists, emergency medicine physicians, simulation instructors and peer tutors, all experienced in medical education and simulation-based learning. In a pre-study,

feasibility for instructors was examined (see preliminary results below) but not compared to an approach without TeamTAG.

Team performance Measurement

For measuring team performance, the "Teamwork Emergency Assessment Measure" (TEAM) will be used. TEAM was developed by Cooper et al., is constructed as an assessment tool and has been applied to both clinical and simulation environments.(15,16,38) It consists of eleven items belonging to the three subscales leadership, teamwork, and task management. Example items are, "the team leader maintained a global perspective" and "the team prioritized tasks," measured with Likert-Scales.

As there was no German version of the TEAM, the English version was translated into German, using elements of the TRAPD methodology(39) (translation, review, adjudication, pre-test, documentation). Two investigators (JF and FS) independently translated the TEAM into German in parallel, reviewed the results and consented to one version, which was translated back by a native English speaker. This new version was compared to the original TEAM and consented by both investigators and the native speaker. All steps of the translation were documented.

After translating TEAM, a rater training was developed. The training involves three aspects, which are important to prepare raters to accurately assess a certain behaviour or skill(40): a *rater error training* in which information is provided on typical rating errors to raise awareness and prevent them, a *performance dimension training* to inform about the targeted dimensions, including definitions and videotaped examples, and a *frame of reference training*, in which videotaped examples of different quality are assessed and discussed. All raters who are responsible for TEAM-ratings in this study (case instructors and additional raters) will receive this rater training and additional written material on teamwork and how to use TEAM.

Group instructors debriefing training

Before data collection, all group instructors receive a teamwork-related training and additional written material with information about how to provide feedback and conduct debriefings, about human factors in general and CRM in particular, which is intended to serve as a framework for all teamwork aspects during debriefing. The training includes videos showing good and bad examples of teamwork and is followed by discussions about opportunities for debriefing in these specific situations (adapted from *frame of reference training*(40)). After this training, which is equal for all group instructors, the instructors will be randomized, stratified by level of academic education and additional professional training (e.g., nurse or paramedic), into the two conditions. Both groups will be instructed by the investigators separately: The intervention-group instructors discuss the „TeamTAG" and are instructed to at least focus on every CRM-principle of TeamTAG once during the first five cases and re-evaluate their previous focus of debriefing after the next simulation if behaviour does not change sufficiently

from instructor's perspective. The control-group instructors are advised to give feedback regarding whatever teamwork-related aspect they deem important during the first five cases and also to re-evaluate the teamwork if needed.

Procedure of data collection

Upon arrival, every student participant defines an individual anonymized study code, which has to be entered on every form and questionnaire and allows linking all measurements during the course of the night. Students also track their role (leader, member, observer) after every case to allow sub-group analyses in relation to these roles. Figure 1 depicts the procedure of data collection during the night shift simulation.

Fig. 1.

Fig. 1. *Study flowchart; GAS = Gather – Analyse – Summarize. CRM = Crisis Resource Management. TEAM = Team Emergency Assessment Measure. TeamTAG =Teamwork Techniques Analysis Grid, R = randomization*

Before starting the simulation, all 35 participants fill in the first questionnaire, which assesses possible confounders such as demographic data, professional training as nurse or paramedic or any training in teamwork/human factors. Next, students will be randomized into seven groups via a computer-generated algorithm by the principal investigator. Four groups serve as intervention group, the remaining ones as control – without knowing in which condition they are. After randomization, all groups gather separately and are asked to discuss already known principles of teamwork and 15 multiple choice questions concerning emergency medicine. A recent study showed that such discussions' results are linked to team performance.(41)

During the simulation, all groups face six simulations where teamwork is measured and teamwork-related feedback is provided. All cases depict common emergency situations where participation of an emergency team in the emergency room is needed. Table 1 gives a brief overview of diagnoses of the six cases and challenges for teamwork.

Table 1.

During every case, team performance will be measured using the "Team Emergency Assessment Measure" (TEAM),(16) which will be filled in by the case instructors and an additional rater. The two TEAM-raters are blinded regarding the debriefing condition the group is assigned to.

After every case (duration about 30min), debriefing starts (about 20min) with checklist-based feedback by the simulated patients (focus: communication skills, empathy) and by the case instructors and peer-observers (focus: factual knowledge, diagnostic skills). As last part of the debriefing process, the teamwork-related debriefing is conducted by the group instructor based on the GAS-method with or without the support of the „TeamTAG" depending on the experimental condition.

After the debriefing process, all group members evaluate the case and rate how helpful the debriefing was. Group instructors in both groups track the main topics of their teamwork debriefing and remind participants to fill in their evaluation forms.

Right after the last case of the night, all participants fill in a final evaluation, which asks to list all CRM-principles on which they received feedback during the night. Furthermore, participants evaluate the importance of each principle for their future work as physicians and provide a general evaluation of the night. Every group tutor fills in a feedback evaluation form, which asks for feasibility of, efficiency of and difficulty with providing feedback.

Collected data

(1) Baseline characteristics: Collected data of the first questionnaire, group discussions and teamwork-discussion results are used to compare the baseline between the two conditions. Discussion results will be analysed qualitatively to identify differences in knowledge and in the personal definition of good teamwork at the beginning of the night. Furthermore, TEAM-scores during the first simulation case serve as baseline team performance.

(2) Hypothesis 1 measurement [team performance, number of discussed CRM-principles]: Team performance is evaluated using the eleven items of the translated TEAM on a 5-point Likert scale (0 = never, 4 = always). Similar to previous studies,(15,16,38,42,43) we will analyse ratings on the item level, the sum score (range 0 to 44), and an overall rating as a single item with a scale from 1-10. All teamwork principles discussed will be tracked after every debriefing by the group instructors. Furthermore, participants state which principles were discussed during the night after completing the last simulation debriefing.

(3) Hypothesis 2 measurement [importance, satisfaction, helpfulness]: Estimated relevance of the learned CRM-principles and overall satisfaction with the simulation are evaluated on 7-point Likert scales at the end of the night. Helpfulness of debriefing from the different providers (simulated patient, peer, case tutor and group tutor) is rated by participants after every case on a 7-point Likert scale.

(4) Hypothesis 3 measurement [instructor ratings]: Debriefing evaluation of the group instructors (feasibility of, efficiency of and difficulty with providing feedback) is measured with 7-point Likert scales and as free-text answers at the end of the night.

(5) Other measures: The general evaluation form asks for pleasure, quality of instruction during the night, difficulty of cases, possibility of applying knowledge on 7-point Likert scales.

All Likert scales are coded from +3 “strongly agree” to -3 “strongly disagree.” All data collection forms are available upon request.

Analyses

Data will be analysed in SPSS and R using descriptive, inferential and explorative statistics. We conducted a calculation of power for our primary research question (team performance). Recent studies, reporting mainly data for well-trained and experienced teams, show TEAM-sum-scores up to 40.(42,43) Only one study provides data for less experienced teams with a TEAM-sum-score of 21.(42) Based on these results and data from a pre-study (see below), we expect a TEAM-sum-score of about 20 for untrained team and a score of around 40 for teams receiving a training related to teamwork skills and/or having a lot of experience in this area). These scores imply a potential increase due to training of up to 20 point on the TEAM-sum-score. As a relevant training effect for a single training event like ours, we estimate a gain in TEAM-sum-score of eleven points (equals one point per item). Using the standard deviation from the last published study on TEAM(43) (SD = 4.4) and $\alpha < 0.05$, about 6 teams are needed to detect a significant difference between the conditions with a power of 80%. Missing data are handled using pairwaise deletion.

(1) Baseline characteristics: Discussion results will be compared between intervention and control group using qualitative methods and confounder analysis (demographics, prior training) with parametric and non-parametric tests for testing equivalence. TEAM-Scores (single items, sum score, overall score) from the first simulation case will be compared using multi-level analyses.
(2) Analyses for hypothesis 1: TEAM-Scores (single items, sum score, overall score) will be compared between intervention and control group during the sixth simulation case using multi-level analyses. The number of discussed CRM-principles will be compared between control and intervention group by using a multi-level model to take the hierarchical structure of data into account. The development of team performance over the course of the night will be analysed using descriptive statistics and plotting "training curves" per team.
(3) Analyses for hypothesis 2: The participants' ratings of the feedback's helpfulness, the importance of CRM-principles and satisfaction with the debriefing will be compared between control and intervention group by using multi-level models to take the hierarchical structure of our data into account.
(4) Analyses for hypothesis 3: Group instructors' evaluations of the instrument will be examined descriptively.
(5) Other measures: The general evaluation will be examined in a descriptive way.

Data sharing statement

Data analysis will be conducted by the investigator's team (data management team). As the study is not a clinical trial, a data monitoring team is not needed. Anonymised full data set will be published together with the journal publication or using the "Dryad Data Repository" (Durham, NC, USA) as

required by the journals guidelines. Data will furthermore be stored in the local data repository at Charité Medical School Berlin according to the local guideline for good scientific practice.

3. Preliminary results

Validation of the German TEAM

The German TEAM can be found as a digital supplementary information. As a preliminary validation, interrater-correlation was checked between three investigators (JF, FS and DE) and an external expert on two video-taped resuscitations. Both resuscitations were simulation based and had similar factual content; however, the first simulation showed good teamwork and the second intermediate teamwork performance. The videotaped simulations were used for group instructors' debriefing trainings and for validity-testing of the German TEAM.

Intraclass correlation coefficients were .99 for the first resuscitation (Mean TEAM-score = 42.3, SD = 1.3) and .85 for the second one (Mean TEAM-score = 22.5, SD = 3.1), which indicates excellent interrater agreement. For this reason, we consider the German TEAM as a valid instrument to assess team performance in our study.

