



# Transfer of learning and patient outcome in simulated crisis resource management: a systematic review

## Transfert de l'apprentissage et évolution des patients dans une gestion de crise des ressources simulée: une revue systématique

Sylvain Boet, MD · M. Dylan Bould, MBChB · Lillia Fung, MD ·  
Haytham Qosa, MD · Laure Perrier, MLIS · Walter Tavares, PhD(c) ·  
Scott Reeves, PhD · Andrea C. Tricco, PhD

Received: 12 September 2013 / Accepted: 7 March 2014 / Published online: 25 March 2014  
© The Author(s) 2014. This article is published with open access at [Springerlink.com](http://Springerlink.com)

### Abstract

**Purpose** Simulation-based learning is increasingly used by healthcare professionals as a safe method to learn and practice non-technical skills, such as communication and leadership, required for effective crisis resource management (CRM). This systematic review was conducted to gain a better understanding of the impact of simulation-based CRM teaching on transfer of learning to

the workplace and subsequent changes in patient outcomes.

**Source** Studies on CRM, crisis management, crew resource management, teamwork, and simulation published up to September 2012 were searched in MEDLINE®, EMBASE™, CINAHL, Cochrane Central Register of Controlled Trials, and ERIC. All studies that used simulation-based CRM teaching with outcomes measured at Kirkpatrick Level 3 (transfer of learning to the workplace) or 4 (patient outcome) were included. Studies measuring only learners' reactions or simple learning (Kirkpatrick Level 1 or 2, respectively) were excluded. Two authors independently reviewed all identified titles and abstracts for eligibility.

**Principal findings** Nine articles were identified as meeting the inclusion criteria. Four studies measured transfer of simulation-based CRM learning into the clinical setting (Kirkpatrick Level 3). In three of these studies, simulation-enhanced CRM training was found significantly more effective than no intervention or didactic teaching. Five studies measured patient outcomes (Kirkpatrick Level 4). Only one of these studies found that simulation-based CRM training made a clearly significant impact on patient mortality.

**Conclusions** Based on a small number of studies, this systematic review found that CRM skills learned at the simulation centre are transferred to clinical settings, and the acquired CRM skills may translate to improved patient outcomes, including a decrease in mortality.

This article is accompanied by an editorial. Please see Can J Anesth 2014; 61: this issue.

**Author contributions** Sylvain Boet was the principal investigator responsible for the concept of this review. He wrote the protocol, wrote and revised the manuscript, and was responsible for coordination of the study and communication with all co-authors. Andrea C. Tricco, senior investigator, reviewed earlier versions of the protocol and critically reviewed earlier versions of the manuscript. Sylvain Boet and Andrea C. Tricco supervised all steps of the study. M. Dylan Bould and Walter Tavares, co-investigators, critically reviewed the protocol. Lillia Fung and Haytham Qosa, co-investigators, participated in data extraction. Laure Perrier, co-investigator, was responsible for the literature search. Scott Reeves, co-investigator, shared his methodological expertise. Sylvain Boet, M. Dylan Bould, Lillia Fung, Haytham Qosa, Laure Perrier, Walter Tavares, and Andrea C. Tricco participated in screening. Sylvain Boet, M. Dylan Bould, Lillia Fung, Haytham Qosa, Laure Perrier, Scott Reeves, and Andrea C. Tricco participated in interpretation of the data. M. Dylan Bould, Lillia Fung, Haytham Qosa, Laure Perrier, Walter Tavares, and Scott Reeves critically reviewed the final manuscript.

**Electronic supplementary material** The online version of this article (doi:[10.1007/s12630-014-0143-8](https://doi.org/10.1007/s12630-014-0143-8)) contains supplementary material, which is available to authorized users.

S. Boet, MD (✉) · L. Fung, MD · H. Qosa, MD  
Department of Anesthesiology, The Ottawa Hospital, University of Ottawa, General Campus, 501 Smyth Rd, Critical Care Wing 1401, Ottawa, ON, Canada  
e-mail: [sboet@toh.on.ca](mailto:sboet@toh.on.ca)

S. Boet, MD · M. D. Bould, MBChB  
The Ottawa Hospital Research Institute, University of Ottawa Skills and Simulation Centre (uOSSC), University of Ottawa, ON, Canada

## Résumé

**Objectif** *L'apprentissage basé sur des simulations est de plus en plus utilisé par les professionnels de santé comme méthodes sécuritaires d'apprentissage et de pratique de compétences non techniques, comme la communication et le leadership, qui sont nécessaires pour une gestion efficace des ressources en situation de crise (CRM). Cette étude systématique a été menée pour mieux comprendre l'impact de l'enseignement à partir de simulations de la CRM sur le transfert des connaissances sur le lieu de travail et les changements ultérieurs sur l'évolution des patients.*

**Source** *Les études sur la CRM, gestion de crise, gestion de ressources d'équipes, travail d'équipe et simulation, publiées jusqu'en septembre 2012 ont été recherchées dans les bases de données MEDLINE®, EMBASE™, CINAHL, Cochrane Central Register of Controlled Trials et ERIC. Toutes les études utilisant un enseignement de la CRM à partir de simulations avec des résultats mesurés au niveau 3 de Kirkpatrick (transfert de l'apprentissage au milieu de travail) ou au niveau 4 (évolution du patient) ont été incluses. Toutes les études ne mesurant que les réactions des apprenants ou le seul apprentissage (respectivement, niveau 1 ou 2 de Kirkpatrick) ont été exclues. Deux auteurs ont revu de façon indépendante tous les titres et résumés identifiés pour évaluer leur admissibilité.*

**Constataions principales** *Neuf articles répondant aux critères d'inclusion ont été identifiés. Quatre études mesuraient le transfert d'apprentissage de la CRM à partir de simulations vers un cadre clinique (niveau 3 de Kirkpatrick). Dans trois de ces études, la formation à la CRM soutenue par des simulations s'est avérée*

