



Randomized controlled trials of simulation-based interventions in Emergency Medicine: a methodological review

Anthony Chauvin^{1,2,3} · Jennifer Truchot^{1,2,6} · Aida Bafeta³ · Dominique Pateron^{4,5} · Patrick Plaisance^{1,2} · Youri Yordanov^{3,4,5}

Received: 31 August 2017 / Accepted: 10 November 2017 / Published online: 16 November 2017
© SIMI 2017

Abstract

The number of trials assessing Simulation-Based Medical Education (SBME) interventions has rapidly expanded. Many studies show that potential flaws in design, conduct and reporting of randomized controlled trials (RCTs) can bias their results. We conducted a methodological review of RCTs assessing a SBME in Emergency Medicine (EM) and examined their methodological characteristics. We searched MEDLINE via PubMed for RCT that assessed a simulation intervention in EM, published in 6 general and internal medicine and in the top 10 EM journals. The Cochrane Collaboration risk of Bias tool was used to assess risk of bias, intervention reporting was evaluated based on the “template for intervention description and replication” checklist, and methodological quality was evaluated by the Medical Education Research Study Quality Instrument. Reports selection and data extraction was done by 2 independents researchers. From 1394 RCTs screened, 68 trials assessed a SBME intervention. They represent one quarter of our sample. Cardiopulmonary resuscitation (CPR) is the most frequent topic (81%). Random sequence generation and allocation concealment were performed correctly in 66 and 49% of trials. Blinding of participants and assessors was performed correctly in 19 and 68%. Risk of attrition bias was low in three-quarters of the studies ($n = 51$). Risk of selective reporting bias was unclear in nearly all studies. The mean MERQSI score was of 13.4/18.4% of the reports provided a description allowing the intervention replication. Trials assessing simulation represent one quarter of RCTs in EM. Their quality remains unclear, and reproducing the interventions appears challenging due to reporting issues.

Keywords Simulation · Randomized controlled trials · Quality · Emergency Medicine · Cardiopulmonary resuscitation · Reproducibility

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11739-017-1770-1>) contains supplementary material, which is available to authorized users.

✉ Anthony Chauvin
anthony.chauvin@aphp.fr

¹ Service d’Accueil des Urgences, Emergency Département, Hôpital Lariboisière, Assistance Publique-Hôpitaux de Paris, 2 Rue Ambroise Paré, 75010 Paris, France

² Faculté de Médecine, Université Diderot, Paris, France

³ INSERM U1153, Statistic and Epidemiologic Research Center Sorbonne Paris Cité (CRESS), METHODS Team, Hotel-Dieu Hospital, Paris, France

⁴ Sorbonne Universités, UPMC Paris Univ-06, Paris, France

⁵ Service des Urgences-Hôpital Saint Antoine, Assistance Publique-Hôpitaux de Paris (APHP), Paris, France

⁶ Ilumens Simulation Department, Paris Descartes University, 45 rue des Saint Pères, 75006 Paris, France

Introduction

Healthcare professionals have the ethical imperative to ensure patients’ safety while delivering optimal treatment. To develop the knowledge, skills, and attitudes of professionals, medical education involves the participation of live patients. These two situations seem difficult to reconcile, but could be reconciled by the use of simulation-based medical education (SBME) [1–3]. To learn and improve knowledge, clinical skills and attitudes [4–7], SBME provides healthcare professionals with a controlled practice environment such as computer-based virtual reality simulators, high-fidelity manikins, simple task trainers or actors posing as standardized patients [8]. Several trials report that SBME is effective in enhancing medical procedures, technical skills (i.e., central venous catheter placement, cardiopulmonary resuscitation), communication, teamwork, leadership and decision-making

[2, 9–20]. In Emergency Medicine (EM), SBME has evolved from the sparse use of low-fidelity manikins a decade ago, to high-fidelity simulation being fully integrated in numerous residency-training programs worldwide [21, 22].

The number of trials assessing SBME interventions has rapidly expanded [23, 24]. To be useful to research users, as for any healthcare intervention, SBME must be assessed by well-designed trials, then fully and transparently reported. Many studies show that potential flaws in design, conduct and reporting of randomized controlled trials (RCTs) can bias their results [25–30]. Even if poor reporting does not mean poor methods [31], adequate reporting allows readers to assess the strength and weakness of studies and improve the replication of interventions in daily practice [32, 33].

To our knowledge, the methodological characteristics of SBME trials in the field of EM have not been assessed. In this study, we aim to (1) assess the proportion of simulation trials among RCTs evaluating interventions in the field of EM, (2) describe and evaluate their methodological characteristics, and (3) estimate whether the reports adequately describe the intervention to allow replication in practice.

Methods

We performed a review of reports of RCTs assessing an SBME intervention in the field of Emergency Medicine that were published over a 4-year period. We used the Cochrane Collaboration risk of bias tool [34] and the Medical Education Research Study Quality Instrument (MERQSI) [35] to assess the risk of bias and methodological quality of included RCTs. We used reporting guidelines to evaluate intervention descriptions. We report this review in accordance with the Preferred Reporting of Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement [36] (Online Appendix 1).

Search strategy

We searched MEDLINE via PubMed for all reports of RCTs published from January 1, 2012 to December 31, 2015 in the 6 general and internal medicine journals (*New England Journal of Medicine*, *The Lancet*, *Journal of the American Medical Association*, *Annals of Internal Medicine*, *British Medical Journal* and *Archives of Internal Medicine*) and the 10 EM journals (*Resuscitation*, *Annals of Emergency Medicine*, *Emergencies*, *Injury*, *Scandinavian Journal of Trauma Resuscitation & Emergency Medicine*, *Academic Emergency Medicine*, *Prehospital Emergency Care*, *European Journal of Emergency Medicine*, *Emergency Medicine Australasia* and *World Journal Of Emergency Surgery*) with the highest impact factor according to the 2014 Web of Knowledge (search date: February 17, 2016). We applied no limitations

on language. We therefore, decided to use a period of 4 years and one database (i.e.; PubMed via Medline), with the hypothesis that this period and database would give us a large picture of simulation research in EM. The search strategy is reported in Online Appendix 2.

Study selection

We included all RCTs assessing SMBE interventions, regardless of type, which were performed in an emergency department or evaluated an emergency situation (e.g., cardiopulmonary situation). We excluded systematic reviews or meta-analyses, methodological publications, editorial style reviews, research letters, secondary analysis, abstracts and posters, correspondence, and protocols.

Two reviewers (AC and JT) independently examined all retained references based on the title and abstracts, then the full text of relevant studies, according to inclusion and exclusion criteria. Any disagreements were resolved by discussion with a third researcher (YY).