TeamTAG

A first version of TeamTAG was used in a pre-study, conducted during the previous simulated night shift in 2016. In this pre-study, all instructors (N=7) used TeamTAG as part of their debriefing (similar to GAS-method plus TeamTAG). They were asked to rate the feasibility and helpfulness of the TeamTAG (7-point Likert scale; -3/+3). Furthermore they could comment on specific aspect of the guideline they liked / disliked (free-text answers). All participants were asked how useful the instructors' feedback was (7-point Likert scale; -3/+3).

Instructors rated the guideline as a feasible tool (M=1.9, SD=0.9) and stated, that it helped them in both observing and giving feedback to the participants of the simulation ($M_{observe}$ = 2.3, SD=0.8; $M_{feedback}$ = 2.3, SD=0.5). The participants declared to have found the feedback to be useful (M = 1.7, SD = 1.0).

4. Ethics and dissemination

The study protocol was designed according to the „Declaration of Helsinki“, the local guideline for good scientific practice at Charité Medical School Berlin and the ICMJE-Recommendations. The study protocol was approved by the institutional office for data protection (AZ 737/16) and the ethics committee at Charité Medical School Berlin (EA2/172/16).

All participants and instructors will provide informed consent. Due to the fact that the simulation is already a well-known event at Charité Medical School Berlin and receives official teaching-funds, participants who refuse to take part in the study must have a chance to participate nevertheless. In this case, students will not provide the informed consent prior to randomization; instead an independent „no-study“ group will then be created, which is identical to the control group but without any teamwork debriefing. We do not expect any harm for students who undergo intervention.

Publication

Results of the study will be presented during national and international scientific meetings. The authors aim to publish all results in a peer-reviewed journal. Anonymised full data set will be published together with the journal publication or using the “Dryad Data Repository” (Durham, NC, USA) as required by the journals guidelines. Data will furthermore be stored in the local data repository at Charité Medical School Berlin according to the local guideline for good scientific practice.

5. Author contributions

JF and FS translated TEAM, designed the study and are responsible for conduction. DE, WEH and JEK contributed to the study design. JEK supervises study design and conduction. JF and FS are responsible for data analyses. JF, FS and JEK wrote the manuscript. JF and FS conducted the pre-study. DE is responsible for funding, local administration at Charité Medical School Berlin and heads the steering committee. All authors carefully read the manuscript, made critical and substantial revisions and gave their approval for publication.

6. Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. The investigators JF and FS are funded by the German „Federal Ministry of Research and Education (BMBF)“. The sponsor did not interfere with the conception and conduction of the study, data analysis or production of the manuscript.

7. Competing interests

WEH received financial compensation for educational consultancy from the AO Foundation, Zurich, Switzerland. All other authors report no competing interests.

8. Acknowledgements

The authors would like to acknowledge Hanno Heuzeroth and David Steinbart for support in conducting the pre-study and Simon Cooper for support in using and translating “TEAM.” Furthermore, the authors thank all instructors and physicians, especially Tobias Deselaers, for organising the simulation and their willingness to participate as raters and instructors etc.

9. References

1. Vries EN De, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. Qual Saf Heal Care. 2008;17:216–23.
2. Linda T. Kohn, Janet M. Corrigan and MSD. To Err Is Human. Building a Safer Health System, Volume 6. International Journal of Public Health. 1999. 93-95 p.
3. Berner ES, Graber ML. Overconfidence as a Cause of Diagnostic Error in Medicine. Am J Med. 2008;121(5A):S2–23.
4. Schwartz A, Elstein AS. Clinical reasoning in medicine. Clin Reason Heal Prof. 2008;223–34.
5. Hautz SC, Schuler L, Kämmer JE, Schauber SK, Ricklin ME, Sauter TC, et al. Factors predicting a change in diagnosis in patients hospitalised through the emergency room: a prospective observational study. BMJ Open. 2016;6(5):e011585.
6. Risser DT, Rice MM, Salisbury ML, Simon R, Jay GD, Berns SD. The Potential for Improved Teamwork to Reduce Medical Errors in the Emergency Department. Ann Emerg Med. 1999;34(3):373–83.
7. Leape LL. Error in medicine. JAMA. 1994;272(23):1851–7.
8. Hautz WE, Kämmer JE, Exadaktylos A, Hautz SC, Kaba E. How thinking about groups is different from groupthink. Med Educ. 2016;ePub ahead:13137.
9. Hautz WE, Kämmer JE, Schauber S, Spieß CD, Gassmeier W. Diagnostic Performance by Medical Students Working Individually or in Teams. JAMA. 2015;313(3):303–4.
10. Knaus WA, Draper EA, Wagner DP, Zimmermann JE. An Evaluation of Outcome from Intensive Care in Major Medical Centers. Ann Intern Med. 1986;104(3):410–8.
11. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. Acta Anaesthesiol Scand. 2009;53:143–51.
12. Frank J, Snell L, Sherbino J, editors. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.
13. Kaba A, Wishart I, Fraser K, Coderre S, Mclaughlin K. Are we at risk of groupthink in our approach to teamwork interventions in health care? Med Educ. 2016;50(4):400–8.
14. Kiesewetter J, Gutmann J, Drossard S, Gurrea Salas D, Prodinger W, Mc Dermott F, et al. The Learning Objective Catalogue for Patient Safety in Undergraduate Medical Education – A Position Statement of the Committee for Patient Safety and Error Management of the German Association for Medical Education. GMS J Med Educ. 2016;33(1):1–14.
15. Cooper S, Cant R, Connell C, Sims L, Porter JE, Symmons M, et al. Measuring teamwork performance: Validity testing of the Team Emergency Assessment Measure (TEAM) with clinical resuscitation teams. Resuscitation. 2016;101:97–101.
16. Cooper S, Cant R, Porter J, Sellick K, Somers G, Kinsman L, et al. Rating medical emergency teamwork performance: Development of the Team Emergency Assessment Measure (TEAM). Resuscitation. 2010;81(4):446–52.
17. Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011 Jan;11(1):99.
18. Hayes CW, Rhee A, Detsky ME, Leblanc VR, Wax RS. Residents feel unprepared and unsupervised as leaders of cardiac arrest teams in teaching hospitals: A survey of internal medicine residents. Crit Care Med. 2007;35(7):1668–72.
19. Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA. Linking Simulation-Based Educational Assessments and Patient-Related Outcomes : A Systematic Review and Meta-Analysis. Acad Med. 2015;90(2):246–56.
20. Cheng A, Eppich W, Grant V, Sherbino J, Zendejas B, Cook DA. Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. Med Educ. 2014;48:657–66.
21. Gaba D, Fish K, Howard S. Crisis Management in Anesthesiology. New York, NY: Churchill Livingstone; 1994.
22. Morey JC, Simon R, Jay GD, Wears RL, Salisbury M, Dukes KA, et al. Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the medteams project. Health Serv Res. 2002;37(6):1553–81.
23. Sawyer T, Eppich W, Brett-Fleegler M, Grant V, Cheng A. More Than One Way to Debrief - A Critical Review of Healthcare Simulation Debriefing Methods. Simul Healthc J Soc Simul Healthc. 2016;11(3):209–17.

24. Van De Ridder JMM, Stokking KM, McGaghie WC, Ten Cate OTJ. What is feedback in clinical education? Med Educ. 2008;42(2):189–97.
25. Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach. 2013 Jan;35(1):e867-98.
26. Phrampus PE, O'Donnell JM. Debriefing Using a Structured and Supported Approach. The Comprehensive Textbook of Healthcare Simulation. Springer; 2013. p. 73–84.
27. Cheng A, Rodgers DL, van der Jagt É, Eppich W, O'Donnell J. Evolution of the Pediatric Advanced Life Support course. Pediatr Crit Care Med. 2012;13(5):589–95.
28. Stanford School of Medicine. What Are Cognitive Aids? - Cognitive Aids in Medicine [Internet]. 2016 [cited 2016 Dec 14]. Available from: http://cogaids.stanford.edu/whatarecogaids.html
29. Harrison TK, Manser T, Howard SK, Gaba DM. Use of Cognitive Aids in a Simulated Anesthetic Crisis. Anesth Analg. 2006;103(3):551–6.
30. Marshall S. The Use of Cognitive Aids During Emergencies in Anesthesia: A Review of the Literature. Anesth Analg. 2013;117(5):1162–71.
31. Haugen AS, Softeland E, Almeland SK, Sevdalis N, Vonen B, Eide GE, et al. Effect of the World Health Organization Checklist on Patient Outcomes: A Stepped Wedge Cluster Randomized Controlled Trial. Ann Surg. 2015;261(5):821–8.
32. Cheng A, Hunt EA, Donoghue A, Nelson-McMillan K, Nishisaki A, LeFlore J, et al. Examining Pediatric Resuscitation Education Using Simulation and Scripted Debriefing: A Multicenter Randomized Trial. JAMA Pediatr. 2013;167(6):528–36.
33. Stegmann K, Pilz F, Siebeck M, Fischer F. Vicarious learning during simulations: is it more effective than hands-on training? Med Educ. 2012 Oct;46(10):1001–8.
34. Weinman J. Psychology as applied to medicine (2nd Ed.). Bristol, England: Wright; 1987.
35. Jenkins V, Fallowfield L. Can communication skills training alter physicians' beliefs and behaviour in clinics? J Clin Oncol. 2002;20(3):765–9.
36. Stroben F, Schröder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE. A simulated night shift in the emergency room increases students ' self-efficacy independent of role taking over during simulation. BMC Med Educ. 2016;16:177.
37. Rall M, Oberfrank S. „Human factors" und „crisis resource management" - Erhöhung der Patientensicherheit. Unfallchirurg. 2013;116:892–9.
38. Bittencourt Couto T, Kerrey BT, Taylor RG, Fitzgerald M, Geis GL. Teamwork Skills in Actual, In Situ, and In-Center Pediatric Emergencies. Simul Healthc. 2015;10:76–84.
39. Dorer B. ESS Round 7 Translation Guidelines [Internet]. European Social Survey. 2014 [cited 2017 Jan 5]. Available from: http://www.europeansocialsurvey.org/docs/methodology/ESS_R7_Translation_Guidelines_FINAL.pdf
40. Feldman M, Lazzara EH, Vanderbilt AA, DiazGranados D. Rater Trainings to Support High-Stakes Simulation-Based Assessments. J Contin Educ Heal Prof. 2012;33(4):279–88.
41. Park H, Kim C, Park J, Park E. Effects of team-based learning on perceived teamwork and academic performance in a healtch assessment subject. Collegian. 2014;
42. Maignan M, Koch F-X, Chaix J, Phellouzat P, Binauld G, Collomb Muret R, et al. Team Emergency Assessment Measure (TEAM) for the assessment of non-technical skills during resuscitation: Validation of the French version. Resuscitation. 2016;101:115–20.
43. Cant RP, Porter JE, Cooper SJ, Roberts K, Wilson I, Gartside C. Improving the non-technical skills of hospital medical emergency teams: The Team Emergency Assessment Measure (TEAM). EMA - Emerg Med Australas. 2016;28(6):641–6.