*significativement plus efficace que l'absence d'intervention ou un enseignement didactique. Cinq études mesuraient les résultats pour les patients (niveau 4 de Kirkpatrick). Une seule de ces études a trouvé que la formation à la CRM basée sur des simulations avait un impact clairement significatif sur la mortalité des patients.*

**Conclusions** *Reposant sur un petit nombre d'études, cette analyse systématique a trouvé que les habiletés en matière de CRM apprises au centre de simulations sont transférées dans des cadres cliniques et que les habiletés acquises de CRM peuvent se traduire par une amélioration de l'évolution, y compris une baisse de la mortalité.*

Simulators are increasingly recognized as useful educational tools in healthcare<sup>1</sup> for both technical and non-technical skills.<sup>2-4</sup> Within acute care specialties, these tools are also used for various training purposes, including simulating rare events<sup>5,6</sup> and teaching technical skills<sup>7</sup> or advanced life support algorithms.<sup>8</sup> The simulation room is an ideal setting for teaching the principles of crisis resource management (CRM).<sup>9</sup> In a simulated crisis, vital non-technical skills, such as task management, teamwork, situation awareness, and decision-making can be safely practiced. The ultimate goal of all CRM simulation training is to increase patient safety and result in better patient outcomes. Although numerous studies have been published on the topic, there is a need for a knowledge synthesis of the impact that simulation-based CRM training has on patient outcomes and on the performance of healthcare providers in the workplace.

There have been previous systematic reviews on simulation-based education and non-technical skills. Gordon *et al.*<sup>10</sup> investigated “any studies involving an educational intervention to improve non-technical skills amongst undergraduate or postgraduate staff in an acute health care environment.” While their review addresses training for non-technical skills, their paper is neither specific to crisis scenarios nor to simulation. To examine CRM programs for postgraduate trainees (i.e., residents), Doumouras *et al.*<sup>11</sup> summarized the design, implementation, and efficacy of simulation-based CRM training programs in the peer-reviewed literature. Nevertheless, this review included simulation-based training only for residents. Their findings supported the utility of CRM programs for residents and a high degree of satisfaction with perceived value reflected by robust resident engagement. They concluded, however, that “a dearth of well-designed, randomized studies preclude the quantification of impact of simulation-based training in the clinical environment.”

The existing literature does not address the downstream effects (i.e., transfer of learning and patient outcome) of CRM simulation-based education. To assess the impact of

S. Boet, MD · M. D. Bould, MBChB  
The Academy for Innovation in Medical Education, University  
of Ottawa, Ottawa, ON, Canada

M. D. Bould, MBChB  
Department of Anesthesiology, The Children's Hospital of  
Eastern Ontario Research Institute, The Children's Hospital of  
Eastern Ontario, Ottawa, ON, Canada

L. Perrier, MLIS · A. C. Tricco, PhD  
Li Ka Shing Knowledge Institute of St Michael's Hospital,  
Toronto, ON, Canada

L. Perrier, MLIS  
Continuing Education & Professional Development, Faculty of  
Medicine, University of Toronto, Toronto, ON, Canada

W. Tavares, PhD(c)  
Wilson Centre and Centennial College, Toronto, ON, Canada

S. Reeves, PhD  
Faculty of Health, Social Care and Education, Kingston  
University and St George's, University of London, London, UK

educational programs, Kirkpatrick's hierarchy<sup>12</sup> can be used as a classification tool to communicate the level of learning outcome, and multiple levels are possible within a single study. In the original Kirkpatrick framework,<sup>12</sup> learning outcomes resulting from educational interventions in healthcare are classified into four levels:<sup>13,14</sup>

Level 1 - Reaction: measures how learners perceive the educational intervention;

Level 2 - Learning: measures acquisition of skills/knowledge/attitudes in a non-clinical setting (e.g., simulation labs);

Level 3 - Behaviour: measures learners' behavioural changes in the professional setting, i.e., transfer of learning to the clinical setting; and

Level 4 - Results: measures the effect of learners' actions, i.e., improved patient outcomes.

In our systematic review, we deliberately focused on the application of learning captured by Kirkpatrick Levels 3 (transfer of learning to the workplace) and 4 (patient outcome); therefore, we excluded studies that investigated only Kirkpatrick Level 1 and 2 outcomes that evaluate learners' reactions or learning, respectively. We aimed to include all healthcare professionals independent of their level of training or specialty. This systematic review was conducted to gain a better understanding of the impact that simulation-based CRM teaching has on transfer of learning to the workplace and on subsequent changes in patient outcomes.

## Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement was used to guide the reporting of this review.<sup>15</sup>

### Protocol

A review protocol and a search strategy following PRISMA guidelines were compiled and revised by the investigators who together have expertise in systematic review methodologies, medical education, and clinical care. They are available from the corresponding author upon request.

### Eligibility criteria

All studies included in this review met predetermined eligibility criteria. The study subjects were healthcare providers, including physicians, nurses, respiratory therapists, physician assistants, perfusionists, and paramedics. All levels of practice were included, from trainees (pre- and post-registration, undergraduate, and

postgraduate) to staff. The following study designs were included in this review: randomized controlled trials (RCTs); quasi-randomized studies (where the method of allocating participants to groups is not strictly random); controlled before-and-after studies (observations measured in both an intervention and a control group before and after the intervention); interrupted time series (ITS) (observations at multiple time points before and after an intervention in a single cohort); cohort studies (following a defined group of people over time); and case control studies (a method that compares people with a specific outcome of interest with a control group that doesn't have the specific outcome).

