Clinical Simulation in Nursing (2017) 13, 347-358

Clinical Simulation in Nursing

www.elsevier.com/locate/ecsn

Review Article

Development of Crisis Resource Management Skills: A Literature Review

Amanda Lucas, RN, BN[a,*], **Marie Edwards, RN, PhD**[b]

[a]*Health Sciences Centre, GE 648, 820 Sherbrook St., Winnipeg, Manitoba R3A 1R9, Canada*
[b]*College of Nursing, Rady Faculty of Health Sciences, Helen Glass Centre for Nursing, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada*

KEYWORDS
crisis resource management;
high-fidelity simulation;
literature review

Abstract: The use of crisis resource management principles (CRM), including problem solving, situational awareness, resource utilization, communication, and leadership, have been thought to reduce adverse patient outcomes and lead to greater teamwork in healthcare settings. Education programs using high-fidelity simulation (HFS) has become an increasingly popular strategy to teach these skills. There is little evidence however demonstrating the effectiveness of this type of education on actual performance of these skills. In order to explore the effectiveness of HFS education on development of CRM skills, a literature review was undertaken to identify evidence available in the healthcare literature. Thirty-one articles were identified that met criteria for this review. Articles were highly variable in methods, population used, educational intervention, evaluative method, and results. The following paper outlines a summary of these results, including synthesis of findings and recommendations for research in this area.

Cite this article:
Lucas, A., & Edwards, M. (2017, August). Development of crisis resource management skills: A literature review. *Clinical Simulation in Nursing, 13*(8), 347-358. http://dx.doi.org/10.1016/j.ecns.2017.04.006.

Although numbers and percentages vary, a vast number of patients die from what are perceived as preventable errors (Rudy, Polomano, Murray, Henry, & Marine, 2007). To avoid these errors, health care professionals must have knowledge of threats to patient safety and experience in caring for patients when extraordinary clinical problems arise. Recently, a focus in application of crisis resource management (CRM) principles has been thought to reduce adverse patient outcomes and lead to greater teamwork in crisis events (Messmer, 2008).

CRM is a set of principles that encompass a range of cognitive and interpersonal skills aimed at creating an environment of improved efficiency, teamwork, and safety (Gaba, 2010; White, 2012). Key CRM skills include (a) problem solving, (b) situational awareness, (c) resource utilization, (d) communication, and (e) leadership. Education programs that focus on these CRM skills have been shown to have a positive impact on learner competence in handling crisis events (Kim, Neilipovitz, Cardinal, Chiu, & Clinch, 2006). By teaching health care providers CRM foundational skills, the cognitive and interpersonal skills that allow them to critically analyze and respond in crisis situations will be developed as well (Andersen, Jensen,

* Corresponding author: alucas@hsc.mb.ca (A. Lucas).

1876-1399/$ - see front matter © 2017 International Nursing Association for Clinical Simulation and Learning. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ecns.2017.04.006

Lippert, & Ostergaard, 2010; Bearman et al., 2012; Gordon, Mendenhall, & Blair O'Connor, 2013; Pearson & McLafferty, 2011; White, 2012). So, how can experience with CRM skills be created? How can professionals practice responses with deteriorating patients without jeopardizing safety? In recent years, growing numbers of health care professions have turned to simulation as a way to answer these questions.

Key Points

- Simulation is a popular way of teaching crew resource management skills.
- Mixed evidence exists proving efficacy of simulation programs teaching crew resource management skills.
- A literature review of 31 relevant articles yielded variable results.

High-fidelity simulation (HFS) has become a popular method of teaching CRM skills. One form of HFS uses advanced human—patient simulators, which are computerized manikins that "can mimic diverse parameters of human physiology, such as changes in cardiovascular, pulmonary, metabolic, and neurological systems" (Lapkin, Levett-Jones, Bellchambers, & Fernandez, 2010, p. e 209).

Gaba (2010) identified that "the most frequent question now asked about CRM and teamwork training in health care, and especially for that using simulation, is 'where is the evidence'?" (p. 4). A literature review was undertaken to identify available evidence on the effectiveness of HFS education interventions on development of CRM skills, meaning, have learners translated behaviours learned in the simulation environment into performance of CRM skills? This was addressed via the following question: what evidence is there to demonstrate the effectiveness of HFS learning programs on health care professionals' acquisition and performance of CRM skills?

A summary of this evidence, including synthesis of findings and recommendations for research in this area, will be outlined.

Review of the Literature

Search Strategy

The search included studies carried out with participants from health care disciplines. All types of research, including qualitative and quantitative studies, were included if the following criteria were met.

a) There was use of HFS learning programs in the study.
b) The study included an intervention affecting performance of a CRM skill.
c) Outcomes measured included one of the CRM skills: problem solving, situational awareness, resource utilization, communication, and/or leadership.
d) Studies were published in English, between 2003 and 2016, and available electronically.

Articles were also excluded if they were descriptive, opinion papers, or commentaries or if results were only available as abstracts.

Search Outcomes

An original search of citations available from 2003 to 2014 yielded 225 papers from PubMed, CINAHL, and Scopus using the following search strategy:

> (MH "Computer Simulation")OR(MH "Simulations") OR(MH "Patient Simulation")AND (MH "Problem Solving+")OR(MH "Decision Making+")OR(MH "Decision Making, Organizational")OR(MH "Leadership")OR(MH "Communication+")OR(MH "Critical Thinking")AND(MH "Education, Nursing+")

There were 30 papers that were not available for review through the University of Manitoba library access system or Google, including 17 unpublished theses. Remaining abstracts were then read, and 81 papers were excluded where inclusion criteria were not met. The remaining papers were then read in full, examined, accepted, or rejected. Articles were excluded if the educational forum used did not include HFS as defined above (e.g., patient actors, task trainers, virtual reality); CRM educational outcomes were not evaluated; they were descriptive, opinion papers, or commentaries; results were only reported in conference abstracts; or they were not available through electronic search mechanisms. This initial search yielded 20 papers for review.