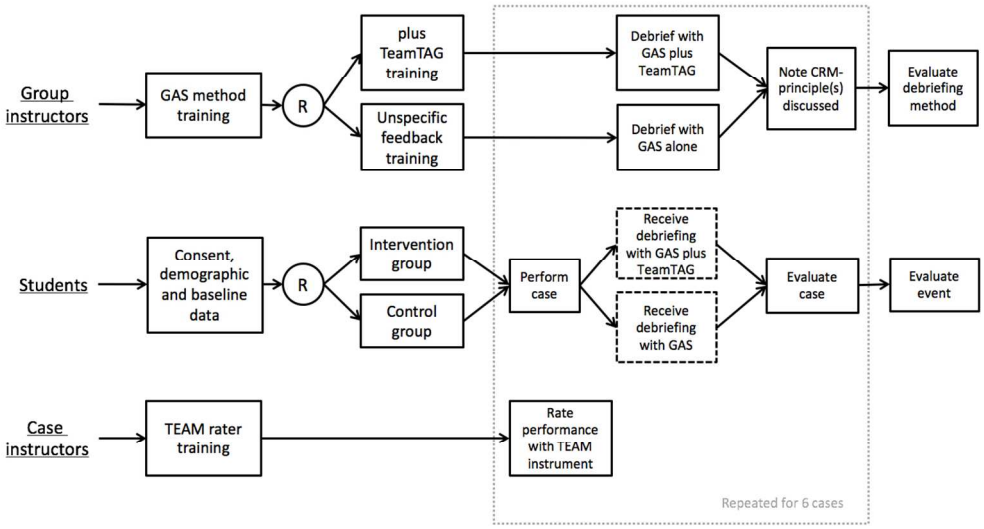

Fig. 1. Study flowchart; GAS = Gather – Analyse – Summarize. CRM = Crisis Resource Management. TEAM = Team Emergency Assessment Measure. TeamTAG =Teamwork Techniques Analysis Grid, R = randomization‼ ┼

485x265mm (72 x 72 DPI)

Table 1. *Teamwork-relevant cases during simulation.*

Case	*Diagnosis*	*Challenges for teamwork*
1	Exacerbated COPD	Conflict management, adequate emotions due to challenging patient
2	Ischemic media-stroke	Task management, communication with colleagues Manage aphasic patient
3	STEMI & non-sustained ventricular tachycardia	Patient deterioration (cardiac arrhythmia) during care
4	Ventricular fibrillation following STEMI	Team leadership, structured ACLS
5	Hemodynamic instable ruptured spleen	Set priorities in evaluation and management, structured ATLS
6	Head laceration with ethanol intoxication	Manage agitated patient

Team Emergency Assessment Measure (TEAM)

Einleitung

Dieser Fragebogen zu nicht-medizinischen Fähigkeiten wurde als Beobachtungsbogen für die valide, reliable und praktikable Bewertung von notfallmedizinischen Teams (z.B. Reanimations- und Traumateams) entwickelt. Der Fragebogen sollte von erfahrenen Klinikerinnen und Klinikern ausgefüllt werden, um akkurate Performanzmessungen und Feedback zur Führungsrolle, Teamarbeit, zum Situationsbewusstsein und Aufgabenmanagement zu ermöglichen. Wo zutreffend, sind Hinweise zur Bewertung angegeben. Die folgende Skala liegt der Bewertung zugrunde:

nie / fast nie	selten	ca. in der Hälfte der Fälle	oft	immer/fast immer
0	1	2	3	4

Angaben zum Team

Datum: Uhrzeit: Ort:

Teamleiter: Team:

Führungsrolle: Es wird angenommen, dass die Teamleitung entweder benannt ist, aus der Situation entsteht oder die/der Erfahrenste ist - falls keine Teamleitung besteht, vergeben Sie „0“ für Frage 1 und 2.	**0**	**1**	**2**	**3**	**4**
1. Die Teamleitung ließ durch Anweisungen das Team wissen, was von ihm erwartet wurde.	☐	☐	☐	☐	☐
2. Die Teamleitung behielt eine globale Perspektive. *Hinweise: Überwachung klinischer Maßnahmen und der Umgebung? Versucht, wenn möglich keine praktischen Aufgaben zu übernehmen (‘Hands off’)? Angemessene Delegation von Aufgaben.*	☐	☐	☐	☐	☐
Teamarbeit: Bewertungen sollten (mehr oder weniger) das Team als Ganzes umfassen, also Leitung und andere Mitglieder als Kollektiv.	**0**	**1**	**2**	**3**	**4**
3. Das Team kommunizierte effektiv. *Hinweise: Verbale, non-verbale und schriftliche Kommunikationsformen?*	☐	☐	☐	☐	☐
4. Das Team arbeitete zusammen um die Aufgaben zeitnah zu lösen.	☐	☐	☐	☐	☐
5. Das Team agierte gefasst und kontrolliert. *Hinweise: angebrachte Emotionen? Probleme beim Konfliktmanagement?*	☐	☐	☐	☐	☐
6. Die Einstellung des Teams war positiv. *Hinweise: angemessene Unterstützung, Zuversicht, Stimmung, Optimismus, Entschlossenheit?*	☐	☐	☐	☐	☐
7. Das Team passte sich an sich verändernde Situationen an. *Hinweise: Anpassung innerhalb der beruflichen Rolle?* *Situationsänderung: Zustandsverschlechterung des Patienten? Veränderungen im Team?*	☐	☐	☐	☐	☐
8. Das Team überwachte und re-evaluierte die Situation.	☐	☐	☐	☐	☐
9. Das Team antizipierte potentiell nötige Maßnahmen. *Hinweise: Vorbereitung der/s Defibrillators, Medikamente, Atemwegsmaterial?*	☐	☐	☐	☐	☐
Aufgabenmanagement:	**0**	**1**	**2**	**3**	**4**
10. Das Team priorisierte die Aufgaben.	☐	☐	☐	☐	☐
11. Das Team folgte anerkannten Standards und Leitlinien. *Hinweise: Sind Abweichungen möglicherweise angebracht?*	☐	☐	☐	☐	☐

Gesamtleistung:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
12. Vergeben Sie eine Gesamtbewertung für die nicht-medizinischen Fähigkeiten des Teams auf einer Skala von 1-10	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Kommentare:_____

_____

Translated by Julia Freytag & Fabian Stroben with friendly permission by Professor Simon Cooper, PhD
Contact: julia.freytag@charite.de, fabian.stroben@charite.de
Simulated Patients Program and Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité-Universitätsmedizin Berlin, Germany

TeamTAG

Key principle	Behavioral Marker	☐	Notes
Anticipate & Plan ahead	- agree on a plan with all team members - think ahead and plan for all contingencies - prepare a Plan B	☐ ☐ ☐	
Set priorities dynamically	- identify and set priorities at the beginning - pay attention towards changes which might become necessary / do not hold on to outdated concepts	☐ ☐	
Call for help early	- be aware of your own limits & the limits of your team - set predefined criteria for asking for help - know who and how you can call for help	☐ ☐ ☐	
Exercise leadership and followership	- as team leader: • allocate team roles & tasks • monitor progress • pay attention to team members • collect all information & make sure everyone is on the same page - as a team member: • be present and alert • share your thoughts/doubts • show appropriate self-care behavior	 ☐ ☐ ☐ ☐ ☐ ☐ ☐	
Communicate effectively	- clear, assertive - use Closed-Loop-Communication - team leader receives all information	☐ ☐ ☐	
Re-evaluate repeatedly	- review the plan regularly, if / how it works - respond to new information / arising problems etc.	☐ ☐	

Julia Freytag & Fabian Stroben – Contact: julia.freytag@charite.de, fabian.stroben@charite.de
Simulated Patients Program and Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité-Universitätsmedizin Berlin, Germany

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	____1_____
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	____n.a._____
	2b	All items from the World Health Organization Trial Registration Data Set	____n.a._____
Protocol version	3	Date and version identifier	_____1_____
Funding	4	Sources and types of financial, material, and other support	_____13_____
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	_____1_____
	5b	Name and contact information for the trial sponsor	_____n.a._____
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	_____13_____
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	___13_____

Introduction

Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	3
	6b	Explanation for choice of comparators	3-4
Objectives	7	Specific objectives or hypotheses	4-5
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	5

Methods: Participants, interventions, and outcomes

Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	5
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	5
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	4, 6-7
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	n.a.
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	9
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	n.a.
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	9-10
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Fig.1, 8-9

Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	10
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	6
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	6
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	6
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	6
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	6-8
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	n.a
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	9
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	n.a.