The intervention must include simulation-based CRM teaching. Interventions that did not explicitly mention the terms "CRM" or "crew resource management" but taught relevant non-technical skills during a medical crisis were also included. We excluded papers where we could not separate out teaching and/or assessment of technical skills from non-technical skills in an acute care context. Outcomes were assessed using a modified Kirkpatrick model of outcomes at four levels.<sup>13,16</sup> Papers were included if they measured identifiable CRM skills at Levels 3 and 4, i.e., behavioural change in the workplace or patient outcome (see above). We excluded papers measuring Kirkpatrick Levels 1 and 2 outcomes because they focus simply on learner reactions and learning measured in a simulated environment. In addition, given the abundant literature on self-assessment inaccuracy,<sup>17,18</sup> papers reporting solely self-assessment data and considered a Level 1 (reaction) outcome, were excluded.

For the purpose of this systematic review, only studies that measured outcomes in humans (either healthcare providers or patients) were included; therefore, we excluded studies that measured only simulated outcomes. Only English and French language publications were included, and only published studies were included.

### Information sources

The literature search was performed by an experienced librarian (L.P.) in close collaboration with the rest of the research team. The literature search was last performed on September 4, 2012 from MEDLINE®, EMBASE™, CINAHL, Cochrane Central Register of Controlled Trials, and ERIC.

### Literature search

Searches were performed without year or language restrictions. Search terms included: crisis resource management, crisis management, crew resource management, teamwork, and simulation. Appropriate wildcards were used in the search to account for plurals and variations in spelling. The

comprehensive search was intended to obtain: (i) all trials investigating crisis resource management with non-technical skills, soft skills, human factors, or only specific types of non-technical skills (leadership, communication, task management, decision-making, situation awareness, team work) applied to emergency/ high stakes situations independent of profession/ discipline; (ii) all trials comparing simulation-based (virtual reality, screen simulator, low-fidelity simulator, high-fidelity simulator, human simulation) education *vs* any other method of education, including traditional training, in-job training, or no training; and (iii) all trials comparing one method of simulation-based education *vs* another method of simulation-based education (e.g. comparison of two different simulators). The detailed search strategy is available in Appendix 1.

### Study selection

All titles and abstracts identified in the literature search were independently reviewed for eligibility by two pairs of authors. Disagreements were recorded and resolved by discussion. The full text articles of potentially eligible abstracts were retrieved and reviewed by two authors independently (H.Q., L.F.). Disagreements were resolved by consensus agreement under the guidance of the third author (D.B. or S.B.).

### Data collection process and data items

Using a data extraction form with inclusion and exclusion criteria, two authors (H.Q. and L.B.) extracted data from included articles. The data extraction form collected general article information, year trial was conducted, study design, sample size, description of study participants, healthcare providers involved, type of case and environment, description of the intervention, nature of the comparison group, data on the primary outcome, methodological quality, and sample-size calculation.

### Risk of bias in individual studies

Two independent reviewers (H.Q. and L.F.) assessed each included study for risk of bias using the Effective Practice and Organisation of Care Group (EPOC) tool<sup>19</sup> for RCT and ITS studies and the Newcastle-Ottawa Quality Assessment Scale<sup>20</sup> for cohort studies, as appropriate.

### Synthesis of results

A meta-analysis was not performed because of heterogeneity of study design and outcome measures; instead, a narrative summary was conducted.

## Results

### Study selection

The search yielded 7,455 publications, which resulted in 5,105 articles after the removal of duplicates. After screening the title and abstracts for the inclusion criteria, 4,646 articles were excluded, leaving 459 published articles. After review of the full text of these articles, another 450 were excluded based on the inclusion/exclusion criteria, resulting in nine articles included in this systematic review (Fig. 1).

### Study characteristics

Details on included study characteristics, participants, interventions, methods, and results are available in Tables 1, 2, and Supplementary Electronic Material (Appendices 2 and 3).

### Training characteristics

Eight studies used a combination of didactic and simulation training approaches in teaching CRM principles,<sup>21-28</sup> and one study used only simulated mock codes.<sup>29</sup>

### Evaluation of outcomes and assessment tools

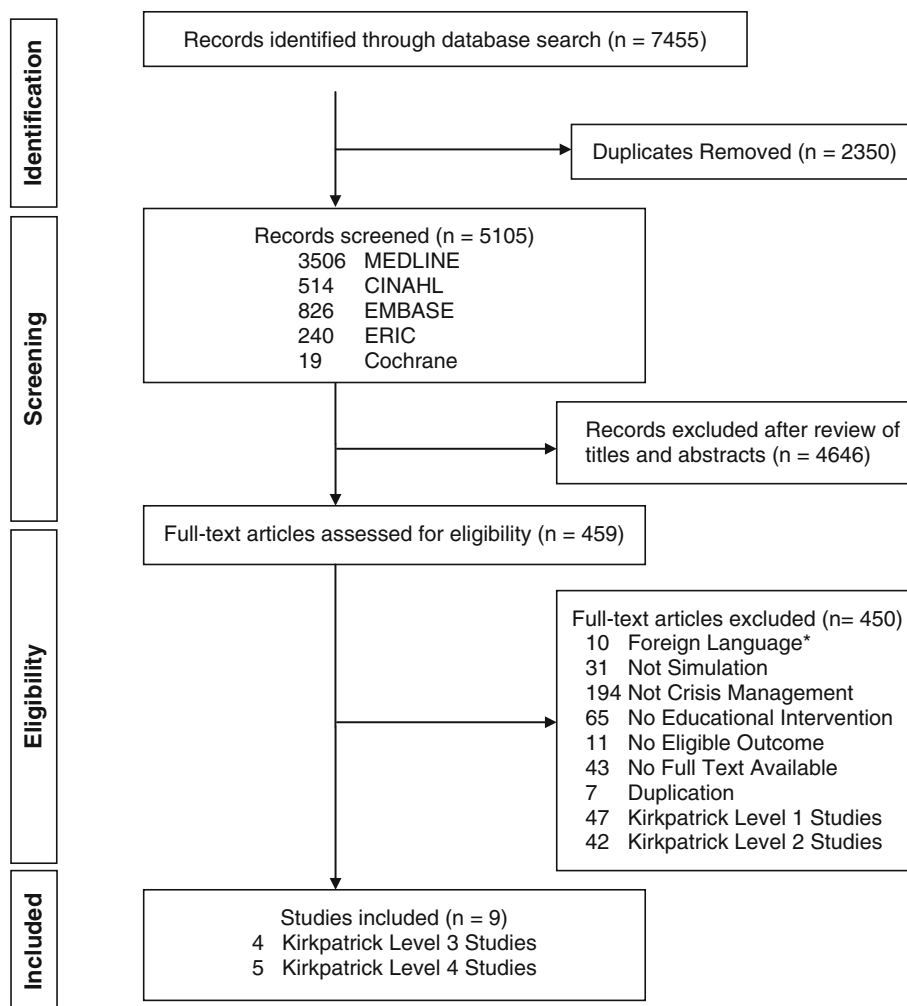
The Kirkpatrick model allows combining several levels into a single study. Two studies investigated Kirkpatrick Levels 3 and 4, with a measure of the performance of team crisis management in the workplace (Level 3) and a measure of patient outcome (Level 4).<sup>23,28</sup> These studies were considered to be both Kirkpatrick Levels 3 and 4 in our analysis; however, in the total count, they were included only in Kirkpatrick Level 4 group.