The search was repeated to include citations from 2013 to 2016. One year of overlap was done (2013-2014) to ensure databases searched yielded late submissions from the final year of the initial search. Using the same criteria and databases, 594 citations were produced, with an additional eleven studies added to this review, for a total 31.

Characteristics of Studies

The 31 studies retrieved were examined using the following categories: methods, sample, educational intervention, assessment measures, and results. Each of these categories will be discussed in some detail, and summaries of this information are presented in Table.

Methods

Seven studies had an experimental design, including two randomized, controlled, and blinded studies (Morgan, Kurrek, Bertram, LeBlanc, & Przybyszewski, 2011; Ten Eyck, Tews, Ballester, & Hamilton, 2010), one two group by two times mixed model design (Sullivan-Mann,

Table Summary of Findings and Results

First Author, Year	Methods	Sample	Educational Intervention	Assessment Measures	CRM Skill and Major Findings	Limitations
Brannan, 2008	Quasi-experimental, single-sample pre—post test	Nursing only (student), N = 107	HFS vs. alternate	Adapted validated tool (AMIQ)	PS (I)—higher posttest scores with intervention group compared with control group	Lack of randomization to groups
Brown, 2009	Experimental, comparative correlational	Nursing only (student), N = 140	HFS vs. alternate	Validated (ECG SimTest)	PS (NI)—no significant differences in critical thinking between measures	Lack of control for previous experience, small sample size, unequal treatment of groups
Buckley, 2011	Nonexperimental	Nursing only (practicing), N = 38	Pretest/posttest with HFS intervention	Unvalidated (checklist or questionnaire)	PS (I)—improved critical abilities with airway management RU (I)—77% of reports, participants were able to coordinate the efforts of responders C (I)—improved handover communication	Performance not directly evaluated difficult to differentiate cause of improvement
Bultas, 2014	Quasi-experimental pretest/posttest with control group	Nursing only (practicing), n = 33	HFS vs. static manikin	Validated (PEARS written exam and MHPT), unvalidated (PEARS BCMT)	PS (M)—HFS group BCMT scores significantly higher than control at posttest SA (I)—experimental group better at recognizing and responding in one scenario but not another	Unvalidated tool, small sample, attrition rate high, lack of blinding, scoring teams versus individuals, same scenarios used at follow-up increasing chance participants memorized responses
Dadiz, 2013	Quasi-experimental, prospective observational	Multidisciplinary (practicing), N = 228	Pretest/posttest with HFS intervention	Validated (checklist or questionnaire)	C (I)—checklist scores improved L (I)—performance of the team leader directing team efforts and decision making improved after training	Reviewers not blinded to time and order of simulations, sample not randomized to groups, inappropriate checklist tool, potential contamination of sample from other education initiatives
DeVita, 2005	Quality improvement	Multidisciplinary (practicing), N = 138	Pretest/posttest with HFS intervention	Unvalidated (time-to-task and measure of simulator	PS (I), RU (I), L (I)—simulators "survived" from 0% of the time to 89% of the time, time-to-task	No interrater reliability of performance ratings, scenarios different

(*continued on next page*)

Table (*continued*)

First Author, Year	Methods	Sample	Educational Intervention	Assessment Measures	CRM Skill and Major Findings	Limitations
				"survival")	measures improved posttraining intervention	between groups
Gilfoyle, 2007	Quasi-experimental, within-sample pretest with staged posttest	Physician only (paediatric residents), N = 15	Pretest/posttest with HFS intervention	Unvalidated (checklist or questionnaire)	C (M)—no learning of occurred during maintenance period, but no decay of learning occurred either L(I)—increase with checklist leadership scores	Unvalidated scenarios and checklist tool; possible "training effect" related to use of same scenario at six-month session
Goodstone, 2013	Quasi-experimental, two-group pretest/posttest	Nursing only (students), N = 42, n = 41	Pretest/posttest with HFS vs. case study	Validated (HSRT)	PS (M)—both groups scored significantly higher on posttest, indicating both style of simulation education equally effective	Lack of randomization, small sample, lack of no-treatment control group, potential testing effect
Hall, 2015	Quasi-experimental, retrospective, comparative	Nursing-only (students), N = 279	HFS with traditional experience vs. traditional	Validated (ATI content mastery series test)	PS (I)—significantly higher scores on posttest with experimental group compared with control	Limited generalizability, differences in sample
Huseman, 2012	Quasi-experimental, single-sample pretest/posttest	Multidisciplinary (direct care providers), N = 178	Pretest/posttest with HFS intervention	Unvalidated (chart review)	PS (M)—statistically significant improvement in some area, but not in other areas, also not sustained over time	Not addressed
Jankouskas, 2007	Nonexperimental	Multidisciplinary (practicing), N = 40	Theory with HFS with pretest/posttest	Validated (ANTS)	SA (NI)—no statistically significant increase RU (NI)—no significant difference noted in the task-management domain C (I)—statistically significant increase L (I)—significant improvements	Element of situation awareness was difficult to visualize on videotaped simulations, making it difficult to evaluate
Jankouskas, 2011	Experimental, randomized, controlled, pretest/posttest experimental	Multidisciplinary (students), N = 96	HFS vs. alternate	Validated (ANTS)	SA (I)—significant improvements in situational awareness RU (I)—significant increase in task management and team-working measures C (I)—a statistically significant improvement L (I)—significant improvement	Limited generalizability; diffusion of treatment over time, history effect, potential bias from unblinded principle investigator, use of inappropriate tool
Johnson,	Quasi-experimental,	Nursing only	HFS vs. Web-	Unvalidated	PS (I), RU (I)—manikin group	Small sample, varied