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	____10_____
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	_____10_____
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	_____10_____
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	_____10_____
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	____n.a. _____
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	____n.a. _____
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	____n.a. _____
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	____n.a. _____
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	___12_____
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	____n.a. _____

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	12
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	n.a.
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	8,12
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	13
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	12
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	n.a.
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	12
	31b	Authorship eligibility guidelines and any intended use of professional writers	12
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	12
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	upon request
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	n.a.

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

BMJ Open

Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective, and randomized study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-015977.R1
Article Type:	Protocol
Date Submitted by the Author:	02-May-2017
Complete List of Authors:	Freytag, Julia; Charite Universitatsmedizin Berlin, Simulated Patients Program, Office of the Vice Dean for Teaching and Learning Stroben, Fabian; Charite Universitatsmedizin Berlin, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning; Inselspital, University of Bern, Department of Emergency Medicine Hautz, Wolf; Inselspital, University of Bern, Department of Emergency Medicine Eisenmann, Dorothea; Charite Universitatsmedizin Berlin, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning; Charite Universitatsmedizin Berlin, Department of Anesthesiology and Operative Intensive Care Medicine CCM & CVK Kämmer, Juliane; Max-Planck-Institut fur Bildungsforschung, Center for Adaptive Rationality; Charite Universitatsmedizin Berlin, Progress Test Medizin
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Emergency medicine, Communication
Keywords:	MEDICAL EDUCATION & TRAINING, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < ANAESTHETICS

Improving patient safety through better teamwork: How effective are different methods of simulation debriefing? Protocol for a pragmatic, prospective, and randomized study

Version 1.1 / 13.1.2017 15:00h
Review: 28.4.2017

Julia Freytag, MSc(1)*; Fabian Stroben, MD (2,3)*; Dr. med. Wolf E Hautz, MME (3); Dr. med. Dorothea Eisenmann (2,4); Dr. rer. nat. Juliane E Kämmer (5,6)

*contributed equally

(1) Simulated Patients Program, Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Germany
(2) Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Germany
(3) Department of Emergency Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland
(4) Department of Anesthesiology and Operative Intensive Care Medicine CCM & CVK, Charité Medical School Berlin, Germany
(5) Progress Test Medizin, Charité Medical School Berlin, Germany
(6) Max Planck Institute for Human Development, Center for Adaptive Rationality, Berlin, Germany

Julia Freytag, MSc, Research Associate, Simulated Patients Program, Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany
Fabian Stroben, MD, Research Associate, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany & Department of Emergency Medicine, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse 4, 3010 Bern, Switzerland
Dr. med. Wolf E Hautz, MME, Consultant, Department of Emergency Medicine, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse 4, 3010 Bern, Switzerland
Dr. med. Dorothea Eisenmann, Head of Lernzentrum, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany & Department of Anesthesiology and Operative Intensive Care Medicine CCM & CVK, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany
Dr. rer nat. Juliane E Kämmer, Research Scientist, Progress Test Medizin, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany & Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany

Correspondence to:
Fabian Stroben, Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité Medical School Berlin, Charitéplatz 1, 10117 Berlin, Germany
Phone: +49 (0) 30 450 576 403
Fax: +49 (0) 30 450 576 922
fabian.stroben@charite.de

Abstract

Introduction: Medical errors have an incidence of 9% and may lead to worse patient outcome. Teamwork training has the capacity to significantly reduce medical errors and therefore improve patient outcome. One common framework for teamwork training is crisis resource management, adapted from aviation and usually trained in simulation settings. Debriefing after simulation is thought to be crucial to learning teamwork-related concepts and behaviours but it remains unclear how best to debrief these aspects. Furthermore, teamwork-training sessions and studies examining education effects on undergraduates are rare. The study aims to evaluate the effects of two teamwork-focused debriefings on team performance after an extensive medical student teamwork training.

Methods and analyses: A prospective experimental study has been designed to compare a well-established three-phase debriefing method (gather–analyse–summarize; the *GAS method*) to a newly developed and more structured debriefing approach that extends the GAS method with *TeamTAG* (teamwork techniques analysis grid). TeamTAG is a cognitive aid listing preselected teamwork principles and descriptions of behavioural anchors that serve as observable patterns of teamwork and is supposed to help structure teamwork-focussed debriefing. Both debriefing methods will be tested during an emergency room teamwork-training simulation comprising six emergency medicine cases faced by 35 final-year medical students in teams of five. Teams will be randomized into the two debriefing conditions. Team performance during simulation and the number of principles discussed during debriefing will be evaluated. Learning opportunities, helpfulness, and feasibility will be rated by participants and instructors. Analyses will include descriptive, inferential, and explorative statistics.

Ethics and dissemination: The study protocol was approved by the institutional office for data protection and the ethics committee of Charité Medical School and registered under EA2/172/16. All students will participate voluntarily and will sign an informed consent after receiving written and oral information about the study. Results will be published.

Strengths and limitations of the study.

- The study design builds on established principles of teaching and assessing teamwork.
- The study will be one of the first to explore the effects of teamwork-focused debriefing on team performance with undergraduate medical students.
- The study will be embedded in a well-established simulation setting with proven efficacy.
- The study will be a pragmatic, randomized comparison of two debriefing methods.
- Only a single centre will be studied.
- Feedback quality will not be externally evaluated.

1. Introduction

Medical errors and adverse events occur with an incidence of about 9% and can seriously harm patients.(1,2) Error rates in emergency settings are even reported to be twice as high.(3–5) Most medical errors originate from human factors and teamwork(6) or medication errors(7) and about half of all medical errors are considered preventable.(1,7)

Empirical evidence(6,8–11) suggests that improving teamwork may be key to reducing medical error. Yet, although teamwork and patient safety are prominent objectives in many national outcome frameworks,(12–14) these topics are insufficiently represented in undergraduate education and are rarely assessed, even though validated teamwork assessment tools exist.(15,16) Consequently, about 60% of junior doctors in Germany reported feeling inadequately prepared for clinical practice(17) and almost half of the residents in a Canadian survey reported feeling overwhelmed when leading a resuscitation team.(18)

In addition, common interventions targeting the quality of teamwork and human factors, such as simulation training and crisis resource management (CRM) training, have produced a variety of effects.(19,20) In both simulation and CRM training, debriefing is considered crucial to enhancing learning(21) but little is known about how best to debrief. In fact, the widely differing effects of simulation may very well result from differences in debriefing. A feasible and beneficial debriefing method, particularly for undergraduates, could lead to more effective simulation sessions and thus ease the transition into clinical practice for junior doctors. This could ultimately lead to a reduction of medical errors and thus improved patient outcome. In this study we will compare the effects of two different debriefing methods on team performance and the acquisition of teamwork skills during teamwork simulations for medical students.

Training and debriefing

The concept of CRM was originally derived from safety training in aviation and has been adapted to the health care sector, another high-stakes environment.(22) The idea of CRM is to guide individuals and teams in emergency situations (crises), encouraging them to use all available resources to manage the situation effectively and prevent critical incidents from occurring in the first place. CRM training has been shown to be a potent tool to improve teamwork and—as a consequence—patient safety.(23–25) In our study, elements of CRM set the framework for teamwork training and debriefing during an emergency room simulation.

Simulation debriefing is defined as a bidirectional and interactive discussion after a simulation in which participants reflect on their actions and analyse their performance.(21) Feedback is a central process element of debriefing that is often used as a conversational technique especially in participants with little experience in debriefing.(26) Feedback is defined as the delivery of information to improve

reasoning or behaviour compared to defined performance standards,(26,27) and it is critical in improving learning.(21) How best to integrate feedback into debriefing, what specific aspects to address, and how to structure debriefing to foster learning are, however, still unknown.(21,28) The goal of this study is thus to evaluate the potential benefit of preselecting certain aspects to be discussed during debriefing and of structuring debriefing with the help of a cognitive aid. To this end, we will compare a well-established debriefing method to a more structured and feedback-focused method to evaluate their effects on teamwork, learning opportunities, feasibility, and helpfulness for participants (and instructors). We will focus on two debriefing methods, the *gather–analyse–summarize* (GAS) method and the GAS method plus a cognitive aid:

(1) **The GAS method**: This debriefing method consists of three parts: **g**athering, **a**nalysing, and **s**ummarizing.(29,30) The GAS method is one of many similar three-step debriefing structures(26) and has been used, for example, in simulation courses run by the American Heart Association.(30) During the first phase (gather), participants are given the opportunity to report their thoughts on the simulated situation. They are encouraged to exchange their views on what actually happened to establish a shared mental model of the situation. This model can afterwards be used to discuss the simulation in a learner-centred way (analyse). During this process, questions tailored towards specific learning objectives are used to facilitate participants' reflection on and analysis of their actions and induce learning. Finally, the debriefing is summed up and critically reviewed by the team and its instructor (summarize).(26,29) Topics discussed during the debriefing using this method are mostly self-selected by the team and instructor, which makes this method highly flexible. A possible drawback with regard to teamwork (or any other specific learning objective) is that its potential to enhance the quality of teamwork is influenced by the instructor's level of experience.(26) A typical question to start the debriefing with the gather step might be "How do you feel now"? followed in the analysis step by "What worked well"? or "Do you see any opportunities for improvement"? The summarize step might be initiated by "What we learned from this session…".

(2) **The GAS method plus a cognitive aid**: This newly developed debriefing method uses the GAS structure detailed above and additionally provides the instructors with a cognitive aid to structure the debriefing in more detail. It further provides a selection of important aspects to address during debriefing. Cognitive aids are "structured pieces of information designed to enhance cognition and adherence to…best practices".(31) Cognitive aids have been shown to be beneficial in different areas of medicine.(32–34) Moreover, cognitive aids are useful for debriefing: Instructors' use of a cognitive aid may improve participants' acquisition of behavioural and cognitive outcomes after simulation—especially so with novice instructors.(35) In practice, such aids are often a pocket card, script, or poster.