Data extraction

Two reviewers (AC and JT) extracted data in duplicate and independently using a standardized data extraction form. When assessments differed, the item was discussed until consensus was reached. If needed, a third reviewer (YY) was consulted.

General characteristic of RCTs assessing a SBME intervention

For each RCT, we assessed the following:

1. General characteristics: year of publication, name of journal, location of studies, publication time, number of centers involved, number of participants randomized and analyzed, type of participants involved (e.g., nurses, medical students), study design (parallel-arm or cross-over study), ethical committee approval, funding sources, reporting of registration number or study protocol available. We extracted primary outcomes as reported in the report. If the primary outcome was unclear, we used the outcome stated in the sample size calculation. We determined the EM topics related to (1) cardiopulmonary resuscitation (CPR), (2) airway management (without CPR situation), (3) triage, (4) surgical intervention (i.e., cricothyroidotomy) and (5) others.
2. Type of comparator (e.g., usual procedure or not) and the tested intervention. We defined the type of simulation as high or low-fidelity: high-fidelity manikins are “those that provide physical findings, display vital signs, physiologically respond to interventions (via computer

interface) and allow for procedures to be performed on them (e.g., bag mask ventilation, intubation, intravenous insertion)”, and low-fidelity manikins are “static mannequins that are otherwise limited in these capabilities” [37]. Simulation studies with cadavers were considered low-fidelity simulation.

Methodological quality of RCTs assessing an SBME intervention

Risk of bias assessment

The risk of bias within each RCT was evaluated by assessing the following key domains of the Cochrane Collaboration risk of bias tool [34]: selection bias (methods for random sequence generation and allocation concealment), performance bias (blinding of participants and personnel), detection bias (blinding of outcome assessors), attrition bias (incomplete outcome data), and reporting bias (selective outcomes reporting). Each domain was rated as low, high, or unclear risk of bias by the Cochrane handbook recommendations [38].

Methodological quality assessment

The methodological quality was appraised using the Medical Education Research Study Quality Instrument (MERSQI). The MERSQI is a 10-item scale developed to measure the methodological quality of trials assessing educational interventions by evaluating six domains [35]. The 10-item scale (range 5–18 because our study involved only RCTs) covers the following domains: study design (1–3 points), number of institutions studied (0.5–1.5 points), response rate (0.5–1.5 points), type of data (1 or 3 points), internal structure (0 or

1 point), content (0 or 1 point), relationship to other variables (0 or 1 point), appropriateness of the analysis (0 or 1 point), complexity of the analysis (1–2 point) and outcomes (1–3 points). A high score indicates high quality. Despite no predefined cut-off for high and low quality, one study used a MERSQI score of ≥ 14.0 as a reference value for high quality.

Intervention description and replication

We assessed how key methodological components of the SBME interventions were reported according to a modified checklist based on the “Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide” [33]: what (materials) (i.e., describe any physical or informational materials used in the intervention), what (procedures) (i.e., describe each of the procedures), who provided (i.e., description of intervention provider), where [i.e., describe the type(s) of location(s)], when and how much (i.e., report the number of times the intervention was delivered), how well (planned) (i.e., the intervention adherence) and how well delivered (actual) (i.e., assessment of the intervention adherence). See the complete descriptions in Table 1.

If authors correctly reported all key items, the intervention was considered reproducible. Items missing from the intervention description or not described in sufficient detail for replication were considered incomplete.

Data synthesis and analysis

Quantitative data with a normal distribution are reported with mean and standard deviation (SD). If the distribution was not normal, they are reported with median and

Table 1 Description of key methodological components of simulation-based medical education

What (materials)	Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers (i.e., a standardized training and information provided on tablets). Provide information on where the materials can be accessed (i.e., Online Appendix, URL)
What (procedures)	Describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities
Who provided	For each category of intervention provider, describe their expertise, background and any specific training given (i.e., training by experienced Basic Life Support instructor, certified by the European Resuscitation Council)
Where	Describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features (i.e., presentation of the simulation room with a photo, and technical references of used materials)
When and how much	Report the number of times the intervention was delivered and over what period of time including the number of sessions, schedule, duration, and intensity
How well (planned)	If intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them (i.e., training investigators in the method of enrollment and data collection until the study protocol could be completed without questions, further instruction, or error)
How well delivered (actual)	If intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned (i.e., to conduct a centralized review of one pre-study pilot session from each institution and all study videotapes to ensure compliance with the methodology)

interquartile range [IQR]. Qualitative data are reported with numbers (%). For risk of bias defined by the Cochrane Collaboration, we determined the frequency of presence of each bias item. We reported individual frequencies or scores for each item on the MERSQI. SAS 9.3 (SAS Institute, Inc., Cary, NC, USA) was used for all analyses.

Results

Search results

Our search identified 1394 reports of RCTs; 270 trials (19%) were in the field of EM and 68 (25%) assessed an SBME intervention (Fig. 1).

General characteristics of RCTs assessing an SBME intervention

The 68 reports of RCTs were published in 9 different journals (Table 2). All included study reports were published in EM journals, with 29 (43%) published in *Resuscitation*. About half of the studies were performed in Europe ($n = 32$, 47%) and 59 (87%) were monocentric. Cross-over trials accounted for 41% of our sample ($n = 28$). Most of the included studies had ethics committee approval ($n = 61$, 90%). Seven (10%) were registered at a public registry, or had an available protocol. The median number of participants randomized and analyzed was 46 [IQR 30; 77] and 41 [27; 72]. The most frequently studied populations were medical students ($n = 16$ studies, 24%), laypersons ($n = 13$, 19%) and Emergency Medical Service ($n = 10$, 15%).

Cardiopulmonary resuscitation (CPR) (Online Appendix 3) Most of the RCTs ($n = 55$, 81%) studied CPR [39–93]. Two-thirds ($n = 36$) focused on basic life support [39–74], and others a specific aspect of CPR (i.e., ventilation, chest