(*continued on next page*)

Table (*continued*)

First Author, Year	Methods	Sample	Educational Intervention	Assessment Measures	CRM Skill and Major Findings	Limitations
2014	two-group pretest/posttest	(practicing), N = 32	based education	(performance checklist)	scored significantly higher on observed performance mean scores	previous experience
Kesten, 2015	Repeated measures pilot study	Nursing only (practicing), N = unknown	Repeated HFS intervention with evaluation over time	Validated (APRN EVAL tool)	PS (I), SA (I), C (I), L (I)—statistically significant improvement in scores over time, continued improvement from time 3-4 in leadership domain	Small sample, lack of standardization of tools for population, differences with HFS scenarios, previous experience affecting CRM learning
Lasater, 2005	Mixed methods	Nursing only (student nurses), N = 48	Theory course with complementary HFS sessions	Unvalidated (LCJSR, LCJPS) with validated (CCTDI)	PS (I)—statistically significant outcomes in critical thinking	Involvement of a single centre, lack of comparable baseline data, and lack of sample variation
Lavigne Fadale, 2014	Quasi-experimental pretest/posttest design	Critical care or emergency room nurses, N = 16	Repeated intervention with repeated testing over time	Unvalidated (time to task, number of interventions [vasopressor titrations])	PS (M)—trend toward significance for most points with only statistically significant improvement in number of vasopressor titrations	Small sample, potential for performance bias, equipment malfunction
Liaw, 2011	Experimental, prospective, randomized, control trial with a pretest/ posttest design	Nursing only (students), N = 31	Pretest/posttest with HFS intervention vs. control without HFS intervention	Unvalidated (RAPIDS tool)	PS (I), C (I)—significant improvement over control group	Not generalizable, unknown long-term retention of skills, single-scenario exposure, lack of applicability to actual clinical setting
Maneval, 2012	Experimental, randomized, controlled, pretest/posttest experimental	Nursing only (novice nurse) N = 26	HFS vs. alternate	Validated (HSRT)	PS (NI)—nonstatistically significant improvements in mean posttest scores	Small sample size, higher than national average pretest scores, and a lack of management support affecting participation
Meurling, 2013	Quasi-experimental exploratory	Medical students, N = 54	3 HFS scenarios with evaluation over time	Unvalidated (a TEAM programme, i.e., mix of time-to-task, frequency of behaviours)	PS (NI), SA (NI), RU (NI), C (I), L (NI)—clinical performance improved modestly with only the frequency of “sum-ups” showing statistically significant improvement	Sample not generalizable

(*continued on next page*)

Table (*continued*)

First Author, Year	Methods	Sample	Educational Intervention	Assessment Measures	CRM Skill and Major Findings	Limitations
Morgan, 2011	Experimental, randomized, controlled, blinded	Physician only (anaesthesiologists), N = 59	HFS alone vs. HFS with debrief	Validated (ANTS)	PS (NI)—no improvement in nontechnical skill performance SA (I)—significant improvement RU (NI)—marginal, not statistically significant, improvements in team working and task-management domains C (NI)—not significant L (NI)—no improvement in leadership	High functional ability of sample prior to intervention, impeding ability for learning activity to affect performance, lengthy interval between simulations may have limited participants' ability to retain learning, question of inappropriate evaluation tool (ANTS)
Przybyl, 2015	Quality improvement, pretest/posttest	Nursing only (practicing), N = 93	CRRT course with addition of HFS education	Unvalidated (knowledge questionnaire)	PS (no comment on significance)—increase in understanding of CRRT principles and critical thinking related to operation of CRRT machine	Challenges with recruitment
Schubert, 2012	Quasi-experimental, within-sample pretest with staged posttest	Nursing only (practicing), N = 58	Pretest/posttest with HFS intervention	Validated (multiple choice) with LTT	PS (I)—statistically significant improvement in problem-solving skills	Untested tool, small sample, high attrition rate, participants working previous night shift may have affected performance
Shapiro, 2004	Quasi-experimental, prospective observational	Multidisciplinary (practicing), N = 20	HFS vs. alternate	Validated (team dimensions rating form)	RU (NI), C (NI), L (NI)—nonstatistically significant increase in scores between pre—post training measures in both experimental and comparison groups; these results suggest a positive impact on performance	Small sample size
Shinnick, 2013	Quasi-experimental, one-group pretest/posttest	Prelicensure nursing students, N = 154	Pretest/posttest with HFS intervention	Validated (HSRT)	PS (NI)—gains in knowledge without statistically significant gains in critical thinking	Previous experience with scenario content and simulation environment
Singer, 2013	Quasi-experimental, prospective cohort	Physician only (students), N = 67	HFS vs. alternate	Adapted validated checklist	PS (I)—first-year medical residents outperformed	Single-centre study with limited numbers,

(*continued on next page*)

Table (*continued*)