Four studies reached Kirkpatrick Level 3 at most, assessing transfer of learning to the workplace (i.e., participants' performance during real clinical context). Five studies reached Kirkpatrick Level 4 (patient outcome) at most. They considered mortality among the patients' clinical outcome data.<sup>23,24,26,28,29</sup> One study also used a patient survey, which was not included in the analysis because it was considered to be self-assessment data.<sup>26</sup> Other clinical performance scores included the Weighted Adverse Outcomes Score,<sup>24</sup> resuscitation time,<sup>23,28</sup> and length of stay.<sup>23,28</sup>

### Effects of intervention

In terms of transfer of learning to the workplace (Kirkpatrick Level 3), all included studies but one<sup>21</sup> (with  $P = 0.07$ ) found a significant effectiveness of simulation-

**Fig. 1** Search and selection of included studies. \*Languages other than English or French considered as Foreign Language



enhanced CRM training,<sup>23,28</sup> including when compared with didactic teaching alone.<sup>22,25,27</sup> Detailed results of the included studies are provided in Table 1 and Supplementary Electronic Material (Appendix 2). In terms of skill preservation, there are conflicting results among studies. In the study by Miller *et al.*, transfer of CRM skills in the workplace was not retained after a month,<sup>25</sup> while transfer was retained for at least five weeks in another study.<sup>27</sup>

In terms of patient outcomes (Kirkpatrick Level 4), all included studies found at least some improved patient outcomes after simulation CRM training,<sup>23,26,28,29</sup> including when compared with didactic teaching alone.<sup>24</sup> Surrogate measures used to approach patient outcomes can be grouped into four main categories: efficiency of patient care (time to perform), complications, length of stay, and survival/mortality. Detailed results of the included studies are provided in Table 2 and Supplementary Electronic Material (Appendix 3). Only one study found that simulation CRM training had a clearly significant impact on mortality for in-hospital pediatric cardiac arrest, where

survival rates increased from 33% to 50% within one year.<sup>29</sup> Capella *et al.*<sup>23</sup> and Steinemann *et al.*<sup>28</sup> both found an improvement in efficiency of patient care after CRM simulation training but no effect on mortality (Supplementary Electronic Material- Appendix 3). Riley *et al.*<sup>24</sup> observed a statistically significant and persistent improvement of 37% in perinatal outcome from pre- to post-intervention in the hospital exposed to the simulation program,<sup>24</sup> while there was no statistically significant change in patient outcome in the two other hospitals (didactic-only, control with no intervention), showing the benefits of simulation CRM teaching. Phipps *et al.*<sup>26</sup> found that the complication rate decreased significantly after teaching.

#### Risk of bias

Overall, the studies included in this systematic review appear to be at intermediate or high risk of bias. In addition, many items remained unclear, including random sequence generation (selection bias), allocation