Fig. 1 Flow chart of the study

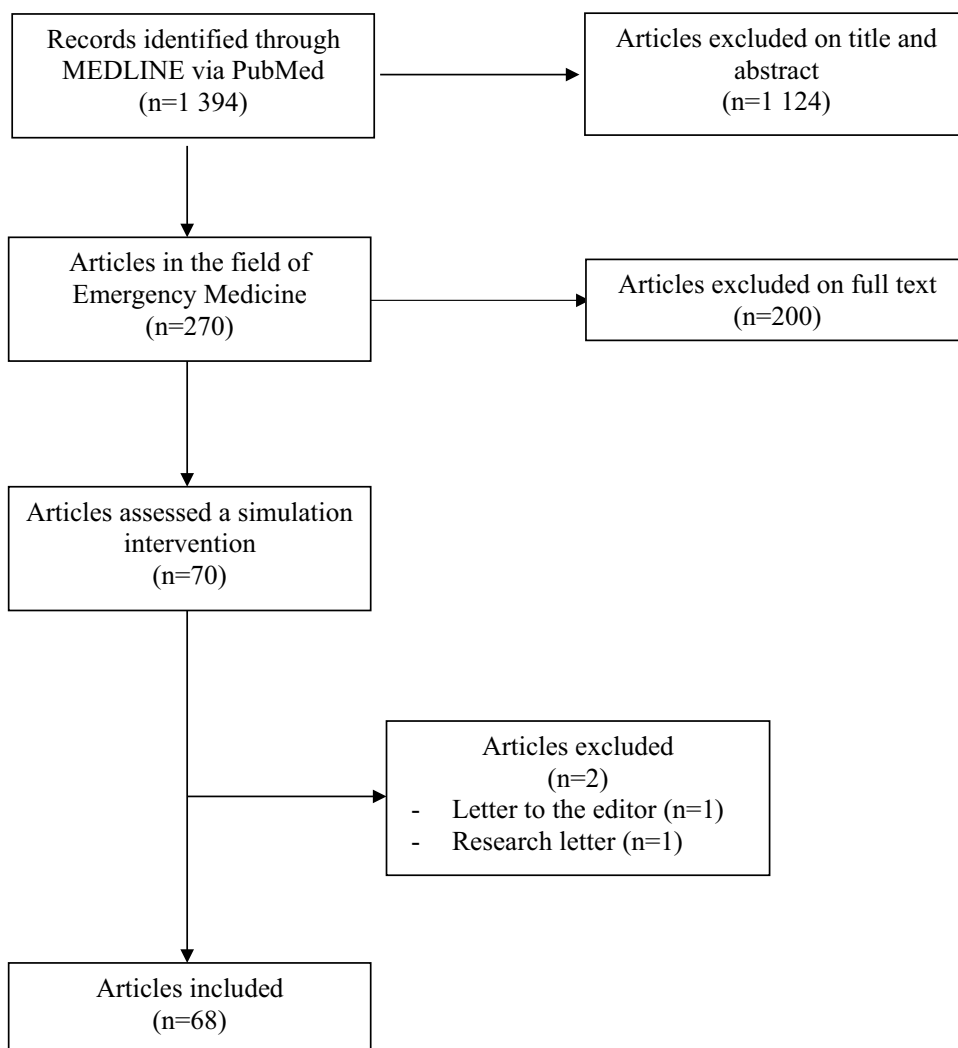


Table 2 General characteristics of the included studies ($n = 68$)

	n (%)
Year of publication	
2012	7 (10.3)
2013	27 (39.7)
2014	25 (36.8)
2015	9 (13.2)
Publishing journal	
<i>Resuscitation</i>	29 (42.6)
<i>Emergency Medicine Journal</i>	11 (16.2)
<i>Scandinavian Journal of Trauma Resuscitation & Emergency Medicine</i>	8 (11.8)
<i>European Journal of Emergency Medicine</i>	6 (8.7)
<i>The Journal of Emergency Medicine</i>	5 (7.4)
<i>Annals of Emergency Medicine</i>	4 (5.9)
<i>Prehospital Emergency Care</i>	3 (4.4)
<i>Academic Emergency Medicine</i>	1 (1.5)
<i>Emergency Medicine Australasia</i>	1 (1.5)
Country	
Europe	32 (47.1)
North America	19 (27.9)
Asia	17 (25)
Publication time, days, median [IQR]	145 [94;174]
Study design	
Parallel arms	40 (58.8)
Cross-over study	28 (41.2)
No. of centers involved	
Monocentric	59 (86.7)
Multicentric	9 (13.3)
Level of fidelity	
High-fidelity	11 (16.2)
Low-fidelity	57 (83.8)
Participants randomized, median [Q1; Q3]	46 [30; 77]
Participants analyzed, median [Q1; Q3]	41 [27; 72]
Participants	
Medical students	16 (23.5)
Laypersons	13 (19.1)
Emergency Medical Service	10 (14.7)
Emergency medical physicians or interns	9 (13.2)
Other students	8 (11.8)
Other medical specialists or paramedicals	6 (8.8)
Others	6 (8.8)
Topics	
CPR	55 (80.9)
Airway management (without CPR)	5 (7.4)
Triage	3 (4.4)
Emergency surgery	2 (2.9)
Other	3 (4.4)
Ethical committee approval	61 (89.7)
Registration or study protocol available	7 (10.3)

CPR cardiopulmonary resuscitation

Table 3 Bias assessment (according to the Cochrane Collaboration risk of bias tool [34]) of included studies ($n = 68$)

	Low risk	Unclear risk	High risk
Sequence generation	45 (66.2)	23 (33.8)	0
Allocation concealment	33 (48.5)	33 (48.5)	2 (3)
Blinding of participants	13 (19.1)	7 (10.3)	48 (70.6)
Blinding of outcome	46 (67.6)	18 (26.5)	4 (5.9)
Incomplete outcome	51 (75)	10 (14.7)	7 (10.3)
Selective reporting outcome	4 (5.9)	63 (92.6)	1 (1.5)

compression), or a specific population (i.e., children/neonates) [88–93]. All trials involved the use of simulators, but high-fidelity simulators were used in only seven studies (12%). In most cases, the comparator was the usual procedure ($n = 44$, 80%). CPR quality outcomes were reported for 41 trials (75%), but their definition varied across the trials (e.g., chest compression rate, mean chest compression, correct compression depth rate, correct recoil rate or correct hand position rate).

Airway management (Online Appendix 4) Ventilation or intubation was evaluated in five studies (7%) [94–98]. The aim of the trials was to test a new non-invasive mask ($n = 1$) [94], compare different airway management approaches ($n = 1$) [95], assess intubation for trauma patients ($n = 2$) [96, 97] or intubation in chemical, biological, radiation and nuclear situations ($n = 1$) [98]. All interventions involved low-fidelity simulators, and the comparator was the usual procedure. The outcome assessed in all five trials was the delay or success with ventilation ($n = 5$).

Triage (Online Appendix 5) Three RCTs (4%) evaluated triage in mass casualty incidents [99–101], using computer simulation or actors. Triage accuracy was the outcome used in each study ($n = 3$).

Emergency surgery (Online Appendix 6) The surgical airway was assessed in two RCTs (3%) on cadavers, and compared novel techniques to the usual procedure of cricothyroidotomy [102, 103].

Others (Online Appendix 7) The remaining RCTs ($n = 3$, 4%) evaluated a Glasgow scale scoring aid [104], the use of a novel medication delivery system during in situ simulation sessions [105], and teleconsultation for EM service teams [106].