We will use a specific cognitive aid called "TeamTAG" (teamwork techniques analysis grid) to foster observation and feedback relevant to teamwork. TeamTAG is a guideline for structuring the feedback process during debriefing and remembering what to address during the analysis step of the GAS method. The TeamTAG lists teamwork-relevant CRM principles together with descriptions of behavioural anchors that serve as directly observable patterns of teamwork and provides space for notes (see online supplementary information). The TeamTAG can be printed on a single sheet of paper (A4) and filled in during observation of the simulation. After the simulation, instructors have the flexibility to set priorities for debriefing based on their observations and structured notes. The debriefing itself will follow the same structure as under the GAS method. However, the TeamTAG might, for example, remind instructors that team leaders "allocate roles & tasks" or are responsible for "monitoring progress" (according to the CRM principle "exercise leadership and followership"). These aspects might be specifically addressed by group instructors to improve group reflection during the analysis step.

Hypotheses

First, we assume that the GAS method plus TeamTAG will be a more effective debriefing tool than the common GAS method alone and will lead to the discussion of more teamwork-relevant principles. Debriefing using the GAS method plus TeamTAG should thus result in more learning opportunities for teams and ultimately in improved team performance. This hypothesis is based on the fact that the TeamTAG is concise and guides observation and feedback with practical examples. Using these examples during observation may help focus the observers' attention(36) and result in the team discussing more teamwork-relevant CRM principles. In undergraduate education, instructors are often novices and vary considerably regarding how experienced they are in debriefing. Because novices were shown to benefit more from structured debriefing scripts than more experienced instructors,(35) we consider our environment (see Methods section) ideal for detecting differences between the two debriefing methods if they exist.

Hypothesis 1a: Participants who receive debriefing based on the GAS method plus TeamTAG will show a greater improvement in team performance than those who discuss the simulation according to the common GAS method alone.

Hypothesis 1b: Participants who receive debriefing based on the GAS method plus TeamTAG will report discussing a higher number of CRM principles than participants who are debriefed with the GAS method alone.

Second, we expect that teams receiving debriefing based on the GAS method plus TeamTAG will perceive teamwork skills as more important after the simulation event, which should increase their sensitivity to a culture of safety and the likelihood of changing their behaviour.(37,38) Moreover,

perceiving the content of the debriefing as more important should lead to higher overall satisfaction with and perception of helpfulness of the debriefing.

Hypothesis 2a: Participants who receive debriefing based on the GAS method plus TeamTAG will report a higher level of perceived importance of teamwork principles than those who are debriefed according to the common GAS method.

Hypothesis 2b: Participants who receive debriefing based on the GAS method plus TeamTAG will report higher satisfaction with and helpfulness of the debriefing they received than those who are debriefed according to the GAS method alone.

Third, we will focus on the satisfaction of the instructors as a measure of feasibility and efficiency. We expect higher satisfaction when they use the GAS method plus TeamTAG as it might facilitate more structured feedback and it provides a better opportunity for instructors to address the learning objectives of their participants.

Hypothesis 3: Instructors who use the GAS method plus TeamTAG will report higher levels of feasibility and efficiency of their debriefing than instructors who use the GAS method alone.

2. Methods and analysis

This investigation is designed as a prospective experimental superiority study with intervention and control groups receiving debriefing during a simulation training based on either the GAS method plus TeamTAG or the GAS method alone, respectively. The study will be executed during an emergency department (ED) simulation at Charité Medical School, Berlin, Germany, on January 14, 2017. The ED simulation has been implemented at the local skills lab since 2013 on a peer-led basis. The main goal of this extensive, 8-h night-shift simulation training is to give students the opportunity to experience being the person in charge of a patient's health care. This event takes place once a year, with about 35 students in their final year of medical studies participating voluntarily. Participants are recruited via newsletter and advertising posters. The students act in randomly assigned teams of five and self-select into different roles (team leader, team member, observer), which they switch during the night. Simulated patients and high-fidelity simulators are used to create realistic case simulations; simulated radiologic and laboratory services are provided. One of the main goals of the event is to improve students' confidence in working with medical emergencies in an ED over the course of the night.(39) The simulation was awarded a project prize by the German Association for Medical Education in 2016.

Each student team has to work on six simulated cases. Each case is staffed with a case instructor who is responsible for the simulation and provides technical help. Each student team is accompanied by a group instructor who guides the participants during the night. After every case, multi-source feedback is provided by simulated patients, observing participants, and case instructors. As part of our study, in 2017 participants will additionally receive a teamwork-based debriefing by the group instructors after every case in one of two conditions (GAS method vs. GAS method plus TeamTAG). Additionally, the quality of teamwork will be rated by trained raters throughout the night.

As group instructors we will choose experienced peer teachers who are advanced in their health-care studies (medicine, nursing) and have completed emergency room courses/electives during their studies. Peer teachers at Charité Medical School frequently give courses in clinical skills training and simulator-based emergency medicine trainings for other medical students. All group instructors undergo extensive feedback training during their studies and are furthermore trained in working with and debriefing groups.

Development of the TeamTAG as cognitive aid

As a basis for this study, the TeamTAG guideline was developed with the goal of having a feasible and time-efficient feedback instrument that supports teaching basic teamwork skills to participants. Two investigators (JF and FS) developed the TeamTAG guidelines that present six common CRM principles,(22,40) each accompanied by the description of behavioural anchors. The six principles are (1) anticipate and plan ahead, (2) set priorities dynamically, (3) call for help early, (4) exercise

leadership and followership, (5) communicate effectively, and (6) re-evaluate repeatedly. The TeamTAG can be found in the online supplementary material. The CRM principles and their behavioural anchors were chosen to fit the following criteria: (a) simulation setting, (b) presumed skills of participants, (c) experience of instructors, and (d) observability. The tool was reviewed and adjusted by an experienced group of anaesthesiologists, emergency medicine physicians, simulation instructors, and peer tutors, all experienced in medical education and simulation-based learning. In a pre-study, feasibility for instructors was examined (see preliminary results below) but not compared to an approach without the TeamTAG.

Team performance measurement

To measure team performance, we will use the Team Emergency Assessment Measure (TEAM).(15) TEAM is an assessment tool that has been applied to both clinical and simulation environments.(15,16,41) It consists of 11 items belonging to the three subscales leadership, teamwork, and task management. Example items are "the team leader maintained a global perspective" and "the team prioritized tasks", measured on a 5-point Likert scale of 0 (*never*) to 4 (*always*). Additionally, it includes an overall rating of team performance [range: 1 (*poorest performance*) to 10 (*best performance*)].

As there was no German version of the TEAM, the English version was translated into German, using elements of the TRAPD (translation, review, adjudication, pre-test, documentation) methodology(42). Two investigators (JF and FS) independently translated the TEAM into German in parallel, reviewed the results, and consented to one version, which was translated back by a native English speaker. This new version was compared to the original TEAM and agreed to by both investigators and the native speaker. All steps of the translation were documented.

After the TEAM was translated, we developed a rater training. The training involves three aspects that are important preparation for accurately assessing a certain behaviour or skill(43): (1) a *rater error training* in which information is provided on typical rating errors to raise awareness and prevent them, (2) a *performance dimension training* to teach raters about the targeted dimensions, including definitions and videotaped examples, and (3) a *frame-of-reference training*, in which videotaped examples showing teamwork of different levels of quality are assessed and discussed. All raters who will be responsible for TEAM ratings in this study (case instructors and additional raters) will receive this rater training and additional written material on teamwork and how to use the TEAM.

Group instructors debriefing training

Before data collection, all group instructors will receive a teamwork-related training and additional written material with information about how to provide feedback and conduct debriefings and about human factors in general and CRM in particular, which is intended to serve as a framework for

discussing all teamwork aspects during debriefing. The training will include videos showing good and bad examples of teamwork and will be followed by discussions about opportunities for debriefing in these specific situations (adapted from *frame of reference training*(43)). After this training, which will be the same for all group instructors, the instructors will be randomly assigned, stratified by level of academic education and additional professional training (e.g., nurse or paramedic), to the two conditions. The two groups will receive separate instruction from the investigators: The intervention group instructors will be told to discuss their groups' performance with the help of the TeamTAG and to focus on each CRM principle of the TeamTAG at least once during the first five cases (i.e., one or two principles per case) so that by Case 6 all CRM principles will have been debriefed and team performance during Case 6 can be compared between conditions. Furthermore, they will be instructed to re-evaluate their previous focus of debriefing after each case if behaviour does not change sufficiently from their perspective. The order of chosen topics can be varied by the instructors and should be adjusted to observed difficulties in teamwork during the simulation. The control group instructors will be advised to give feedback regarding whatever teamwork-related aspect they deem important during the first five cases and also to re-evaluate the teamwork if needed. Instructors will stay with their groups during the whole simulation event to guarantee coordinated, consistent, and longitudinal feedback.

Data collection

Upon arrival, every student participant will create an individual anonymized study code, which will be entered on every form and questionnaire and will allow us to link all measurements during the course of the night. Students will also track their role (leader, member, observer) after every case to allow sub-group analyses in relation to these roles. Figure 1 depicts the data collection procedure during the night-shift simulation.

Fig. 1.

Fig. 1. *Study flowchart; GAS = gather–analyse–summarize. CRM = crisis resource management. TEAM = Team Emergency Assessment Measure. TeamTAG = teamwork techniques analysis grid, R = randomization.*

Before starting the simulation, all 35 participants will be asked to fill in a first questionnaire that assesses possible confounders such as demographic data, professional training as a nurse or paramedic, or any training in teamwork/human factors. Next, students will be randomly assigned to seven groups via a computer-generated algorithm by the principal investigator. Four groups will serve as intervention groups and the remaining three as controls; participants will not know to which condition they are assigned. After randomization, all groups will gather separately and will be asked to discuss already known principles of teamwork and 15 multiple choice questions concerning

emergency medicine. A recent study showed that the results of such discussions are linked to team performance.(44)

During the simulation, all groups will face six simulations where teamwork will be measured and teamwork-related feedback provided. All cases depict common emergency situations where the participation of an emergency team in the emergency room is needed. Table 1 gives a brief overview of the diagnoses of the six cases and challenges for teamwork.