**Table 1** Included studies with Kirkpatrick Level 3 outcomes (behavioural changes)

Primary author (Year)	Clinical context and participants	Study design	Results	Conclusions
Shapiro <i>et al.</i> (2004)	4 Emergency Medicine Teams (1 team = 1 MD, 1 resident, 3 RNs)	RCT – teams randomized to simulation-enhanced CRM teaching or no extra teaching. Outcomes during real trauma activations were compared	Simulation group showed a non-significant improvement in the quality of team behaviour during the two observation periods ( $P = 0.55$ )	Simulation appears to be promising to improve CRM team behaviours
Knudson <i>et al.</i> (2008)	Trauma (18 surgical residents)	RCT—subjects randomized to scenario-based didactic sessions or scenario-based, simulator-enhanced teaching. Outcomes during real trauma activations were compared	The simulation group performed better than the didactic group at behavioural skills level (increased performance by around 9%) but no difference at technical skills level	Simulation may be more effective than didactic teaching for transfer of learning of CRM skills
Bruppacher <i>et al.</i> (2010)	Anesthesia (20 post-graduate trainees)	RCT—subjects randomized to simulation teaching or didactic teaching. Outcomes during weaning from cardiopulmonary bypass in the operating room.	The simulation group scored significantly higher than the seminar group at both post-test (Global Rating Scale: $P < 0.001$ ; checklist: $P < 0.001$ ) and retention test (Global Rating Scale: $P < 0.001$ ; checklist: $P < 0.001$ )	High-fidelity simulation-based training leads to improved patient care during cardiopulmonary bypass weaning when compared with interactive seminars
Miller <i>et al.</i> (2012)	39 Trauma activations (various staff MDs, residents, RNs, technicians, pharmacists, clerks, and RTs in an ED)	ITS - teamwork behaviours were observed during real trauma activations and compared over four periods: pre-test (baseline, didactic-based phase) and post-test (simulation and decay phase)	CTS measurements for teamwork improved in 12 out of 14 domains during ISTS phase compared with baseline, whereas only one CTS improved during the didactic phase. All CTS measures returned to baseline during the decay phase	Teamwork and communication in the clinical setting may be improved during an <i>in situ</i> simulation program, but these potential benefits are lost if the simulation program is not continued
Capella <i>et al.</i> (2010)*	Trauma (114; 28 surgery residents, 6 faculty surgeons, 80 ED nurses)	Uncontrolled Before-and-After study; Pre/post training intervention study design Compared clinical outcome and efficiency of care pre and post team training	Significant improvement from pre-training to post-training in all teamwork domain ratings (leadership, $P = 0.003$ ; situation monitoring, $P = 0.009$ ; mutual support, $P = 0.004$ ; communication, $P = 0.001$ ); and overall ratings ( $P < 0.001$ ) A significant decreased time was observed after the training for: arrival to CT scanner (26.4-22.1 min, $P = 0.005$ ), endotracheal intubation (10.1-6.6 min, $P = 0.049$ ), and operating room (130.1-94.5 min, $P = 0.021$ )	Structured trauma resuscitation team training augmented by simulation improves team performance
Steinemann <i>et al.</i> (2011)*	Trauma (137; 9 staff surgeons, 21 staff ED physicians, 24 residents, 3 physician assistants, 44 RNs, 23 RTs, 13 technicians)	Uncontrolled Before-and-After study; Pre/post training intervention study design compared clinical outcome pre and post team training during actual trauma resuscitations	Significant improvements in speed (reduction by 16% of mean overall ED resuscitation time, $P < 0.05$ ) and completeness of resuscitation (76% increase in the frequency of near-perfect task completion, $P < 0.001$ ) Significant increase in CRM performance scores from the pre-to post-training periods (T-NOTECHS score 16.7 to 17.7, $P < 0.05$ )	A relatively brief (4 hr) simulation-based curriculum can improve clinical performance of multidisciplinary trauma teams

CRM = crisis resource management; CTS = Clinical Teamwork Score; ED = emergency department; ITS = interrupted time series; ISTS = *in situ* trauma simulation; MD = medical doctor; RCT = randomized controlled trial; RN = registered nurse; RT = respiratory therapist; T-NOTECHS = trauma non-technical skills

\*Study includes both Kirkpatrick Level 3 and Level 4 outcomes. Please see Table 2 for details on Kirkpatrick Level 4 outcomes

**Table 2** Included studies with Kirkpatrick level 4 outcomes (patient outcomes)

Primary author (Year)	Clinical context and participants	Study design	Results	Conclusions
Capella <i>et al.</i> (2010)*	Trauma (114; 28 surgery residents, 6 faculty surgeons, 80 ED nurses)	Uncontrolled Before-and-After study; Pre/post training intervention study design Compared clinical outcome and efficiency of care pre and post team training	A significant decrease in time was observed after the training for: arrival to CT scanner (26.4-22.1 min, $P = 0.005$ ), endotracheal intubation (10.1-6.6 minutes, $P = 0.049$ ), and operating room (130.1-94.5 min, $P = 0.021$ )  No significant difference was observed after the training for: intensive care unit LOS (5.5-6.3 days, $P = 0.445$ ), hospital LOS (7.6-6.3 days, $P = 0.210$ ), absence of complication rate (70.5-76.8, $P = 0.113$ ), and survival rate (86.9-91.5, $P = 0.121$ ) and times from arrival to FAST examination (8.3-9.6 min, $P = 0.131$ ) and time in the ED (186.1-187.4, $P = 0.93$ )	Structured trauma resuscitation team training augmented by simulation resulted in improved efficiency of patient care in the trauma bay
Steinemann <i>et al.</i> (2011)*	Trauma (137; 9 staff surgeons, 21 staff ED physicians, 24 residents, 3 physician assistants, 44 RNs, 23 RTs, 13 technicians)	Uncontrolled Before-and-After study; Pre/post training intervention study design compared clinical outcome pre and post team training during actual trauma resuscitations	Significant improvements in speed (reduction by 16% of mean overall ED resuscitation time, $P < 0.05$ ) and completeness of resuscitation (76% increase in the frequency of near-perfect task completion, $P < 0.001$ ) The mortality rate, mean ICU and hospital LOS were not significantly different before and after the training all $P > 0.05$	A relatively brief (4 hr) simulation-based curriculum can improve clinical performance and patients' outcomes
Riley <i>et al.</i> (2011)	Obstetrics and perinatal (134 from 3 hospitals; 13 obstetricians, 23 family practitioners, 14 pediatricians, 65 registered nurses, 18 certified RNs anesthetist, 1 physician assistant)	RCT – Cluster randomization of hospitals Randomized to simulation-based, didactic-based, or no intervention. Groups were compared using clinical outcome scores.	A statistically significant and persistent improvement of 37% ( $P < 0.05$ ) in perinatal morbidity was observed between the pre- and post-intervention for the hospital exposed to the simulation program. There were no statistically significant differences in the didactic-only or the control hospitals ( $P > 0.05$ ).  No significant change in the perception of culture of safety ( $P > 0.05$ ) at the three hospitals	Interdisciplinary <i>in situ</i> simulation training is effective in decreasing perinatal morbidity and mortality for perinatal emergencies Didactics alone were not effective in improving perinatal outcomes