Risk of bias assessment

Random sequence generation and allocation concealment were classified at unclear or high risk of bias for 34% ($n = 23$) and 49% ($n = 33$) of trials (Table 3). Improper reporting or absence of participant blinding led to classifying 81% ($n = 55$) of the trials at unclear or high risk of bias. Risk of bias due to assessor blinding was classified at

high or unclear risk of bias for 33% ($n = 22$) of the RCTs, with risk of attrition bias classified at high or unclear risk of bias for 25% ($n = 17$). Because trials were mostly unregistered (90%, $n = 61$), the risk of selective outcome reporting was unclear in most reports ($n = 63$, 93%).

Methodological quality assessment

Most studies were monocentric ($n = 59$; 86%) (Table 1). The response rate was $\geq 75\%$ in 56 studies (82%) (Table 4). Assessment of an objective and measurable outcome (i.e., chest compression depth, success rate of intubation) was

Table 4 Ratings from measures of study educational quality

Domain	Item (points awarded)	Present n (%)
Study design (maximum 3)	1. Study design	
	Single group cross-sectional or single group post-test only (1)	0 (0)
	Single group pre-/post-test (1.5)	0 (0)
	Observational two-groups (2)	0 (0)
	Randomized two-group (3)	68 (100)
Sampling (maximum 3)	2. No. of institutions studied	
	Not available	1 (1.5)
	1 (0.5)	59 (86.7)
	2 (1)	6 (8.8)
	> 2 (1.5)	2 (2.9)
	3. Response rate	
	Not available	2 (2.9)
< 50% or not reported (0.5)	7 (10.3)	
50–74% (1)	3 (4.4)	
$\geq 75\%$ (1.5)	56 (82.4)	
Type of data: outcome assessment (maximum 3)	4. Type of data	
	Subjective (1)	15 (22.1)
	Objective (3)	53 (77.9)
Validity of evaluation instrument (maximum 3)	5. Internal structure	
	Not applicable	0 (0)
	Not reported (0)	1 (1.5)
	Reported (1)	67 (98.5)
	6. Content	
	Not applicable	0 (0)
	Not reported (0)	4 (5.9)
	Reported (1)	64 (94.1)
	7. Relationship to other variables	
	Not applicable	9 (13.2)
	Not reported (0)	18 (26.5)
	Reported (1)	41 (60.3)
	Data analysis (maximum 3)	8. Appropriateness of analysis
Inappropriate (0)		0 (0)
Appropriate (1)		68 (100)
9. Complexity of analysis		
Descriptive analysis data (1)		68 (100)
Beyond descriptive analysis (2)	0 (0)	
Outcomes (maximum 3)	10. Outcomes	
	Satisfaction, attitudes, perceptions, opinions, general facts (1)	2 (2.9)
	Knowledge, skills (1.5)	66 (97.1)
	Behaviors (2)	0 (0)
	Patient/healthcare outcome (3)	0 (0)

used in most included studies ($n = 53$; 78%). The internal structure and the content of the evaluation instrument was correctly reported for 67 (99%) and 64 (94%) studies. Authors report the relationship to other variables for only 41 (60%) studies. The analysis of data is appropriate in all studies. Knowledge or skills is the outcome used in almost all studies ($n = 66$; 97%). The mean (SD) MERSQI score was 13.4 (1.3)/18. The lowest score was 10 and the highest was 15.5.

Intervention description and replication

Only three articles (4%) correctly report all the items of the modified TIDieR checklist for intervention descriptions (Table 5). The items most reported are the procedures ($n = 61$; 88%) and who provided the intervention ($n = 59$; 86%). Elements of materials used are reported for only 28 trials (41%). One-third ($n = 26$; 38%) report where the intervention occurred. The “when” and “how much” items are completely reported for 48 studies (70%). The adherence (planned and actual) to the intervention is correctly reported in only 10 (15%) and 5 (7%) articles.

Discussion

We aimed to conduct a methodological review of published RCTs assessing a simulation-based intervention in EM. Simulation represents up to 25% of our RCTs. The most frequent topic is CPR. Only half of the studies have low risk of bias for allocation concealment. Only 10 are registered

or have an available protocol. Despite these methodological difficulties, the methodological quality is high, with a mean (SD) MERQSI score 13.4 (1.3)/18. The description of the intervention is correctly reported for only 4% of studies.

One of the results of our study is the importance of simulation in EM research. Almost 15 years ago, a report by the Commonwealth Fund Task Force emphasized that the quality of care that patients receive could be to some extent determined by the quality of medical education students and residents receive [107, 108]. However, as stressed by Stephenson et al. [109] “Medical education is not short of excellent ideas about how to improve courses and create the professionals needed by society. What is in much shorter supply is evidence about the effectiveness of such teaching (...).” Another finding of our study is that, within our sample, more than one quarter of RCTs published in EM appear to be assessing educational interventions rather than patient-centered. To our knowledge, no such assessment has been performed in another medical specialty. Moreover, our analysis shows that the majority of the educational content in EM is CPR-based. This may indirectly suggest that simulation-based education is not used to its full potential to allow assessment of other interventions such as decision-making, communication and team work skills.

Several other studies evaluate the internal validity of published articles in certain specific simulation situations [110–112]. These studies indicate that simulation research is poorly reported. Systematic reviews of SBME also quantitatively document missing elements in abstracts, study design, definitions of variables, and study limitations [113–115]. The MERSQI score for our studies

Table 5 Reporting of key items for reproducibility of the intervention

Items for interventions	<i>n</i> (%)	All check list items, <i>n</i> (%)	“What, Who, Where and When” items, <i>n</i> (%)
What (materials): physical or informational materials used, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed	28 (40.6%)	3 (4.3)	11 (15.9)
What (procedures): the procedures, activities, and/or processes used in the intervention, including any enabling or support activities	61 (88.4%)		
Who provided: for each category of intervention provider, describe their expertise, background and any specific training given	59 (85.5%)		
Where: describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features	26 (37.7%)		
When and how much: describe the number of times the intervention was delivered and over what period of time including the number of sessions, their schedule, and their duration, intensity or dose	48 (69.6%)		
How well (planned): intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them	10 (14.5%)		
How well (actual): intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned	5 (7.2%)		

indicate that the educational quality of the studies appears fair. Bias related to the lack of blinding remains problematic when designing RCTs. Concealing study procedures from study participants is difficult, and probably impossible for some SBME interventions. However, alternate methods exist, such as having participants blinded to the hypothesis. For example, Philippon et al. assess whether the death of the manikin increases anxiety among learners as compared with a similar simulation-based course in which the manikin stays alive [116]. Participants were blinded to the study's objectives, and were advised that they were participating in a study designed to assess emotions while managing life-threatening situations. Additionally, when blinding trial participants is not possible, outcome assessors can often still be blinded to limit the risk of bias due to open RCTs. As developed by Kahan et al. [117], when blinding outcome assessment is not possible, strategies exist for reducing the possibility of bias (i.e., modify the outcome definition or method of assessment).