Table 1. *Teamwork-relevant cases presented in the emergency department simulation.*

Case	*Diagnosis*	*Challenges for teamwork*
1	Exacerbated COPD	Conflict management, control of emotions due to challenging patient
2	Ischemic stroke of middle cerebral artery	Task management, communication with colleagues Manage aphasic patient
3	STEMI & non-sustained ventricular tachycardia	Patient deterioration (cardiac arrhythmia) during care
4	Ventricular fibrillation following STEMI	Team leadership, structured ACLS
5	Haemodynamically unstable ruptured spleen	Set priorities in evaluation and management, structured ATLS
6	Head laceration with ethanol intoxication	Manage agitated patient

Note. COPD = chronic obstructive pulmonary disease, STEMI = ST-elevation myocardial infarction, ACLS = advanced cardiac life support, ATLS = advanced trauma life support.

During every case, team performance will be measured using the TEAM,(16) which will be filled in by the case instructors and an additional rater. The two TEAM raters will be blind to the debriefing condition the group is assigned to.

After every case (duration about 30 min), debriefing will start (duration about 20 min) with checklist-based feedback from the simulated patients (focus: communication skills, empathy) and the case instructors and peer observers (focus: factual knowledge, diagnostic skills). As the last part of the debriefing process, the teamwork-related debriefing will be conducted by the group instructor using the GAS method with or without the support of the TeamTAG depending on the experimental condition. The strict timing, which will be centrally coordinated, will be necessary for a smooth transition of groups between cases and to ensure that the total length of the simulation does not exceed 8 h.

After the debriefing process, all group members will be asked to evaluate the case and rate how helpful the debriefing was. Group instructors in both conditions will track the main topics of their teamwork debriefing in a debriefing protocol as free text. After the simulation the content of these debriefing protocols will be clustered independently (JF and FS) and matched with CRM principles.

Right after the last case of the night, all participants will fill in a final evaluation, which will ask them to list all the CRM principles on which they received feedback during the night. Participants will also evaluate the importance of each principle for their future work as physicians and provide a general evaluation of the night. Every group tutor will rate the feasibility, efficiency, and difficulty of providing feedback.

Collected data

(1) Baseline characteristics: The data collected on the first questionnaire and the results of group and teamwork discussions will be used to compare the baseline between the two conditions. Discussion results will be analysed qualitatively to identify differences in knowledge and in the personal definition of good teamwork at the beginning of the night. Furthermore, the TEAM scores during the first simulation case will serve as the baseline team performance.

(2) Hypothesis 1 measurement (team performance, number of CRM principles discussed): Team performance will be evaluated using the 11 items of the translated TEAM. Similar to previous studies,(15,16,41,45,46) we will analyse ratings on the item level (range: 0 to 4), the sum score (range: 0 to 44), and the overall rating per case (range: 1 to 10). The number of CRM principles discussed will be derived from two sources, namely, the debriefing protocols of the group instructors and participants.

(3) Hypothesis 2 measurement (importance, satisfaction, helpfulness): Estimated relevance of the CRM principles learned and overall satisfaction with the simulation will be evaluated on 7-point Likert scales at the end of the night. Helpfulness of the debriefing from the different providers (simulated patient, peer, case tutor, and group tutor) will be rated by participants after every case on a 7-point Likert scale.

(4) Hypothesis 3 measurement (instructor ratings): Debriefing evaluation of the group instructors (feasibility, efficiency, and difficulty of providing feedback) will be measured with 7-point Likert scales and as free-text answers at the end of the night.

(5) Other measures: The general evaluation form will ask participants to rate pleasure, quality of instruction during the night, difficulty of cases, possibility of applying knowledge on 7-point Likert scales.

All Likert scales will be coded from +3 (*strongly agree*) to -3 (*strongly disagree*). All data collection forms will be available upon request.

Analyses

Data will be analysed in SPSS and R using descriptive, inferential, and explorative statistics. We conducted a calculation of power for our primary research question (team performance). Recent studies, reporting mainly data for well-trained and experienced teams, showed TEAM sum scores up

to 40.(45,46) Only one study provided data for less experienced teams with a TEAM sum score of 21.(45) On the basis of these results and data from a pre-study (see the TeamTAG section in the Preliminary results), we expect a TEAM sum score of about 20 for an untrained team and a score of around 40 for teams that receive a training related to teamwork skills and/or have a lot of experience in this area. These scores indicate a potential increase due to training of up to 20 points on the TEAM sum score. As a relevant training effect for a single training event such as ours, we estimate a gain in the TEAM sum score of 11 points (i.e., one point per item). Using the standard deviation from the last published study on the TEAM(46) ($SD = 4.4$) and $\alpha < 0.05$, we have determined that about six teams are needed to detect a significant difference between the conditions with a power of 80%. Missing data will be handled using pairwise deletion.

(1) Baseline characteristics: Discussion results of the intervention and control groups will be compared using qualitative methods and confounder analysis (demographics, prior training) with parametric and non-parametric tests for testing equivalence. The TEAM scores (single items, sum score, overall score) from the first simulation case will be compared between conditions using multi-level analyses to take the hierarchical structure of data into account.

(2) Analyses for Hypothesis 1: The TEAM scores (single items, sum score, overall score) of the intervention and control groups during the sixth simulation case will be compared using multi-level analyses. The development of team performance over the six cases will be analysed using descriptive statistics and plotting "training curves" for each team. The total number of CRM principles discussed in the control and intervention groups will be compared using a multi-level model.

(3) Analyses for Hypothesis 2: The participants' ratings of the feedback's helpfulness, the importance of CRM principles and satisfaction with the debriefing will be compared between the control and intervention groups using multi-level models.

(4) Analyses for Hypothesis 3: Group instructors' evaluations of the instrument will be examined descriptively.

(5) Other measures: The general evaluation will be examined in a descriptive way.

Methodological limitations

Group instructors will not be observed while debriefing due to our limited labour force. Therefore, we cannot be sure the quality of the debriefing will be comparable among the seven participating groups. Further studies could use debriefing assessment tools such as the Observational Structured Assessment of Debriefing Tool,(47) which might help distinguish between effects of overall debriefing quality and our approach. In our study, we will try to address this limitation with extensive group instructor training to ensure an equal qualification level regarding debriefing and with a randomization of

instructors to conditions. Furthermore, participants will be asked to state the debriefing topic and to rate the quality of debriefing after every simulation case, which will be reported in later publications.

The time for debriefing after every case will be relatively short due to the design of our 8-h simulation, where all groups will rotate through six cases to give participants a broad overview of emergency medicine and application areas of CRM. To use this limited time most productively, we have added additional specifications for debriefing (e.g., focus on one or two principles per debriefing session, as described in the Methods section) because some instructors stated in a pre-study that the time allowed for debriefing was not sufficient. Future studies could investigate whether results of this study hold if all CRM principles are being discussed and thus repeated after every case/more often during the night and if time for debriefing is longer. Until now, there has been no strong evidence for the superiority of a longer debriefing.(21)

The study will focus only on short-term effects of two different debriefing approaches. Further research should investigate long-term effects on performance or changes in behaviour during clinical practice. A last limitation of this study is that it is a single-centre study and so results might be limited to local circumstances.

Data sharing statement

Data analysis will be conducted by the investigator's team (data management team). As the study is not a clinical trial, a data-monitoring team is not needed. The anonymised full data set will be published together with the journal publication or using the Dryad Data Repository (Durham, NC, USA) as required by the journal's guidelines. Data will furthermore be stored in the local data repository at Charité Medical School Berlin according to the local guidelines for good scientific practice.

3. Preliminary results

Validation of the German TEAM

The German TEAM can be found in the online supplementary information. As a preliminary validation, interrater correlation was checked between three investigators (JF, FS, and DE) and an external expert on two videotaped resuscitations. Both resuscitations were simulation based and had similar factual content; however, the first simulation showed good teamwork and the second intermediate teamwork performance. The videotaped simulations were used for group instructors' debriefing training and for validity testing of the German TEAM.

Intraclass correlation coefficients were .99 for the first resuscitation (Mean TEAM score = 42.3, *SD* = 1.3) and .85 for the second (Mean TEAM score = 22.5, *SD* = 3.1), which indicates excellent interrater agreement. For this reason, we consider the German TEAM a valid instrument for assessing team performance in our study.

TeamTAG

A first version of TeamTAG was used in a pre-study, conducted during the previous simulated night shift in 2016. In this pre-study, all instructors (*N* = 7) used TeamTAG as part of their debriefing (similar to the GAS method plus TeamTAG). They were asked to rate the feasibility and helpfulness of the TeamTAG (7-point Likert scale; -3 to +3), as well as whether time for debriefing was sufficient [7-point Likert scale; -3 (*strongly insufficient*) to +3 (*strongly sufficient*)]. Furthermore they could comment on specific aspect of the guideline they liked or disliked (free-text answers). All participants were asked how useful the instructors' feedback was (7-point Likert scale; -3 to +3).

Instructors rated the guideline as a feasible tool (*M* = 1.9, *SD* = 0.9) and stated that it helped them in both observing and giving feedback to the participants of the simulation ($M_{observe}$ = 2.3, *SD* = 0.8; $M_{feedback}$ = 2.3, *SD* = 0.5). They had a heterogeneous view of the adequacy of time available for debriefing (*M* = -0.3, *SD* = 1.1) The participants declared having found the feedback to be useful (*M* = 1.7, *SD* = 1.0).

4. Ethics and dissemination

The study protocol was designed according to the Declaration of Helsinki, the local guidelines for good scientific practice at Charité Medical School Berlin, and the ICMJE (International Committee of Medical Journal Editors) recommendations. The study protocol was approved by the institutional office for data protection (AZ 737/16) and the ethics committee at Charité Medical School Berlin (EA2/172/16).