First Author, Year	Methods	Sample	Educational Intervention	Assessment Measures	CRM Skill and Major Findings	Limitations
					third-year residents after completing an HFS intervention	not blinded and limited number of competencies assessed
Sittner, 2009	Quasi-experimental, within-sample pretest with staged posttest	Nursing only (practicing), N = 11	Pretest/posttest with HFS intervention	Validated (checklist or questionnaire)	PS (NI)—no significant statistical difference over time in knowledge and clinical judgment	Small sample with no control group, participants previous experience potentially affecting pretest scores
Straka, 2012	Quasi-experimental, single-sample pretest/posttest	Nursing only (novice nurse), N = 26	Pretest/posttest with HFS intervention	Unvalidated (checklist or questionnaire)	PS (I)—significant improvement in problem-solving skills	Time frame, small sample, lack of standardized tool
Sullivan-Mann, 2009	Experimental, two group by two times mixed model	Nursing only (student nurse), N = 53	3 HFS sessions vs. 5 HFS sessions	Validated (HSRT)	PS (I)—statistically significant improvement in problem-solving skills	Small sample, differences in instruction between groups, lack of no intervention control, may not be generalizable
Ten Eyck, 2010	Experimental, randomized, controlled, blinded	Physician only (students), n = 68	HFS vs. alternate	Unvalidated (time to task)	PS (M)—statistically significant improvement in times in 4 of 8 critical tasks L (I)—showed more significant improvement as the program progressed	Lack of generalizability to "real" setting, other institutions, and different situations; lack of validated instrument; single evaluator of cases
Wolf, 2008	Quasi-experimental, single-sample pretest/posttest	Nursing only (practicing), n = 6	Pretest/posttest with HFS intervention	Unvalidated (chart review)	PS (I)—improvement from 40% accurate triage to 70%-100% postsimulation intervention	Not addressed
Wunder, 2016	Quasi-experimental, pretest/posttest	Nursing only (students), n = 32	Repeated HFS with repeated testing and CRM lecture	Validated (ANTS)	PS (I), SA (I), RU (I), C (I), L (I)—increase in total score overall, breakdown not given	Lack of interrater reliability and rater familiarity with students

Note. ANTS = Anaesthetists' Nontechnical Skills; ATI = Assessment Technologies Institute; C = communication; CRM = crew resource management; CRRT = Continuous renal replacement therapy; ECG = electrocardiogram; HFS = high-fidelity simulation; I = statistically significant improvement; L = leadership; M = mixed results; NI = no improvement, nonstatistically significant improvement; PEARS = Paediatric Emergency Assessment, Recognitions, and Stabilization; PS = problem solving; RAPIDS = Rescuing a Patient in Deteriorating Situations; RU = resource utilization; SA = situational awareness.

Perron, & Fellner, 2009), three randomized, controlled, pretest/posttest experimental studies (Jankouskas, Haidet, Hupcey, Kolanowski, & Murray, 2011; Liaw, Rethans, Scherpbier, & Piyanee, 2011; Maneval et al., 2012), and one comparative correlational study (Brown & Chronister, 2009).

Nineteen studies used a quasi-experimental design including two prospective observational (Dadiz et al., 2013; Shapiro et al., 2004), an exploratory (Meurling, Hedman, Fellander-Tsai, & Wallin, 2013), a prospective cohort (Singer et al., 2013), ten pretest/posttest (Brannan, White, & Bezanson, 2008; Gilfoyle, Gottesman, & Razack, 2007; Huseman, 2012; Lavigne Fadale, Tucker, Dungan, & Sabol, 2014; Schubert, 2012; Shinnick & Woo, 2013; Sittner, Schmaderer, Zimmerman, Hertzog, & George, 2009; Straka, Burkett, Capan, & Eswein, 2012; Wolf, 2008; Wunder, 2016), and four two-group comparative studies (Bultas, Hassler, Ercole, & Rea, 2014; Goodstone, Goodstone, Glaser, Kupferman, & Dember-Neal, 2013; Hall, 2015; Johnson et al., 2014). There were also two nonexperimental studies (Buckley & Gordon, 2011; Jankouskas, Chaska Bush, Murray, Rudy, & Henry, 2007), a mixed methods dissertation (Lasater, 2005), a repeated measures pilot (Kesten, Brown, & Meeker, 2015), and two quality improvement studies (DeVita, Schaefer, Wang, & Dongilli, 2005; Przybyl, Androwich, & Evans, 2015).

Sample

All the studies used convenience sampling, with most researchers recruiting from continuing education or university/college courses. Participants from all studies fell into three categories.

Physician Only

Five studies used participants belonging to the medical profession. Morgan et al. (2011) studied anaesthesiologists (N = 59), whereas Gilfoyle et al. (2007) used paediatric residents (N = 15). Meurling et al. (2013), Singer et al. (2013), and Ten Eyck et al. (2010) studied medical students (N = 54, N = 67, and n = 68, respectively). Demographics were offered in the Morgan et al. (2011) and the Singer et al. (2013) studies, but samples were not comparable (i.e., practicing anaesthesiologists and medical students).

Nursing Only

Practicing nurses were featured in eight studies: Sittner et al. (2009) (N = 11), Schubert (2012) (N = 58), Buckley and Gordon (2011) (N = 38), Johnson et al. (2014) (N = 32), Bultas et al. (2014) (n = 33), Lavigne Fadale et al. (2014) (N = 16), Przybyl et al. (2015) (N = 93), and Wolf (2008) (n = 6). Graduate nurses or novice nurses were studied by Straka et al. (2012) (N = 26) and Maneval et al. (2012) (N = 26). Ten studies used student nurses: Brannan et al. (2008) (N = 107), Brown and Chronister (2009) (N = 140), Sullivan-Mann et al. (2009) (N = 53), Hall (2015) (N = 279), Wunder (2016) (n = 32), Shinnick and Woo (2013) (N = 154), Kesten et al. (2015) (N = unknown), Liaw et al. (2011) (N = 31), Goodstone et al. (2013) (n = 41), and Lasater (2005) (N = 48). Demographics of both practicing and student nurses reflected a female majority with differing ages and levels of experience.