Table 2 continued

Primary author (Year)	Clinical context and participants	Study design	Results	Conclusions
Andreatta <i>et al.</i> (2011)	Pediatrics resuscitation (228, junior and senior pediatric medicine resident with code team members: RNs, medical students, pediatric hospitalists, pharmacists)	Longitudinal cohort study for 4 years Observed patients' outcome as the training occurred over several years	After the routine integration of the formal mock code program into residency curriculum, resuscitation survival rates significantly increased from 33% to 50% within 1 year, in increments that correlated with the increasing number of mock code events ( $r = 0.87$ ) and held steady for three consecutive years	Simulation-based mock codes can provide a sustainable and transferable learning context for advanced clinical training and assessment that ultimately decreased mortality for pediatric resuscitations
Phipps <i>et al.</i> (2012)	Obstetrics and perinatal (~185; obstetricians, perinatologists, labour and delivery RNs, certified nurse midwives, anesthesiologists, certified RN anesthetists, resident physician/fellows)	ITS - Patient outcomes were assessed using data collected quarterly for 8 quarters prior to initiating the program and for the 6 quarters after implementing the program hospital wide, multidisciplinary simulation - based CRM intervention was applied to assess clinical outcome data collected 8 quarters pre-interruption and 6 quarters post interruption.	AOIs significantly decreased from 0.052 (95%CI: 0.048 to 0.055) at baseline to 0.043 (95%CI: 0.04 to 0.047). Overall, the frequency of event reporting and the overall perception of safety did not change significantly. No change in patient perception but were satisfied > 90% even before the intervention	Using the combination of a didactic and simulation-based CRM training was noted to improve patient outcomes

AOI = adverse outcome index; CT = computed tomography; CI = confidence interval; CRM = crisis resource management; ED = emergency department; ICU = intensive care unit; ITS = interrupted time series; LOS = length of stay; FAST = focused assessment with sonography for trauma; MD = medical doctor; RCT = randomized controlled trial; RN = registered nurse; RT = respiratory therapist

\*Study includes both Kirkpatrick Level 3 and Level 4 outcomes. Please see Table 1 for details on Kirkpatrick Level 3 outcomes



	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Baseline outcome measurements similar	Baseline characteristics similar	Incomplete outcome data (attrition bias)	Blinding of outcome assessment (detection bias)	Contamination	Selective reporting (reporting bias)	Other bias	intervention independent of other changes	Shape of intervention effect pre-specified	intervention unlikely to affect data collection
Bruppacher 2010	+	?	+	?	?	+	?	+	+			
Kudson 2008	?	?	?	?	?	+	?	+	-			
Miller 2012					?	-		+	-	?	+	+
Phipps 2012					-	+		+	+	?	+	+
Riley 2011	?	?	+	-	?	+	-	+	+			
Shapiro 2004	?	?	+	?	+	-	?	+	+			

**Fig. 2** Risk of bias summary. Other biases include large inter-rater reliability of 0.2 for part II outcome assessments<sup>23</sup> and sampling bias.<sup>26</sup> Green = low risk; yellow = intermediate risk or unclear; red = high risk

concealment, baseline characteristics, contamination, and intervention independent of other changes, suggesting room for improvement in the way studies are reported. Figure 2 shows a risk of bias summary for six studies using the EPOC tool,<sup>19</sup> and Table 3 presents risk of bias for three studies using the Newcastle-Ottawa Quality Assessment Scale.<sup>20</sup>

**Discussion**

Despite an abundance of existing literature on simulation-based education and CRM, we identified only nine articles that examined transfer of learning to the workplace by healthcare providers or changes in patient outcome after simulation-based CRM training. The vast majority of the literature has been limited to lower-level outcomes, such as reaction of participants and learning that has been measured using further simulation scenarios. This approach leaves the studies open to the criticism that learners may have been taught to perform well only in the simulator and not necessarily in real life.

These findings are relevant to various stakeholders such as healthcare providers, researchers, educators, policy makers, healthcare institutions, and broader organizations. Although limited in quantity and quality, the literature suggests that simulation CRM training may have a significant impact on transfer of learning to the workplace and on patient outcome.

Currently, no consensus exists on the learning outcomes unique to simulation (i.e., simulated patient outcome, simulated behaviours, etc...) and how best to assess these factors. For example, Kirkpatrick does not adequately capture studies like that of DeVita *et al.*<sup>30</sup> where the main outcome measure was survival of the simulated patient. This may be because the Kirkpatrick model for evaluating learning interventions was not originally developed for simulation education.<sup>12</sup> Although Kirkpatrick’s model is most often used to appraise the quality of educational research, we agree with Yardley and Dornan<sup>31</sup> that other frameworks may be relevant for appraising the quality of educational research. They write,<sup>31</sup> “Aggregative or interpretive methods of evidence synthesis that mix qualitative with quantitative evidence, or synthesize qualitative evidence alone, give better knowledge support and start from constructionist rather than positivist epistemological assumptions.” Medical education is pluralistic, and a positivist paradigm lens alone cannot capture its complexity. As a widely adopted framework specific to simulation education outcomes is presently lacking, it is important to recognize that the Kirkpatrick classification may not accurately capture all higher-level learning outcomes in simulation education. An ideal framework for simulation and education interventions would account for complexity of interventions, maintenance of behaviour changes, and differentiate between self- and external skill assessment and between simulated and real practice.

The data from this review provide evidence that CRM simulation training can improve behaviour at the workplace; however, whether this kind of training directly improves patient outcome is not as clear. Various measures to approach patient outcomes were used in the papers included in our review, including patient care efficiency (time to perform), complications, length of stay, and mortality. While most would agree that complications, length of stay, and mortality are appropriate criteria to assess patient outcome, it is debatable whether patient care efficiency is appropriate. Only one study found that simulation CRM training had a clearly significant impact on mortality following in-hospital pediatric cardiac arrest.<sup>29</sup> This study was simply a cohort study in a single hospital with no control group, thus results may potentially be due to other concomitant hidden interventions, and therefore, no strong definitive conclusion can be made regarding the causal relationship between the teaching intervention and mortality. Only an RCT with a control group could show that the teaching intervention is the reason for better survival. The practical requirements for designing studies that examine improvements in patient outcome can be difficult due to the need for larger sample sizes and a control group. For example, one of the studies included in