All participants and instructors will provide informed consent. Because the simulation is already a well-known event at Charité Medical School Berlin and receives official teaching funds, participants who refuse to take part in our study must have a chance to participate nevertheless. In this case, students will not provide the informed consent prior to randomization; instead, an independent "no-study" group will then be created, which will be identical to the control group but without any teamwork debriefing. We do not expect any harm for students who undergo the intervention.

Publication

Results of the study will be presented during national and international scientific meetings. The authors aim to publish all results in a peer-reviewed journal. Part of the protocol has been previously presented at the Research in Medical Education (RIME) conference in Duesseldorf, Germany, in March 2017 and was awarded the RIME Award: Best Research Protocol 2017.(48)

5. Author contributions

JF and FS translated the TEAM, designed the study, and will be responsible for conduction. DE, WEH, and JEK contributed to the study design. JEK supervised the study design and will supervise conduction. JF and FS are responsible for data analyses. JF, FS, and JEK wrote the manuscript. JF and FS conducted the pre-study. DE is responsible for funding and local administration at Charité Medical School Berlin and heads the steering committee. All authors carefully read the manuscript, made critical and substantial revisions, and gave their approval for publication.

6. Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The investigators JF and FS are funded by the German Federal Ministry of Research and Education (BMBF). The sponsor did not interfere with the conception and conduction of the study, data analysis, or production of the manuscript.

7. Competing interests

WEH received financial compensation for educational consultancy from the AO Foundation, Zurich, Switzerland. All other authors report no competing interests.

8. Acknowledgements

The authors would like to acknowledge Hanno Heuzeroth and David Steinbart for support in conducting the pre-study and Simon Cooper for support in using and translating the TEAM. Furthermore, the authors thank all instructors and physicians, especially Tobias Deselaers, for organising the simulation and their willingness to participate as raters and instructors, and such. In addition, we would like to acknowledge Jane Runnacles (London), Suzanne Bentley (New York), and Christopher Timmis (Wolverhampton) for their constructive critique on an earlier version of the manuscript. The authors thank Anita Todd for editing the manuscript.

9. References

1. de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. Qual Saf Health Care. 2008;17:216–23.
2. Kohn LT, Corrigan JM, Donaldson MS, editors. To Err Is Human: Building a Safer Health System. Washington: National Academies Press; 2000.
3. Berner ES, Graber ML. Overconfidence as a Cause of Diagnostic Error in Medicine. Am J Med. 2008;121(5 Suppl):S2–23.
4. Schwartz A, Elstein AS. Clinical reasoning in medicine. In: Higgs J, Jones M, Loftus S, Christensen N, editors. Clinical reasoning in the health professions. 3rd ed. Oxford (UK); 2008. p. 223–34.
5. Hautz SC, Schuler L, Kämmer JE, Schauber SK, Ricklin ME, Sauter TC, et al. Factors predicting a change in diagnosis in patients hospitalised through the emergency room: a prospective observational study. BMJ Open. 2016;6(5):e011585.
6. Risser DT, Rice MM, Salisbury ML, Simon R, Jay GD, Berns SD. The Potential for Improved Teamwork to Reduce Medical Errors in the Emergency Department. Ann Emerg Med. 1999;34(3):373–83.
7. Leape LL. Error in medicine. JAMA. 1994;272(23):1851–7.
8. Hautz WE, Kämmer JE, Exadaktylos A, Hautz SC, Kaba E. How thinking about groups is different from groupthink. Med Educ. 2016;ePub:13137.
9. Hautz WE, Kämmer JE, Schauber S, Spieß CD, Gassmeier W. Diagnostic Performance by Medical Students Working Individually or in Teams. JAMA. 2015;313(3):303–4.
10. Knaus WA, Draper EA, Wagner DP, Zimmermann JE. An Evaluation of Outcome from Intensive Care in Major Medical Centers. Ann Intern Med. 1986;104(3):410–8.
11. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. Acta Anaesthesiol Scand. 2009;53:143–51.
12. Frank J, Snell L, Sherbino J, editors. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.
13. Kaba A, Wishart I, Fraser K, Coderre S, Mclaughlin K. Are we at risk of groupthink in our approach to teamwork interventions in health care? Med Educ. 2016;50(4):400–8.
14. Kiesewetter J, Gutmann J, Drossard S, Gurrea Salas D, Prodinger W, Mc Dermott F, et al. The Learning Objective Catalogue for Patient Safety in Undergraduate Medical Education–A Position Statement of the Committee for Patient Safety and Error Management of the German Association for Medical Education. GMS J Med Educ. 2016;33(1):1–14.
15. Cooper S, Cant R, Connell C, Sims L, Porter JE, Symmons M, et al. Measuring teamwork performance: Validity testing of the Team Emergency Assessment Measure (TEAM) with clinical resuscitation teams. Resuscitation. 2016;101:97–101.
16. Cooper S, Cant R, Porter J, Sellick K, Somers G, Kinsman L, et al. Rating medical emergency teamwork performance: Development of the Team Emergency Assessment Measure (TEAM). Resuscitation. 2010;81(4):446–52.
17. Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011 Jan;11(1):99.
18. Hayes CW, Rhee A, Detsky ME, Leblanc VR, Wax RS. Residents feel unprepared and unsupervised as leaders of cardiac arrest teams in teaching hospitals: A survey of internal medicine residents. Crit Care Med. 2007;35(7):1668–72.
19. Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA. Linking Simulation-Based Educational Assessments and Patient-Related Outcomes: A Systematic Review and Meta-Analysis. Acad Med. 2015;90(2):246–56.
20. Morgan L, Hadi M, Pickering S, Robertson E, Griffin D, Collins G, et al. The effect of teamwork training on team performance and clinical outcome in elective orthopaedic surgery: a controlled interrupted time series study. BMJ Open. 2015;5:e006216.
21. Cheng A, Eppich W, Grant V, Sherbino J, Zendejas B, Cook DA. Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. Med Educ. 2014;48:657–66.
22. Gaba D, Fish K, Howard S. Crisis Management in Anesthesiology. New York: Churchill Livingstone; 1994.

23. Morey JC, Simon R, Jay GD, Wears RL, Salisbury M, Dukes KA, et al. Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the medteams project. Health Serv Res. 2002;37(6):1553–81.
24. DeVita MA, Schaefer J, Lutz J, Wang H, Dongilli T. Improving medical emergency team (MET) performance using a novel curriculum and a computerized human patient simulator. Qual Saf Health Care. 2005;14:326–31.
25. Theilen U, Fraser L, Jones P, Leonard P, Simpson D. Regular in-situ simulation training of paediatric Medical Emergency Team leads to sustained improvements in hospital response to deteriorating patients, improved outcomes in intensive care and financial savings. Resuscitation. 2017;115:61–7.
26. Sawyer T, Eppich W, Brett-Fleegler M, Grant V, Cheng A. More Than One Way to Debrief–A Critical Review of Healthcare Simulation Debriefing Methods. Simul Healthc. 2016;11(3):209–17.
27. Van De Ridder JMM, Stokking KM, McGaghie WC, Ten Cate OTJ. What is feedback in clinical education? Med Educ. 2008;42(2):189–97.
28. Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach. 2013;35(1):e867-98.
29. Phrampus PE, O'Donnell JM. Debriefing Using a Structured and Supported Approach. In: Levine AL, DeMaria SJ, Schwartz AD, Sim AJ, editors. The Comprehensive Textbook of Healthcare Simulation. New York: Springer; 2013. p. 73–84.
30. Cheng A, Rodgers DL, van der Jagt É, Eppich W, O'Donnell J. Evolution of the Pediatric Advanced Life Support course. Pediatr Crit Care Med. 2012;13(5):589–95.
31. Stanford School of Medicine. What Are Cognitive Aids? Cognitive Aids in Medicine [Internet]. 2016 [cited 2016 Dec 14]. Available from: http://cogaids.stanford.edu/whatarecogaids.html
32. Harrison TK, Manser T, Howard SK, Gaba DM. Use of Cognitive Aids in a Simulated Anesthetic Crisis. Anesth Analg. 2006;103(3):551–6.
33. Marshall S. The Use of Cognitive Aids During Emergencies in Anesthesia: A Review of the Literature. Anesth Analg. 2013;117(5):1162–71.
34. Haugen AS, Softeland E, Almeland SK, Sevdalis N, Vonen B, Eide GE, et al. Effect of the World Health Organization Checklist on Patient Outcomes: A Stepped Wedge Cluster Randomized Controlled Trial. Ann Surg. 2015;261(5):821–8.
35. Cheng A, Hunt EA, Donoghue A, Nelson-McMillan K, Nishisaki A, LeFlore J, et al. Examining Pediatric Resuscitation Education Using Simulation and Scripted Debriefing: A Multicenter Randomized Trial. JAMA Pediatr. 2013;167(6):528–36.
36. Stegmann K, Pilz F, Siebeck M, Fischer F. Vicarious learning during simulations: is it more effective than hands-on training? Med Educ. 2012 Oct;46(10):1001–8.
37. Weinman J. Psychology as applied to medicine. 2nd ed. Bristol, England: Wright; 1987.
38. Jenkins V, Fallowfield L. Can communication skills training alter physicians' beliefs and behaviour in clinics? J Clin Oncol. 2002;20(3):765–9.
39. Stroben F, Schröder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE. A simulated night shift in the emergency room increases students ' self-efficacy independent of role taking over during simulation. BMC Med Educ. 2016;16:177.
40. Rall M, Oberfrank S. "Human factors" und "crisis resource management"–Erhöhung der Patientensicherheit. Unfallchirurg. 2013;116:892–9.
41. Bittencourt Couto T, Kerrey BT, Taylor RG, Fitzgerald M, Geis GL. Teamwork Skills in Actual, In Situ, and In-Center Pediatric Emergencies. Simul Healthc. 2015;10:76–84.
42. Dorer B. ESS Round 7 Translation Guidelines [Internet]. European Social Survey. 2014 [cited 2017 Jan 5]. Available from: http://www.europeansocialsurvey.org/docs/methodology/ESS_R7_Translation_Guidelines_FINAL.pdf
43. Feldman M, Lazzara EH, Vanderbilt AA, DiazGranados D. Rater Trainings to Support High-Stakes Simulation-Based Assessments. J Contin Educ Health Prof. 2012;33(4):279–88.
44. Park H, Kim C, Park J, Park E. Effects of team-based learning on perceived teamwork and academic performance in a health assessment subject. Collegian. 2015;22(3):299–305.