Multidisciplinary

The remaining six studies used multidisciplinary samples. Three of these consisted of practicing nurses and physicians: Dadiz et al. (2013), Jankouskas et al. (2007), and Shapiro et al. (2004) with sample sizes of 228 (total sample varying between 56% and 70% physician participation and 30% to 64% nursing participation throughout a three-year time frame), 40 (50% physicians and 50% nurses), and 20 (40% physicians and 60% nurses), respectively. Jankouskas et al. (2011) studied student nurses (n = 50) and physicians (n = 46) (total sample N = 96). Huseman (2012) studied registered nurses (n = 112), nurses' aides/nursing assistants (n = 66), respiratory therapists (n not provided), student nurses (n not provided), and pharmacists (n not provided). DeVita et al. (2005) studied registered nurses (n = 69), physicians (n = 48), and respiratory therapists (n = 21). Demographics in this category showed a higher mean age among those already practicing compared with student samples, regardless of profession.

Educational Intervention

Eleven studies compared an HFS intervention with alternate activities. Singer et al. (2013), Hall (2015), Maneval et al. (2012), and Jankouskas et al. (2011) compared standard education with standard education with the addition of HFS learning. Brannan et al. (2008), Shapiro et al. (2004), Brown and Chronister (2009), Bultas et al. (2014), Johnson et al. (2014), Goodstone et al. (2013), and Ten Eyck et al. (2010) studied outcomes of HFS learning compared with an alternate form of education such as low-fidelity simulations.

Another 13 studies compared results of a pre—post test after an HFS learning program (Buckley & Gordon, 2011; Dadiz et al., 2013; DeVita et al., 2005; Gilfoyle et al., 2007; Huseman, 2012; Liaw et al. (2011); Przybyl et al. (2015); Schubert, 2012; Shinnick and Woo (2013); Sittner et al., 2009; Straka et al., 2012; Wolf, 2008; Wunder (2016). One study (Morgan et al., 2011) looked at outcomes of simulation alone versus simulation with a guided debrief. Sullivan-Mann et al. (2009) studied results of five simulation sessions versus three. Lasater (2005) used a theory course with complementary simulation activities and ongoing evaluation. Kesten et al. (2015), Meurling et al. (2013), and Lavigne Fadale et al. (2014) studied effects of multiple simulation interventions over time. The final study paired a theory portion with simulation and subsequent combined pretest/posttest (Jankouskas et al., 2007).

Assessment Measures

Fourteen of the 31 studies used a validated tool. Four of these used the Anaesthetists' Nontechnical Skills (ANTS) tool (Jankouskas et al., 2007, 2011; Morgan et al., 2011; Wunder, 2016). ANTS involves assessment of task management, team working, situational awareness, and decision making (Jankouskas et al., 2007, 2011). Two of these studies used this tool on physicians and nurses demonstrating generalizability to disciplines other than anaesthesia (Jankouskas et al., 2007, 2011). According to Jankouskas et al. (2007), using the tool is appealing because it represents the ‘‘generic competencies of any effective health care team’’ (p. 99).

Four studies used the Health Sciences Reasoning Test (Goodstone et al., 2013; Maneval et al., 2012; Shinnick & Woo, 2013; Sullivan-Mann et al., 2009). This multiple-choice test is designed to assess critical thinking of health care profession students (Maneval et al., 2012, p. 129).

Shapiro et al. (2004) studied teamwork through a tool validated in aviation studies called the Team Dimensions Rating Form, consisting of five seven-point, behaviourally anchored, rating scales. Brown and Chronister (2009) also measured critical thinking using an Elsevier product called electrocardiogram SimTest. This is a multiple-choice exam on rhythm strip interpretation. Hall (2015) used the Assessment Technologies Institute content mastery series, a multiple-choice exam testing critical thinking skills. Liaw et al. (2011) used the Rescuing a Patient in Deteriorating Situations (RAPIDS) tool, a 42-item measure evaluating clinical performance. They reported results of a previous study where construct validity and interrater reliability were established.

In the remaining studies, assessment measures varied. There were four studies (DeVita et al., 2005; Lavigne Fadale et al., 2014; Meurling et al., 2013 Ten Eyck et al., 2010) that used ‘‘time to task’’ and number of tasks completed. They measured the time that elapsed before performance of key behaviours. Ten Eyck et al. (2010) commented that tasks were selected based on the criteria that they were ‘‘clear and measurable’’ (p. 141). Although this was not a previously validated assessment scale, this method of evaluation has been used in multiple studies. DeVita et al. (2005), also using ‘‘time to task’’ measures, commented that task completion is an objective measure and is ‘‘less susceptible to inter-rater differences’’ (p. 330). In fact, Lavigne Fadale et al. (2014) reported that 100% interrater reliability for all data sets was achieved.

Another two studies used chart reviews to assess problem solving. Huseman (2012) looked at response times during codes, whereas Wolf (2008) looked at rates of undertriage in an emergency department. In both of these studies, benchmark data were collected prior to the intervention and again afterwards, then compared with initial results.

The remainder of the studies used unvalidated tools. Buckley and Gordon (2011) created a Likert scale survey asking participants to self-identify practice changes post intervention. Dadiz et al. (2013) created a survey and checklist to measure communication. These authors reported that validity was established through expert review, feedback, and pilot testing. They also reported the checklist showed ‘‘excellent inter-rater reliability when the reviewers used it to assess team communication’’ (p. 284). Gilfoyle et al. (2007) created a score-based unvalidated checklist to evaluate performance of key items. This was in addition to a retrospective five-item questionnaire testing knowledge pre—post intervention. Johnson et al. (2014) used a weighted checklist modeled after previously validated checklists. This tool was designed to evaluate performance and interrater reliability was established in a pilot study. Straka et al. (2012) and Przybyl et al. (2015) had participants answer a knowledge questionnaire pre—post test. Sittner et al. (2009) created a multiple-choice test with content validity established by expert panel.