**Table 3** Risk of bias summary for cohort studies analyzed according to the Newcastle-Ottawa Quality Assessment Scale

Study		Capella (2010)	Andreatta (2011)	Steinemann (2011)
<b>Selection</b>				
1) Representativeness of the exposed cohort	a) Truly representative of the average			
	b) Somewhat representative of the average	*		*
	c) Selected group of users		*	
	d) No description of the cohort			
2) Selection of the non exposed cohort	a) Drawn from the same community as the exposed cohort			
	b) Drawn from a different source			
	c) No description of the non exposed cohort			
3) Ascertainment of exposure	a) Secure record	*	*	*
	b) Structured interview			
	c) Written self report			
	d) No description			
4) Demonstration that outcome of interest was not present at start of study	a) Yes	*	*	*
	b) No			
<b>Comparability</b>				
1) Comparability of cohorts on the basis of the design or analysis	a) Study controls for design or analysis			
	b) Study controls for any additional factor			
<b>Outcome</b>				
1) Assessment of outcome	a) Independent blind assessment			
	b) Record linkage	*	*	*
	c) Self report			
	d) No description			
2) Was follow-up long enough for outcomes to occur	a) Yes	*	*	*
	b) No			
3) Adequacy of follow up of cohorts	a) Complete follow up			
	b) Subjects lost to follow up unlikely to introduce bias	*		*
	c) Follow up rate < % and no description of those lost			
	d) No statement		*	
Summary		Selection ***	Selection ***	Selection ***
		Comparability	Comparability	Comparability
		Outcome ***	Outcome ***	Outcome ***

\* = satisfied, blank = unsatisfied

this review did not have a sample-size calculation, and this likely resulted in an underpowered study. All of the studies included in this review involved one time-limited intervention on a small number of subjects.<sup>21</sup> It is possible that modification of patient outcome requires a whole series of interventions on many subjects. Finally, although CRM programs without simulation teaching have

been linked to decreased surgical mortality,<sup>32</sup> we could not find a multicentre RCT that evaluated simulation CRM training on patient outcome. Nevertheless, if we compare with other high-stake industries, like aviation, despite several studies showing an improvement in pilots' behaviour in the cockpit, studies showing the benefit of CRM pilot training on client safety are lacking.<sup>33</sup>

Another potential reason for our small sample size of studies may be the conservative nature of our inclusion criteria. The decision to include objectively measured change of behaviour at the workplace and to exclude self-assessment (Kirkpatrick Levels 1 and 2 – reaction and knowledge and skills learning, respectively) may have limited our analysis; however, self-assessment is largely recognized as inaccurate for healthcare professionals.<sup>18,34</sup> The initial literature search was performed without any language restriction. Nevertheless, we included studies published in English or French only. Of course, we cannot ignore that a few papers were excluded because they were published in other languages. Given that the vast majority of scientific journals are published in English and all high-impact factor journals are in English, in our view, it is unlikely that the conclusions of our review would be significantly different if more languages had been included.

Overall, we found that the studies were at an intermediate or high risk of bias and reporting was suboptimal. First, there is clearly room for improvement in the approach used to report studies. For example, random sequence generation and allocation concealment were almost never reported properly in the included studies. We cannot determine if the studies were performed incorrectly or if “only” the reporting was poor. Second, it may be challenging to design studies on simulation-based CRM without risk of bias when investigating transfer of learning to the workplace and patient outcome. For example, when working in increasingly complex organizations, it is very difficult to ensure that risk of contamination is nonexistent and intervention is independent of other changes. We suggest that, as a field, the simulation community needs to commit to rigorous research reports. Also, larger and multicentre studies could balance the risk of contamination. In order to decrease the risk of bias as much as possible in future studies, we also suggest that researchers consider the risk of bias at an early stage when designing the protocol.

Moving forward, larger sample sizes, more multicentre studies, and studies with less risk of bias are required to provide a precise measure of the effect that simulation-based education has on healthcare provider skills in the workplace and patient outcome. Other systematic reviews show that there is no need for more Kirkpatrick Level 1 (reaction) and Level 2 (learning) studies, since learners are virtually constantly positive toward simulation training<sup>10,11</sup> and learning occurs when measured in a simulated environment.<sup>1</sup> Frequency of retraining, skill retention, and instructional design remain research priorities in studies investigating Kirkpatrick Level 3 (transfer of learning at the workplace) and Level 4 (patient outcome) outcomes. Universally recognized rigorous assessment tools are necessary to compare the effect of various

teaching interventions and to assess CRM regardless of the clinical context. Finally, simulation training is often underused, potentially due to its cost. Future research could better explore the cost-effectiveness of simulation CRM training.

## Conclusions

A limited number of studies have examined the true impact of simulation-based CRM training on Kirkpatrick Level 3 (transfer of learning at the workplace) and Level 4 (patient outcome) outcomes. Based on the nine studies included, this systematic review illustrates that CRM skills acquired at the simulation centre are transferred to clinical settings and lead to improved patient outcomes. Given these findings, we suggest the need for an internationally recognized interprofessional simulation-based CRM training certification for healthcare professionals that would teach CRM independently of the clinical context. Findings from this review may help guide future research in CRM simulation-based education.

**Acknowledgements** Andrea C. Tricco is funded by a Canadian Institutes for Health Research (CIHR)/Drug Safety and Effectiveness Network New Investigator Award in Knowledge Synthesis Methodology. The authors thank Laura Bekes for her contribution to data extraction and Ashlee-Ann Pigford for her constructive review of the manuscript.

**Financial support** Supported by the Department of Anesthesiology of The Ottawa Hospital, University of Ottawa (Ottawa, ON, Canada).

**Conflicts of interest** None declared.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## Appendix 1 Search strategy

Database: Ovid MEDLINE® <1950 to September Week 1 2012>, Ovid MEDLINE® In-Process & Other Non-Indexed Citations <September 4, 2012>

Search Strategy:

- 1 simulat\$.mp.
- 2 “crisis resource management”.tw.
- 3 (crisis adj management).tw.
- 4 “crew resource management”.tw.
- 5 CRM.tw.
- 6 team\$.mp.