45. Maignan M, Koch F-X, Chaix J, Phellouzat P, Binauld G, Collomb Muret R, et al. Team Emergency Assessment Measure (TEAM) for the assessment of non-technical skills during resuscitation: Validation of the French version. Resuscitation. 2016;101:115–20.
46. Cant RP, Porter JE, Cooper SJ, Roberts K, Wilson I, Gartside C. Improving the non-technical skills of hospital medical emergency teams: The Team Emergency Assessment Measure (TEAM). Emerg Med Australas. 2016;28(6):641–6.
47. Arora S, Ahmed M, Paige J, Nestel D, Runnacles J, Hull L, et al. Objective structured assessment of debriefing: bringing science to the art of debriefing in surgery. Ann Surg. 2012;256(6):982–8.
48. Freytag J, Stroben F, Kämmer JE, Hautz WE, Eisenmann DE. Teamwork-related debriefings in undergraduate emergency-medicine simulations : Evaluation of the „TeamTAG" using a prospective- randomized pilot-study design. Paper presented at 5th International Conference for Research in Medical Education (RIME 2017). Düsseldorf: German Medical Science GMS Publishing House; 2017.

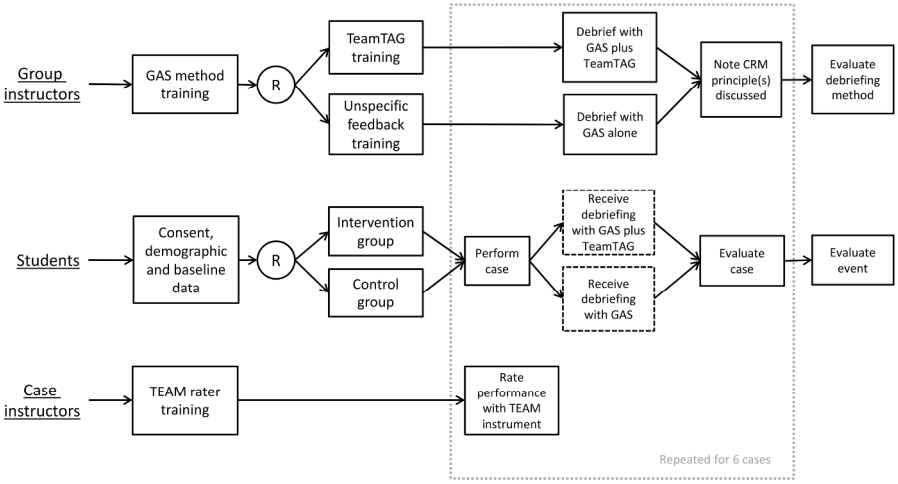

Fig. 1. Study flowchart; GAS = gather–analyse–summarize. CRM = crisis resource management. TEAM = Team Emergency Assessment Measure. TeamTAG = teamwork techniques analysis grid, R = randomization.‼ ┼

122x86mm (600 x 600 DPI)

TeamTAG

Key principle	Behavioral Marker	☐	Notes
Anticipate & Plan ahead	- agree on a plan with all team members - think ahead and plan for all contingencies - prepare a Plan B	☐ ☐ ☐	
Set priorities dynamically	- identify and set priorities at the beginning - pay attention towards changes which might become necessary / do not hold on to outdated concepts	☐ ☐	
Call for help early	- be aware of your own limits & the limits of your team - set predefined criteria for asking for help - know who and how you can call for help	☐ ☐ ☐	
Exercise leadership and followership	- as team leader: • allocate team roles & tasks • monitor progress • pay attention to team members • collect all information & make sure everyone is on the same page - as a team member: • be present and alert • share your thoughts/doubts • show appropriate self-care behavior	 ☐ ☐ ☐ ☐ ☐ ☐ ☐	
Communicate effectively	- clear, assertive - use Closed-Loop-Communication - team leader receives all information	☐ ☐ ☐	
Re-evaluate repeatedly	- review the plan regularly, if / how it works - respond to new information / arising problems etc.	☐ ☐	

Julia Freytag & Fabian Stroben – Contact: julia.freytag@charite.de, fabian.stroben@charite.de
Simulated Patients Program and Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité-Universitätsmedizin Berlin, Germany

Team Emergency Assessment Measure (TEAM)

Einleitung

Dieser Fragebogen zu nicht-medizinischen Fähigkeiten wurde als Beobachtungsbogen für die valide, reliable und praktikable Bewertung von notfallmedizinischen Teams (z.B. Reanimations- und Traumateams) entwickelt. Der Fragebogen sollte von erfahrenen Klinikerinnen und Klinikern ausgefüllt werden, um akkurate Performanzmessungen und Feedback zur Führungsrolle, Teamarbeit, zum Situationsbewusstsein und Aufgabenmanagement zu ermöglichen. Wo zutreffend, sind Hinweise zur Bewertung angegeben. Die folgende Skala liegt der Bewertung zugrunde:

nie / fast nie	selten	ca. in der Hälfte der Fälle	oft	immer/fast immer
0	1	2	3	4

Angaben zum Team

Datum: Uhrzeit: Ort:

Teamleiter: Team:

Führungsrolle: Es wird angenommen, dass die Teamleitung entweder benannt ist, aus der Situation entsteht oder die/der Erfahrenste ist - falls keine Teamleitung besteht, vergeben Sie „0“ für Frage 1 und 2.	**0**	**1**	**2**	**3**	**4**
1. Die Teamleitung ließ durch Anweisungen das Team wissen, was von ihm erwartet wurde.	☐	☐	☐	☐	☐
2. Die Teamleitung behielt eine globale Perspektive. *Hinweise: Überwachung klinischer Maßnahmen und der Umgebung? Versucht, wenn möglich keine praktischen Aufgaben zu übernehmen (‘Hands off’)? Angemessene Delegation von Aufgaben.*	☐	☐	☐	☐	☐
Teamarbeit: Bewertungen sollten (mehr oder weniger) das Team als Ganzes umfassen, also Leitung und andere Mitglieder als Kollektiv.	**0**	**1**	**2**	**3**	**4**
3. Das Team kommunizierte effektiv. *Hinweise: Verbale, non-verbale und schriftliche Kommunikationsformen?*	☐	☐	☐	☐	☐
4. Das Team arbeitete zusammen um die Aufgaben zeitnah zu lösen.	☐	☐	☐	☐	☐
5. Das Team agierte gefasst und kontrolliert. *Hinweise: angebrachte Emotionen? Probleme beim Konfliktmanagement?*	☐	☐	☐	☐	☐
6. Die Einstellung des Teams war positiv. *Hinweise: angemessene Unterstützung, Zuversicht, Stimmung, Optimismus, Entschlossenheit?*	☐	☐	☐	☐	☐
7. Das Team passte sich an sich verändernde Situationen an. *Hinweise: Anpassung innerhalb der beruflichen Rolle?* *Situationsänderung: Zustandsverschlechterung des Patienten? Veränderungen im Team?*	☐	☐	☐	☐	☐
8. Das Team überwachte und re-evaluierte die Situation.	☐	☐	☐	☐	☐
9. Das Team antizipierte potentiell nötige Maßnahmen. *Hinweise: Vorbereitung der/s Defibrillators, Medikamente, Atemwegsmaterial?*	☐	☐	☐	☐	☐
Aufgabenmanagement:	**0**	**1**	**2**	**3**	**4**
10. Das Team priorisierte die Aufgaben.	☐	☐	☐	☐	☐
11. Das Team folgte anerkannten Standards und Leitlinien. *Hinweise: Sind Abweichungen möglicherweise angebracht?*	☐	☐	☐	☐	☐

Gesamtleistung:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
12. Vergeben Sie eine Gesamtbewertung für die nicht-medizinischen Fähigkeiten des Teams auf einer Skala von 1-10	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Kommentare:_____

_____

Translated by Julia Freytag & Fabian Stroben with friendly permission by Professor Simon Cooper, PhD
Contact: julia.freytag@charite.de, fabian.stroben@charite.de
Simulated Patients Program and Lernzentrum (skills lab), Office of the Vice Dean for Teaching and Learning, Charité-Universitätsmedizin Berlin, Germany

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	n.a.
	2b	All items from the World Health Organization Trial Registration Data Set	n.a.
Protocol version	3	Date and version identifier	1
Funding	4	Sources and types of financial, material, and other support	13
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	1
	5b	Name and contact information for the trial sponsor	n.a.
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	13
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	13

Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	3
	6b	Explanation for choice of comparators	3-4
Objectives	7	Specific objectives or hypotheses	4-5
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	5
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	5
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	5
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	4, 6-7
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	n.a.
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	9
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	n.a.
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	9-10
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Fig.1, 8-9

Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	10
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	6
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	6
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	6
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	6
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	6-8
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	n.a
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	9
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	n.a.

Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	____10_____
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	_____10_____
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	_____10_____
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	_____10_____
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	____n.a. _____
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	____n.a. _____
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	____n.a. _____
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	____n.a. _____
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	___12_____
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	____n.a. _____

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	12
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	n.a.
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	8,12
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	13
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	12
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	n.a.
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	12
	31b	Authorship eligibility guidelines and any intended use of professional writers	12
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	12
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	upon request
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	n.a.

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.