Lasater (2005) developed two quantitative tools to measure clinical judgment. These are the Lasater Clinical Judgment in Simulation Rubric and the Lasater Clinical Judgment in Practice Survey. The author reported that both require refinement. This same author also used a critical thinking measure called the California Critical Thinking Dispositions Inventory. This tool tests for truth-seeking, open-mindedness, analyticity, systematicity, self-confidence, inquisitiveness, and maturity and is an aptitude test toward critical thinking. Bultas et al. (2014) also used a combination of validated and unvalidated tools testing knowledge retention and team performance. The Paediatric Emergency Assessment, Recognitions, and Stabilization (PEARS) course written examination (validated) is a multiple-choice knowledge test and the Mayo High Performance Teamwork Scale (validated) evaluates performance of a team during a crisis. The third tool, the PEARS Behavioural Measures Check-Off Tool (unvalidated), was developed to score team behaviours, skills, and tasks performed during simulations (Bultas et al., 2014).

Schubert (2012) created a critical thinking multiple-choice test in combination with the Learning Transfer Tool, assessing nurses' overall problem-solving skills. Brannan et al. (2008) developed two versions of the Acute Myocardial Infarction Questionnaire: Cognitive Skills Test with content validity established by experts and reliability tested on nursing students prior to the study. Kesten et al. (2015) developed the Advanced Practice Registered Nurse Competency Evaluation Tool. They reported interrater reliability was established and content validity was established through literature review and expert consultation. Singer et al. (2013) adapted an unnamed validated checklist.

Results

CRM training is a method of teaching and practicing team processes (Jankouskas et al., 2007; Morgan et al., 2011). Because these skills are interrelated, most of the studies measured performance of more than one CRM skill. Results for each study and limitations are reported in Table.

Summary

This literature review included 31 papers that measured translation of the CRM outcomes, problem solving, situational awareness, resource utilization, communication, and leadership, after HFS interventions. Study designs included experimental, quasi-experimental, nonexperimental, and quality improvement initiatives. Populations were highly variable in size and composition. Subjects included practicing individuals from medicine, nursing, support staff (health care aides or nurse's aides), and respiratory therapy, as well as students from each of these disciplines. Educational interventions also varied with pre—post test designs as well as comparisons with other educational modalities.

Validated tools measured outcomes in 19 of the 31 studies and the remainder used unvalidated tools. Tools measured a mixture of CRM skills with problem solving being the most frequently studied outcome. Researchers chose a mixture of instruments targeting their outcome of interest.

Results were also mixed, with 22 studies showing improvement in at least one CRM skill. However, there was crossover in some studies, where statistically significant improvement was shown in some areas and no improvement or nonstatistically significant improvement was shown in others. Table presents a summary of each of the categories discussed above for each study included in this review.

Discussion

This literature review has revealed several gaps. First, there is a lack of studies measuring translation of educational interventions on acquisition of CRM skills. Original search methods in 2014 only yielded 20 papers that met criteria for this review. A repeat of this search in 2016 yielded an additional 11 studies that met inclusion criteria. It is interesting to note that the second search yielded more than half the number of studies of the first search and in only a three-year time frame. This indicates a greater interest in observable performance of CRM skills. Continued research may provide insight into the actual effect of HFS on behaviour, which may lead to insight into feasibility of simulation education and effects on patient outcomes.

Second, there was only one meta-analysis found that looked at the effects of HFS in nursing education (Lee & Oh, 2015). This paper presented evidence of the effect of HFS on the cognitive domain of learning, including outcomes like knowledge acquisition, problem solving, critical thinking, clinical judgment, and communication. There was a "significant treatment effect" (p. 504) for all outcomes except communication. With a relatively small number of studies and samples that consisted exclusively of nursing students, the authors concluded HFS "might have beneficial effects on cognitive and psychomotor domain of learning" (Lee & Oh, 2015, p. 505). Combined effects of data sets from all health care disciplines and levels of practice (i.e., practicing professionals and students) may reveal more definitive conclusions. This may be difficult to achieve, however, without consistency between studies. This lack of consistency is perhaps because research in simulation education and using CRM as a framework is still new in health care. Increased interest in these areas may yield a greater volume and quality of evidence that will add to overall results.

Third, the lack of consistency between studies highlights the need for ongoing research into optimal educational formats and evaluative methods. Most of the authors described one-time—only interventions, showing mixed results. Those that did present multiple interventions stated they may have been too few or too widely spaced in time to reflect positive results (Morgan et al., 2011). Incorporating multiple simulation episodes would increase opportunities to practice skills, and spaced over shorter time spans would also allow learners to build and reinforce skills.

Several authors listed their choice of tool as a limitation, citing applicability to population and lack of validation. According to Oermann, Kardong-Edgren, and Rizzolo (2016), important steps to conducting simulation-based education and evaluation include careful consideration of learning objectives with appropriately designed simulations. Assessment tools need to be selected based on these targeted objectives, while also ensuring validity and reliability of the instruments. Evaluators require training to ensure consistency throughout evaluation. Although these recommendations were made for summative assessments (e.g., end-of-program comprehensive competencies), these principles can be applied to other education or assessment activities as well. It is clear there is a need to match proven tools to educational outcomes, but it is difficult to choose an evaluation method when there is insufficient evidence to fully illustrate which tool is the best choice for any given outcome. Future studies comparing and correlating the achievement of targeted outcomes to the use of validated or unvalidated evaluation methods may illustrate gaps in study design and further guide tool selection for future educators.