With only 4% of the interventions being fully reported, dissemination of the studies' findings is placed at risk. The lack of part of the information on the interventions may affect the ability of other researchers and educators to replicate them. Clear and precise recommendations on how SBME interventions should be reported should help improve trial transparency and enable the dissemination of efficient interventions and discard ineffective ones. A steering committee of 12 experts in simulation-based education and research recently developed specific reporting guidelines for simulation-based research [23]. These guidelines are an extension to the Consolidated Standards of Reporting Trials (CONSORT) [118]. The application and impact of these guidelines on quality of reporting will be of interest. However, the experts focused on the reporting of only key items of the CONSORT and not on the description of the intervention. Without a complete published description of the intervention, other researchers cannot replicate or build on research findings. The objective of the TIDieR is to improve the reporting of interventions, make it easier for readers to use the information, reduce wasteful research and increase the potential impact of research on health. Many editors endorse the CONSORT statement to improve the reporting of RCTs. However, the completeness of the reporting is only one aspect of the quality of the methodology. To avoid waste of research with lack of information on methods, authors, editors, and peer-reviewers must pay better attention to the reporting of key elements of the reproducibility [119]. Of note, biased and misreported studies contribute to an important waste in medical research, estimated at up to 85% of research investments each year [120–123].

Strength and limitations

Our study has several limitations. First, we searched only one database (MEDLINE via PubMed) without searching ERIC or EMBASE. However, our search was exhaustive and performed according to the Cochrane standards. Second, for the assessment of the methodological quality, the authors might have omitted key information from reports that was deleted during the publication process, and we were able to assess published reports only. Our convenient sample of journals might also have overestimated the overall quality because we arbitrarily selected journals with the highest methodological quality for selecting articles.

Conclusions

Trials assessing simulation account for one quarter of published RCTs' sample. Their quality remains unclear, and requires great caution in drawing conclusions from their results. In our sample, authors failed to correctly describe the blinding process, allocation concealment and key elements essential to insure the reproducibility of the intervention. Guidelines for improving the reproducibility of Simulation-Based Medical Education research are needed to help improve the interventions replication in daily practice.

Acknowledgements We thank Laura Smales (BioMedEditing, Toronto, ON) for copyediting the manuscript.

Author contributions Conception and design: AC, JT, PP, DP and YY; acquisition of data, AC and JT; analysis: AC; interpretation of data: AC, JT, PP, DP and YY; drafting the article: AC, AB, JT and YY; critical revision for important intellectual content: DP, AB and PP; final approval of the version to be published: AC, JT, PP, AB, DP and YY.

Compliance with ethical standards

Conflict of interest The authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare no support from any organization other than the funding agency listed above for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Statement of human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent For this type of study formal consent is not required.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Ziv A, Wolpe PR, Small SD et al (2003) Simulation-based medical education: an ethical imperative. *Acad Med* 78(8):783–788
- Cook DA (2014) How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. *Med Educ* 48(8):750–760
- Zendejas B, Brydges R, Wang AT et al (2013) Patient outcomes in simulation-based medical education: a systematic review. *J Gen Intern Med* 28(8):1078–1089
- Society for simulation in Healthcare (2016) About simulation. <http://ssih.org/About-Simulation>. Accessed 26 Dec 2016
- Issenberg SB, McGaghie WC, Petrusa ER et al (2005) Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 27(1):10–28
- Issenberg SB, McGaghie WC, Hart IR et al (1999) Simulation technology for health care professional skills training and assessment. *JAMA* 282(9):861–866
- Fincher RME, Lewis LA (2002) Simulations used to teach clinical skills. In: Norman GR, van der Vleuten CPM, Newble DI (eds) *International handbook of research in Medical Education*, part one. Kluwer Academic, Dordrecht
- Haute Autorité de santé. Rapport de mission-Etat de l'art (national et international) en matière de pratiques de simulation dans le domaine de santé (2012). http://www.has-sante.fr/portail/jcms/c_930641/fr/simulation-en-sante. Accessed 22 Mar 2017
- Wayne DB, Butter J, Siddall VJ et al (2005) Simulation-based training of internal medicine residents in advanced cardiac life support protocols: a randomized trial. *Teach Learn Med* 17(3):210–216
- Barsuk JH, McGaghie WC, Cohen ER et al (2009) Use of simulation based mastery learning to improve the quality of central venous catheter placement in a medical intensive care unit. *J Hosp Med* 4(7):397–403
- Murray DJ, Boulet JR, Avidan M et al (2007) Performance of residents and anesthesiologists in a simulation based skill assessment. *Anesthesiology* 107(5):705–713
- Roter DL, Larson S, Shinitzky H et al (2004) Use of an innovative video feedback technique to enhance communication skills training. *Med Educ* 38(2):145–157
- Hardoff D, Gefen A, Sagi D et al (2016) Dignity in adolescent health care: a simulation-based training programme. *Med Educ* 50(5):570–571
- Mashiach R, Mezhybovsky V, Nevler A et al (2014) Three-dimensional imaging improves surgical skill performance in a laparoscopic test model for both experienced and novice laparoscopic surgeons. *Surg Endosc* 28(12):3489–3493
- Douglass AM, Elder J, Watson R et al (2017) A Randomized controlled trial on the effect of a double check on the detection of medication errors. *Ann Emerg Med* S0196–0644(17):30318–30319
- Unterman A, Achiron A, Gat I et al (2014) A novel simulation-based training program to improve clinical teaching and mentoring skills. *Isr Med Assoc J* 16(3):184–190
- Cohen AG, Kitai E, David SB et al (2014) Standardized patient-based simulation training as a tool to improve the management of chronic disease. *Simul Healthc* 9(1):40–47
- Harnof S, Hadani M, Ziv A et al (2013) Simulation-based interpersonal communication skills training for neurosurgical residents. *Isr Med Assoc J* 15(9):489–492
- Nelson D, Ziv A, Bandali KS (2013) Republished: going glass to digital: virtual microscopy as a simulation-based revolution in pathology and laboratory science. *Postgrad Med J* 89(1056):599–603
- Reis S, Sagi D, Eisenberg O et al (2013) The impact of residents' training in Electronic Medical Record (EMR) use on their competence: report of a pragmatic trial. *Patient Educ Couns* 93(3):515–521
- Bristol Medical Simulation Centre. Worldwide sim database. <http://www.bmsc.co.uk/>. Accessed 26 Dec 2016
- Okuda Y, Bond W, Bonfante G et al (2011) National growth in simulation training within Emergency Medicine Residency Program, 2003–2008. *Acad Emerg Med* 12:455–460
- Cheng A, Kessler D, Mackinnon R et al (2016) Reporting guidelines for health care simulation research: extensions to the CONSORT and STROBE statements. *Adv Simul* 1(1):25
- Cook DA, Hatala R, Brydges R et al (2011) Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 306(9):978–988
- Nuesch E, Reichenbach S, Trelle S et al (2009) The importance of allocation concealment and patient blinding in osteoarthritis trials: a meta-epidemiologic study. *Arthritis Rheum* 61(12):1633–1641
- Savovic J, Jones HE, Altman DG et al (2012) Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 157(6):429–438
- Wood L, Egger M, Gluud LL et al (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 336(7644):601–605
- Psaty BM, Prentice RL (2010) Minimizing bias in randomized trials: the importance of blinding. *JAMA* 304(7):793–794
- Nuesch E, Trelle S, Reichenbach S et al (2009) The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 339:b3244
- Tierney JF, Stewart LA (2005) Investigating patient exclusion bias in meta-analysis. *Int J Epidemiol* 34(1):79–87
- Soares HP, Daniels S, Kumar A et al (2004) Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 328(7430):22–24
- Ioannidis JP (2016) Why most clinical research is not useful. *PLoS Med* 13(6):e1002049
- Hoffmann TC, Glasziou PP, Boutron I et al (2014) Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 7(348):g1687
- Higgins JP, Altman DG, Gotzsche PC et al (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343:d5928
- Reed DA, Cook DA, Beckman TJ et al (2007) Association between funding and quality of published medical education research. *JAMA* 298(9):1006–1009
- Moher D, Liberati A, Tetzlaff J et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6(7):e1000097
- Cheng A, Lockey A, Bhanji F et al (2015) The use of high-fidelity manikins for advanced life support training—a systematic review and meta-analysis. *Resuscitation* 93(8):142–149
- Higgins JPT, Green S (2011) *Cochrane handbook for systematic reviews of interventions* version 5.1.0 [updated March 2011]: Wiley-Blackwell, Chapter 9.5. <http://handbook.cochrane.org/>. Accessed 21 Mar 2017
- Allan KS, Wong N, Aves T et al (2013) The benefits of a simplified method for CPR training of medical professionals: a randomized controlled study. *Resuscitation* 84(8):1119–1124
- Birkenes TS, Myklebust H, Kramer-Johansen J (2013) New pre-arrival instructions can avoid abdominal hand placement for chest compressions. *Scand J Trauma Resusc Emerg Med* 22(6):21–47