7 \*Clinical Competence/  
 8 or/2-7 (126860)  
 9 1 and 8 (2817)  
 10 Animals/ not (Animals/ and Humans/)  
 11 9 not 10

## References

- Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 2011; 306: 978-88.
- Gurusamy KS, Aggarwal R, Palanivelu L, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev* 2009; 1: CD006575.
- Haycock A, Koch AD, Familiari P, et al. Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. *Gastrointest Endosc* 2010; 71: 298-307.
- Boet S, Bould MD, Sharma B, et al. Within-team debriefing versus instructor-led debriefing for simulation-based education: a randomized controlled trial. *Ann Surg* 2013; 258: 53-8.
- Savoldelli GL, Naik VN, Park J, Joo HS, Chow R, Hamstra SJ. Value of debriefing during simulated crisis management: oral versus video-assisted oral feedback. *Anesthesiology* 2006; 105: 279-85.
- Boet S, Bould MD, Bruppacher HR, Desjardins F, Chandra DB, Naik VN. Looking in the mirror: self-debriefing versus instructor debriefing for simulated crises. *Crit Care Med* 2011; 39: 1377-81.
- Boet S, Bould MD, Schaeffer R, et al. Learning fibreoptic intubation with a virtual computer program transfers to 'hands on' improvement. *Eur J Anaesthesiol* 2010; 27: 31-5.
- Borges BC, Boet S, Siu LW, et al. Incomplete adherence to the ASA difficult airway algorithm is unchanged after a high-fidelity simulation session. *Can J Anesth* 2010; 57: 644-9.
- Gaba DM, Howard SK, Fish KJ, Smith BE, Sowb YA. Simulation-Based Training in Anesthesia Crisis Resource Management (ACRM): A Decade of Experience. *Simulation Gaming* 2001; 32: 175-93.
- Gordon M, Darbyshire D, Baker P. Non-technical skills training to enhance patient safety: a systematic review. *Med Educ* 2012; 46: 1042-54.
- Doumouras AG, Keshet I, Nathens AB, Ahmed N, Hicks CM. A crisis of faith? A review of simulation in teaching team-based, crisis management skills to surgical trainees. *J Surg Educ* 2012; 69: 274-81.
- Kirkpatrick DL, Kirkpatrick JD. *Evaluating Training Programs: the Four Levels*. 3rd ed. San Francisco, CA: Berrett-Koehler; 2006 .
- Boet S, Sharma S, Goldman J, Reeves S. Review article: Medical education research: an overview of methods. *Can J Anesth* 2012; 59: 159-70.
- Hammick M, Freeth D, Koppel I, Reeves S, Barr H. A best evidence systematic review of interprofessional education: BEME Guide no. 9. *Med Teach* 2007; 29: 735-51.
- Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; 151: 264-9.
- Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005; 27: 10-28.
- Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. *Adv Health Sci Educ Theory Pract* 2002; 7: 63-80.
- Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006; 296: 1094-102.
- EPOC. Risk of Bias Criteria. EPOC (Effective Practice and Organisation of Care Group). Available from URL: <http://epoc.cochrane.org/> (accessed January 2014).
- Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute. Available from URL: [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp) (accessed January 2014).
- Shapiro MJ, Morey JC, Small SD. Simulation based teamwork training for emergency department staff: does it improve clinical team performance when added to an existing didactic teamwork curriculum? *Qual Saf in Health Care* 2004; 13: 417-21.
- Knudson MM, Khaw L, Bullard MK, et al. Trauma training in simulation: translating skills from SIM time to real time. *J Trauma* 2008; 64: 255-64.
- Capella J, Smith S, Philp A, et al. Teamwork training improves the clinical care of trauma patients. *J Surg Educ* 2010; 67: 439-43.
- Riley W, Davis S, Miller K, Hansen H, Sainfort F, Sweet R. Didactic and simulation nontechnical skills team training to improve perinatal patient outcomes in a community hospital. *Jt Comm J Qual Patient Saf* 2011; 37: 357-64.
- Miller D, Crandall C, Washington C 3rd, McLaughlin S. Improving teamwork and communication in trauma care through in situ simulations. *Acad Emerg Med* 2012; 19: 608-12.
- Phipps MG, Lindquist DG, McConaughy E, O'Brien JA, Raker CA, Paglia MJ. Outcomes from a labor and delivery team training program with simulation component. *Am J Obstet Gynecol* 2012; 206: 3-9.
- Bruppacher HR, Alam SK, LeBlanc VR, et al. Simulation-based training improves physicians' performance in patient care in high-stakes clinical setting of cardiac surgery. *Anesthesiology* 2010; 112: 985-92.
- Steinemann S, Berg B, Skinner A, et al. In situ, multidisciplinary, simulation-based teamwork training improves early trauma care. *J Surg Educ* 2011; 68: 472-7.
- Andreatta P, Saxton E, Thompson M, Annich G. Simulation-based mock codes significantly correlate with improved pediatric patient cardiopulmonary arrest survival rates. *Pediatr Crit Care Med* 2011; 12: 33-8.
- 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.
- Yardley S, Dorman T. Kirkpatrick's levels and education 'evidence'. *Med Educ* 2012; 46: 97-106.
- Neily J, Mills PD, Young-Xu Y, et al. Association between implementation of a medical team training program and surgical mortality. *JAMA* 2010; 304: 1693-700.
- Salas E, Burke CS, Bowers CA, Wilson KA. Team training in the skies: does crew resource management (CRM) training work? *Hum Factors* 2001; 43: 641-74.
- Eva KW, Cunningham JP, Reiter HI, Keane DR, Norman GR. How can I know what I don't know? Poor self assessment in a well-defined domain. *Adv Health Sci Educ Theory Pract* 2004; 9: 211-24.