Other implications for future research may include a focus on evaluating performance of CRM behaviours in actual clinical areas. For this to be possible, a focus on teaching CRM concepts to a broader audience, making it part of health care culture in a way that is similar to aviation culture, is needed. Simulation education would need to be a standard in crisis education for all disciplines, along with an emphasis on team outcomes versus discipline specific outcomes. Once CRM culture has been broadly achieved, the impact on patient outcomes may be observable as well.

References

Andersen, P., Jensen, M., Lippert, A., & Ostergaard, D. (2010). Identifying non-technical skills and barriers for improvement of teamwork in cardiac arrest teams. *Resuscitation*, *81*, 695-702. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.resuscitation.2010.01.024.

Bearman, M., O'Brien, R., Anthony, A., Civil, I., Flanagan, B., Jolly, B., …, & Nestel, D. (2012). Learning surgical communication, leadership and teamwork through simulation. *Journal of Surgical Education*, *69*(2), 201-207. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.jsurg.2011.07.014.

Brannan, J., White, A., & Bezanson, J. (2008). Simulator effects on cognitive skills and confidence levels. *Journal of Nursing Education*, *47*(11), 495-500. http://dx.doi.org.uml.idm.oclc.org/10.3928/01484834-20081101-01.

Brown, D., & Chronister, C. (2009). The effect of simulation learning on critical thinking and self-confidence when incorporated into an electrocardiogram nursing course. *Clinical Simulation in Nursing*, *5*(1), e45-e52. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.ecns.2008.11.001.

Buckley, T., & Gordon, C. (2011). The effectiveness of high fidelity simulation on medical-surgical registered nurses' ability to recognize and respond to clinical emergencies. *Nurse Education Today*, *31*, 716-721. http://dx.doi.org.uml.idm.oclc.org/10.3928/00220124-20091023-06.

Bultas, M., Hassler, M., Ercole, P., & Rea, G. (2014). Effectiveness of high-fidelity simulation for pediatric staff nurse education. *Pediatric Nursing*, *40*(1), 27-32.

Dadiz, R., Weinschreider, J., Schriefer, J., Arnold, C., Greves, C., Crosby, E., …, & Guillet, R. (2013). Interdisciplinary simulation-based training to improve delivery room communication. *The Society for Simulation in Healthcare*, *8*, 279-291.

DeVita, M., Schaefer, J., Wang, H., & Dongilli, T. (2005). Improving medical emergency team (MET) performance using a novel curriculum and a computerized human patient simulator. *Quality and Safety in Healthcare*, *14*, 326-331. http://dx.doi.org.uml.idm.oclc.org/10.1136/qshc.2004.011148.

Gaba, D. (2010). Crisis resource management and teamwork training in anaesthesia. *British Journal of Anaesthesis*, *105*(1), 3-6. https://doi.org/10.1093/bja/aeq124.

Gilfoyle, E., Gottesman, R., & Razack, S. (2007). Development of a leadership skills workshop in paediatric advanced resuscitation. *Medical Teacher*, *29*(9-10), e276-e283. http://dx.doi.org.uml.idm.oclc.org/10.1080/01421590701663287.

Goodstone, L., Goodstone, K., Glaser, C., Kupferman, K., & Dember-Neal, T. (2013). Effect of simulation on the development of critical thinking in associate degree nursing students. *Nursing Education Perspectives*, *34*(3), 159-162.

Gordon, S., Mendenhall, P., & Blair O'Connor, B. (2013). *Beyond the checklist. What else health care can learn from aviation teamwork and safety.* Ithaca, NY: ILR Press.

Hall, S. (2015). High-fidelity simulation for senior maternity nursing students. *Nursing Education Perspective*, *36*(2), 124-126.

Huseman, K. (2012). Improving code blue response through the use of simulation. *Journal for Nurses in Staff Development*, *28*(3), 120-124.

Jankouskas, T., Chaska Bush, M., Murray, B., Rudy, S., & Henry, J. (2007). Crisis resource management: Evaluating outcomes of a multidisciplinary team. *Society for Simulation in Healthcare*, *2*(2), 96-101.

Jankouskas, T. S., Haidet, K., Hupcey, J., Kolanowski, A., & Murray, M. (2011). Targeted crisis resource management improves performance among randomized nursing and medical students. *Journal of the Society for Simulation in Healthcare*, *6*(6), 316-326.

Johnson, M., Hickey, K., Scopa-Goldman, J., Andrews, T., Boerem, P., Covec, M., & Larson, E. (2014). Manikin versus web-based simulation for advanced practice nursing students. *Clinical Simulation in Nursing*, *10*(6), e317-e323. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.ecns.2014.02.004.

Kesten, K., Brown, H., & Meeker, M. (2015). Assessment of APRN student competency using simulation: A pilot study. *Nursing Education Perspectives*, *36*(5), 332-334.

Kim, J., Neilipovitz, D., Cardinal, P., Chiu, M., & Clinch, J. (2006). A pilot study using high-fidelity simulation to formally evaluate performance in the resuscitation of critically ill patients: The University of Ottawa critical care medicine, high-fidelity simulation and crisis resource management I study. *Critical Care Medicine*, *34*(8), 2167-2174.

Lapkin, S., Levett-Jones, T., Bellchambers, H., & Fernandez, R. (2010). Effectiveness of patient simulation manikins in teaching clinical reasoning skills to undergraduate nursing students: A systematic review. *Clinical Simulation in Nursing*, *6*(6), e207-e222. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.ecns.2010.05.005.

Lasater, K. (2005). *The impact of high fidelity simulation on the development of clinical judgment in nursing students: An exploratory study.* (doctoral dissertation). Portland, OR: Portland State University.