41. Birkenes TS, Myklebust H, Neset A et al (2014) Quality of CPR performed by trained bystanders with optimized pre-arrival instructions. *Resuscitation* 85(1):124–130
42. Godfred R, Huszti E, Fly D et al (2013) A randomized trial of video self-instruction in cardiopulmonary resuscitation for lay persons. *Scand J Trauma Resusc Emerg Med* 10(5):21–36
43. Hafner JW, Jou AC, Wang H et al (2015) Death before disco: the effectiveness of a musical metronome in layperson cardiopulmonary resuscitation training. *J Emerg Med* 48(1):43–52
44. Hong CK, Park SO, Jeong HH et al (2014) The most effective rescuer's position for cardiopulmonary resuscitation provided to patients on beds: a randomized, controlled, crossover mannequin study. *J Emerg Med* 46(5):643–649
45. Hunt EA, Heine M, Shilkofski NS et al (2015) Exploration of the impact of a voice activated decision support system (VADSS) with video on resuscitation performance by lay rescuers during simulated cardiopulmonary arrest. *Emerg Med J* 32(3):189–194
46. Iserbyt P, Byra M (2013) The design of instructional tools affects secondary school students' learning of cardiopulmonary resuscitation (CPR) in reciprocal peer learning: a randomized controlled trial. *Resuscitation* 84(11):1591–1595
47. Iserbyt P, Charlier N, Mols L (2014) Learning basic life support (BLS) with tablet PCs in reciprocal learning at school: are videos superior to pictures? A randomized controlled trial. *Resuscitation* 85(6):809–813
48. Iserbyt P, Mols L, Charlier N et al (2014) Reciprocal learning with task cards for teaching Basic Life Support (BLS): investigating effectiveness and the effect of instructor expertise on learning outcomes. A randomized controlled trial. *J Emerg Med* 46(1):85–94
49. Jones CM, Thorne CJ, Colter PS et al (2013) Rescuers may vary their side of approach to a casualty without impact on cardiopulmonary resuscitation performance. *Emerg Med J* 30(1):74–75
50. Kovic I, Lulic D, Lulic I (2013) CPR PRO(R) device reduces rescuer fatigue during continuous chest compression cardiopulmonary resuscitation: a randomized crossover trial using a manikin model. *J Emerg Med* 45(4):570–577
51. Krogh KB, Høyer CB, Ostergaard D et al (2014) Time matters—realism in resuscitation training. *Resuscitation* 85(8):1093–1098
52. Li Q, Zhou RH, Liu J et al (2013) Pre-training evaluation and feedback improved skills retention of basic life support in medical students. *Resuscitation* 84(9):1274–1278
53. Min MK, Yeom SR, Ryu JH et al (2013) A 10-s rest improves chest compression quality during hands-only cardiopulmonary resuscitation: a prospective, randomized crossover study using a manikin model. *Resuscitation* 84(9):1279–1284
54. Mpotos N, Calle P, Deschepper E et al (2013) Retraining basic life support skills using video, voice feedback or both: a randomized controlled trial. *Resuscitation* 84(1):72–77
55. Mpotos N, De Wever B, Cleymans N et al (2014) Repetitive sessions of formative self-testing to refresh CPR skills: a randomized non-inferiority trial. *Resuscitation* 85(9):1282–1286
56. Na JU, Lee TR, Kang MJ et al (2014) Basic life support skill improvement with newly designed renewal programme: cluster randomised study of small-group-discussion method versus practice-while-watching method. *Emerg Med J* 31(12):964–969
57. Nishiyama C, Iwami T, Kitamura T et al (2014) Long-term retention of cardiopulmonary resuscitation skills after shortened chest compression-only training and conventional training: a randomized controlled trial. *Acad Emerg Med* 21(1):47–54
58. Nishiyama C, Iwami T, Murakami Y et al (2015) Effectiveness of simplified 15-min refresher BLS training program: a randomized controlled trial. *Resuscitation* 90(5):56–60
59. Painter I, Chavez DE, Ike BR et al (2014) Changes to DA-CPR instructions: can we reduce time to first compression and improve quality of bystander CPR? *Resuscitation* 85(9):1169–1173
60. Panchal AR, Meziab O, Stolz U et al (2014) The impact of ultra-brief chest compression-only CPR video training on responsiveness, compression rate, and hands-off time interval among bystanders in a shopping mall. *Resuscitation* 9:1287–1290
61. Park SO, Hong CK, Shin DH et al (2013) Efficacy of metronome sound guidance via a phone speaker during dispatcher-assisted compression-only cardiopulmonary resuscitation by an untrained layperson: a randomised controlled simulation study using a manikin. *Emerg Med J* 30(8):657–661
62. Rössler B, Ziegler M, Hüpfel M et al (2013) Can a flowchart improve the quality of bystander cardiopulmonary resuscitation? *Resuscitation* 84(7):982–986
63. Semeraro F, Frisoli A, Loconsole C et al (2013) Motion detection technology as a tool for cardiopulmonary resuscitation (CPR) quality training: a randomised crossover mannequin pilot study. *Resuscitation* 84(4):501–507
64. Shin J, Hwang SY, Lee HJ et al (2014) Comparison of CPR quality and rescuer fatigue between standard 30:2 CPR and chest compression-only CPR: a randomized crossover manikin trial. *Scand J Trauma Resusc Emerg Med* 28(10):22–59
65. Sopka S, Biermann H, Rossaint R et al (2013) Resuscitation training in small-group setting—gender matters. *Scand J Trauma Resusc Emerg Med* 16(21):30
66. Taelman DG, Huybrechts SA, Peersman W et al (2014) Quality of resuscitation by first responders using the 'public access resuscitator': a randomized manikin study. *Eur J Emerg Med* 21(6):409–417
67. Van de Velde S, Roex A, Vangronsveld K et al (2013) Can training improve laypersons helping behaviour in first aid? A randomized controlled deception trial. *Emerg Med J* 30(4):292–297
68. Van Raemdonck V, Monsieurs KG, Aerenhouts D et al (2014) Teaching basic life support: a prospective randomized study on low-cost training strategies in secondary schools. *Eur J Emerg Med* 21(4):284–290
69. van Tulder R, Roth D, Krammel M et al (2014) Effects of repetitive or intensified instructions in telephone assisted, bystander cardiopulmonary resuscitation: an investigator-blinded, 4-armed, randomized, factorial simulation trial. *Resuscitation* 85(1):112–118
70. van Tulder R, Roth D, Havel C et al (2014) Push as hard as you can instruction for telephone cardiopulmonary resuscitation: a randomized simulation study. *J Emerg Med* 46(3):363–370
71. van Tulder R, Laggner R, Kienbacher C et al (2015) The capability of professional- and lay-rescuers to estimate the chest compression-depth target: a short, randomized experiment. *Resuscitation* 89(4):137–141
72. Wutzler A, Bannehr M, von Ulmenstein S et al (2015) Performance of chest compressions with the use of a new audio-visual feedback device: a randomized manikin study in health care professionals. *Resuscitation* 87(2):81–85
73. Yeung J, Davies R, Gao F et al (2014) A randomised control trial of prompt and feedback devices and their impact on quality of chest compressions—a simulation study. *Resuscitation* 85(4):553–559
74. Zapletal B, Greif R, Stumpf D et al (2014) Comparing three CPR feedback devices and standard BLS in a single rescuer scenario: a randomised simulation study. *Resuscitation* 85(4):560–566
75. Goliash G, Ruetzler A, Fischer H et al (2013) Evaluation of advanced airway management in absolutely inexperienced hands: a randomized manikin trial. *Eur J Emerg Med* 20(5):310–314
76. Gruber C, Nabecker S, Wohlfarth P et al (2013) Evaluation of airway management associated hands-off time during cardiopulmonary resuscitation: a randomised manikin follow-up study. *Scand J Trauma Resusc Emerg Med* 21:10
77. Jensen JL, Walker M, LeRoux Y et al (2013) Chest compression fraction in simulated cardiac arrest management by primary