Lavigne Fadale, K., Tucker, D., Dungan, J., & Sabol, V. (2014). Improving nurses' vasopressor titration skills and self-efficacy via simulation-based learning. *Clinical Simulation in Nursing*, *10*(6), e291-e299. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.ecns.2014.02.002.

Lee, J., & Oh, P. (2015). Effects of the use of high-fidelity human simulation in nursing education: A meta-analysis. *Journal of Nursing Education*, *54*(9), 501-507.

Liaw, S., Rethans, J., Scherpbier, A., & Piyanee, K. (2011). Rescuing a patient in deteriorating situations (RAPIDS): A simulation-based educational program on recognizing, responding and reporting of physiological signs of deterioration. *Resuscitation*, *82*, 1224-1230. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.resuscitation.2011.04.014.

Maneval, R., Fowler, K., Kays, J., Boyd, T., Shuey, J., Harne-Britner, S., & Mastrine, C. (2012). The effect of high-fidelity patient simulation on the critical thinking and clinical decision making skills of new graduate nurses. *The Journal of Continuing Education in Nursing*, *43*(3), 125-134.

Messmer, P. R. (2008). Enhancing nurse-physician collaboration using pediatric simulation. *The Journal of Continuing Education in Nursing*, *39*(7), 319-327. http://dx.doi.org.uml.idm.oclc.org/10.3928/00220124-20080701-07.

Meurling, L., Hedman, L., Fellander-Tsai, L., & Wallin, C. (2013). Leaders' and followers' individual experiences during the early phase of simulation-based training: An exploratory study. *British Medical Journal Quality and Safety*, *22*, 459-467. http://dx.doi.org.uml.idm.oclc.org/10.1136/bmjqs-2012-000949.

Morgan, P., Kurrek, M., Bertram, S., LeBlanc, V., & Przybyszewski, T. (2011). Nontechnical skills assessment after simulation-based continuing medical education. *Society for Simulation in Healthcare*, *6*(5), 255-259.

Oermann, M. H., Kardong-Edgren, S., & Rizzolo, M. (2016). Summative simulated-based assessment in nursing programs. *Journal of Nursing Education*, *55*(6), 323-328.

Pearson, E., & McLafferty, I. (2011). The use of simulation as a learning approach to non-technical skills awareness in final year student nurses. *Nurse Education in Practice*, *11*, 399-405. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.nepr.2011.03.023.

Przybyl, H., Androwich, I., & Evans, J. (2015). Using high-fidelity simulation to assess knowledge, skills, and attitudes in nurses performing CRRT. *Nephrology Nursing Journal*, *42*(2), 135-147.

Rudy, S., Polomano, R., Murray, W., Henry, J., & Marine, R. (2007). Team management training using crisis resource management results in perceived benefits by healthcare workers. *The Journal of Continuing Education in Nursing*, *38*(5), 219-226.

Schubert, C. (2012). Effect of simulation on nursing knowledge and critical thinking in failure to rescue events. *The Journal of Continuing Education in Nursing*, *43*(10), 467-471.

Shapiro, M., Morey, J., Small, S., Langford, V., Kaylor, C., Jagminas, L., …, & Jay, G. (2004). Simulation based teamwork training for emergency department staff: Does it improve clinical team performance when added to an existing didactic teamwork curriculum? *Quality and Safety in Healthcare*, *13*, 417-421. http://dx.doi.org.uml.idm.oclc.org/10.1136/qshc.2003.005447.

Shinnick, M., & Woo, M. (2013). The effect of human patient simulation on critical thinking and its predictors in prelicensure nursing students. *Nurse Education Today*, *33*, 1062-1067. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.nedt.2012.04.004.

Singer, B., Corbridge, T., Schroedl, C., Wilcox, J., Cohen, E., McGaghie, W., & Wayne, D. (2013). First-year residents outperform third-year residents after simulation-based education in critical care medicine. *The Society for Simulation in Healthcare*, *8*(2), 67-71.

Sittner, B. J., Schmaderer, M., Zimmerman, L., Hertzog, M., & George, B. (2009). Rapid response team simulated training for enhancing patient safety (STEPS). *Clinical Simulation in Nursing*, *5*(3), 119-127. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.ecns.2009.02.007.

Straka, K., Burkett, M., Capan, M., & Eswein, J. (2012). The impact of education and simulation on pediatric novice nurses' response and recognition to deteriorating. *Journal for Nurses in Staff Development*, *28*(6), e5-e8.

Sullivan-Mann, J., Perron, C., & Fellner, A. (2009). The effects of simulation on nursing students' critical thinking scores: A quantitative study. *Newborn and Infant Nursing Reviews*, *9*(2), 111-116. http://dx.doi.org.uml.idm.oclc.org/10.1053/j.nainr.2009.03.006.

Ten Eyck, R., Tews, M., Ballester, J., & Hamilton, G. (2010). Improved fourth-year medical student clinical decision-making performance as a resuscitation team leader after a simulation-based curriculum. *Society for Simulation in Healthcare*, *5*(3), 139-145.

White, N. (2012). Understanding the role of non-technical skills in patient safety. *Nursing Standard*, *26*(26), 43-48.

Wolf, L. (2008). The use of human patient simulation in ED triage can improve nursing confidence and patient outcomes. *Journal of Emergency Nursing*, *34*(2), 169-171. http://dx.doi.org.uml.idm.oclc.org/10.1016/j.jen.2007.11.005.

Wunder, L. (2016). Effect of a nontechnical skills intervention on first-year student registered nurse anesthetists skills during crisis simulation. *AANA Journal*, *84*(1), 46-51.