- care paramedics: king laryngeal tube airway versus basic airway management. *Prehosp Emerg Care* 17(2):285–290
78. Ott T, Fischer M, Limbach T et al (2015) The novel intubating laryngeal tube (iLTS-D) is comparable to the intubating laryngeal mask (Fastrach)—a prospective randomised manikin study. *Scand J Trauma Resusc Emerg Med* 23:44
 79. Otten D, Liao MM, Wolken R et al (2014) Comparison of bag-valve-mask hand-sealing techniques in a simulated model. *Ann Emerg Med* 63(1):6–12
 80. Reiter DA, Strother CG, Weingart SD (2013) The quality of cardiopulmonary resuscitation using supraglottic airways and intraosseous devices: a simulation trial. *Resuscitation* 84(1):93–97
 81. Shin DH, Choi PC, Na JU et al (2013) Utility of the Pentax-AWS in performing tracheal intubation while wearing chemical, biological, radiation and nuclear personal protective equipment: a randomised crossover trial using a manikin. *Emerg Med J* 30(7):527–531
 82. Tandon N, McCarthy M, Forehand B et al (2014) Comparison of intubation modalities in a simulated cardiac arrest with uninterrupted chest compressions. *Emerg Med J* 31(10):799–802
 83. Cheng A, Overly F, Kessler D et al (2015) Perception of CPR quality: influence of CPR feedback, just-in-time CPR training and provider role. *Resuscitation* 87(2):44–50
 84. Chung TN, Bae J, Kim EC et al (2013) Induction of a shorter compression phase is correlated with a deeper chest compression during metronome-guided cardiopulmonary resuscitation: a manikin study. *Emerg Med J* 30(7):551–554
 85. Eisenberg Chavez D, Meischke H, Painter I et al (2013) Should dispatchers instruct lay bystanders to undress patients before performing CPR? A randomized simulation study. *Resuscitation* 84(7):979–981
 86. Foo NP, Chang JH, Su SB et al (2013) A stabilization device to improve the quality of cardiopulmonary resuscitation during ambulance transportation: a randomized crossover trial. *Resuscitation* 84(11):1579–1584
 87. Oh J, Chee Y, Lim T et al (2014) Chest compression with kneeling posture in hospital cardiopulmonary resuscitation: a randomised crossover simulation study. *Emerg Med Australas* 26(6):585–590
 88. Fuerch JH, Yamada NK, Coelho PR et al (2015) Impact of a novel decision support tool on adherence to Neonatal Resuscitation Program algorithm. *Resuscitation* 88(3):52–56
 89. Kim S, You JS, Lee HS et al (2013) Quality of chest compressions performed by inexperienced rescuers in simulated cardiac arrest associated with pregnancy. *Resuscitation* 84(1):98–102
 90. Krogh LQ, Bjørnshave K, Vestergaard LD et al (2015) E-learning in pediatric basic life support: a randomized controlled non-inferiority study. *Resuscitation* 90(5):7–12
 91. Martin P, Theobald P, Kemp A et al (2013) Real-time feedback can improve infant manikin cardiopulmonary resuscitation by up to 79%—a randomised controlled trial. *Resuscitation* 84(8):1125–1130
 92. Rodriguez SA, Sutton RM, Berg MD et al (2014) Simplified dispatcher instructions improve bystander chest compression quality during simulated pediatric resuscitation. *Resuscitation* 85(1):119–123
 93. Kim MJ, Lee HS, Kim S et al (2015) Optimal chest compression technique for paediatric cardiac arrest victims. *Scand J Trauma Resusc Emerg Med* 23:36
 94. Lee HY, Jeung KW, Lee BK et al (2013) The performances of standard and ResMed masks during bag-valve-mask ventilation. *Prehosp Emerg Care* 17(2):235–240
 95. Leventis C, Chalkias A, Sampanis MA et al (2014) Emergency airway management by paramedics: comparison between standard endotracheal intubation, laryngeal mask airway, and I-gel. *Eur J Emerg Med* 21(5):371–373
 96. Park SO, Shin DH, Lee KR et al (2013) Efficacy of the Disposcope endoscope, a new video laryngoscope, for endotracheal intubation in patients with cervical spine immobilisation by semirigid neck collar: comparison with the Macintosh laryngoscope using a simulation study on a manikin. *Emerg Med J* 30(4):270–274
 97. Schober P, Krage R, van Groeningen D et al (2014) Inverse intubation in entrapped trauma casualties: a simulator based, randomised cross-over comparison of direct, indirect and video laryngoscopy. *Emerg Med J* 31(12):959–963
 98. Shin DH, Han SK, Choi PC et al (2013) Tracheal intubation during chest compressions performed by qualified emergency physicians unfamiliar with the Pentax-Airwayscope. *Eur J Emerg Med* 20(3):187–192
 99. Jones N, White ML, Tofil N et al (2014) Randomized trial comparing two mass casualty triage systems (JumpSTART versus SALT) in a pediatric simulated mass casualty event. *Prehosp Emerg Care* 18(3):417–423
 100. Luigi Ingrassia P, Ragazzoni L, Carezzo L et al (2015) Virtual reality and live simulation: a comparison between two simulation tools for assessing mass casualty triage skills. *Eur J Emerg Med* 22(2):121–127
 101. Wolf P, Bigalke M, Graf BM et al (2014) Evaluation of a novel algorithm for primary mass casualty triage by paramedics in a physician manned EMS system: a dummy based trial. *Scand J Trauma Resusc Emerg Med* 22:50
 102. Helm M, Hossfeld B, Jost C et al (2013) Emergency cricothyroidotomy performed by inexperienced clinicians—surgical technique versus indicator-guided puncture technique. *Emerg Med J* 30(8):646–649
 103. Mabry RL, Nichols MC, Shiner DC et al (2014) A comparison of two open surgical cricothyroidotomy techniques by military medics using a cadaver model. *Ann Emerg Med* 63(1):1–5
 104. Feldman A, Hart KW, Lindsell CJ et al (2015) Randomized controlled trial of a scoring aid to improve Glasgow Coma Scale scoring by emergency medical services providers. *Ann Emerg Med* 65(3):325–329
 105. Moreira ME, Hernandez C, Stevens AD et al (2015) Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated emergency department pediatric resuscitations. *Ann Emerg Med* 66(2):97–106
 106. Rörtgen D, Bergrath S, Rossaint R et al (2013) Comparison of physician staffed emergency teams with paramedic teams assisted by telemedicine—a randomized, controlled simulation study. *Resuscitation* 84(1):85–92
 107. The Commonwealth fund task force on academic health centers (2002) Training tomorrow's doctors: the Medical Education Mission of Academic Health Centers. The Commonwealth Fund, New York. <http://www.commonwealthfund.org/publications/fund-reports/2002/apr/training-tomorrows-doctors-the-medical-education-mission-of-academic-health-centers>. Accessed 02 June 2017
 108. Epstein RM (2007) Assessment in medical education. *N Engl J Med* 356(4):387–396
 109. Stephenson A, Higgs R, Sugarman J (2001) Teaching professional development in medical schools. *Lancet* 357(9259):867–870
 110. Cheng A, Lang T, Starr S et al (2014) Technology-enhanced simulation and pediatric education: a meta-analysis. *Pediatrics* 133(5):e1313–e1323
 111. Mundell WC, Kennedy CC, Szostek JH et al (2013) Simulation technology for resuscitation training: a systematic review and meta-analysis. *Resuscitation* 84(9):1174–1183
 112. Cook DA, Hamstra SJ, Brydges R et al (2013) Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. *Med Teacher* 35(1):e867–e898

113. Cook DA, Beckman TJ, Bordage G (2007) A systematic review of titles and abstracts of experimental studies in medical education: many informative elements missing. *Med Educ* 41(11):1074–1081
114. Cook DA, Beckman TJ, Bordage G (2007) Quality of reporting of experimental studies in medical education: a systematic review. *Med Educ* 41(8):737–745
115. Cook DA, Levinson AJ, Garside S (2011) Method and reporting quality in health professions education research: a systematic review. *Med Educ* 45(3):227–238
116. Philippon AL, Bokobza J, Bloom B et al (2016) Effect of simulated patient death on emergency worker's anxiety: a cluster randomized trial. *Ann Intensive Care* 6(1):60
117. Kahan BC, Cro S, Doré CJ et al (2014) Reducing bias in open-label trials where blinded outcome assessment is not feasible: strategies from two randomised trials. *Trials* 15:456
118. Schulz KF, Altman DG, Moher D, For the CONSORT Group (2010) CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 8:18
119. Friedman LP, Cockburn IM, Simcoe TS (2015) The economics of reproducibility in preclinical research. *PLoS Biol* 13(6):e1002165
120. Chalmers I, Glasziou P (2009) Avoidable waste in the production and reporting of research evidence. *Lancet* 374:86–89
121. Glasziou P, Altman DG, Bossuyt P et al (2014) Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 383:267–276
122. Ioannidis JP, Greenland S, Hlatky MA et al (2014) Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 383:166–175
123. Macleod MR, Michie S, Roberts I et al (2014) Biomedical research: increasing value, reducing waste. *Lancet* 383:101–104