Development of Crisis Resource Management Skills: A Literature Review

Amanda Lucas, RN, BN\textsuperscript{a,}* Marie Edwards, RN, PhD\textsuperscript{b}

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

\textbf{KEYWORDS}
crisis resource management; high-fidelity simulation; literature review

\textbf{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.


© 2017 International Nursing Association for Clinical Simulation and Learning. Published by Elsevier Inc. All rights reserved.

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, Neulipovitz, 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,
Lippert, & Østergaard, 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.

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:


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>
<thead>
<tr>
<th>First Author, Year</th>
<th>Methods</th>
<th>Sample</th>
<th>Educational Intervention</th>
<th>Assessment Measures</th>
<th>CRM Skill and Major Findings</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brannan, 2008</td>
<td>Quasi-experimental, single-sample pre–post test</td>
<td>Nursing only (student), N = 107</td>
<td>HFS vs. alternate</td>
<td>Adapted validated tool (AMIQ)</td>
<td>PS (I)—higher posttest scores with intervention group compared with control group</td>
<td>Lack of randomization to groups</td>
</tr>
<tr>
<td>Brown, 2009</td>
<td>Experimental, comparative correlational</td>
<td>Nursing only (student), N = 140</td>
<td>HFS vs. alternate</td>
<td>Validated (ECG SimTest)</td>
<td>PS (NI)—no significant differences in critical thinking between measures</td>
<td>Lack of control for previous experience, small sample size, unequal treatment of groups</td>
</tr>
<tr>
<td>Buckley, 2011</td>
<td>Nonexperimental</td>
<td>Nursing only (practicing), N = 38</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Unvalidated (checklist or questionnaire)</td>
<td>PS (I)—improved critical abilities with airway management</td>
<td>Performance not directly evaluated difficult to differentiate cause of improvement</td>
</tr>
<tr>
<td>Bultas, 2014</td>
<td>Quasi-experimental, prospective observational pretest/posttest with control group</td>
<td>Nursing only (practicing), n = 33</td>
<td>HFS vs. static manikin</td>
<td>Validated (PEARS written exam and MHPT), unvalidated (PEARS BCMT)</td>
<td>PS (M)—HFS group BCMT scores significantly higher than control at posttest</td>
<td>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</td>
</tr>
<tr>
<td>Dadiz, 2013</td>
<td>Quasi-experimental, prospective observational</td>
<td>Multidisciplinary (practicing), N = 228</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Validated (checklist or questionnaire)</td>
<td>C (I)—checklist scores improved</td>
<td>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</td>
</tr>
<tr>
<td>DeVita, 2005</td>
<td>Quality improvement</td>
<td>Multidisciplinary (practicing), N = 138</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Unvalidated (time-to-task and measure of simulator)</td>
<td>PS (I), RU (I), L (I)—simulators “survived” from 0% of the time to 89% of the time, time-to-task</td>
<td>No interrater reliability of performance ratings, scenarios different</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Methods</th>
<th>Sample</th>
<th>Educational Intervention</th>
<th>Assessment Measures</th>
<th>CRM Skill and Major Findings</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilfoyle, 2007</td>
<td>Quasi-experimental, within-sample pretest with staged posttest</td>
<td>Physician only (paediatric residents), N = 15</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Unvalidated (checklist or questionnaire)</td>
<td>&quot;survival&quot;) measures improved posttraining intervention C (M)—no learning of occurred during maintenance period, but no decay of learning occurred either L(I)—increase with checklist leadership scores</td>
<td>Unvalidated scenarios and checklist tool; possible &quot;training effect&quot; related to use of same scenario at six-month session between groups</td>
</tr>
<tr>
<td>Goodstone, 2013</td>
<td>Quasi-experimental, two-group pretest/posttest</td>
<td>Nursing only (students), N = 42, n = 41</td>
<td>Pretest/posttest with HFS vs. case study</td>
<td>Validated (HSRT)</td>
<td>PS (M)—both groups scored significantly higher on posttest, indicating both style of simulation education equally effective</td>
<td>Lack of randomization, small sample, lack of no-treatment control group, potential testing effect</td>
</tr>
<tr>
<td>Hall, 2015</td>
<td>Quasi-experimental, retrospective, comparative</td>
<td>Nursing-only (students), N = 279</td>
<td>HFS with traditional experience vs. traditional</td>
<td>Validated (ATI content mastery series test)</td>
<td>PS (I)—significantly higher scores on posttest with experimental group compared with control</td>
<td>Limited generalizability, differences in sample</td>
</tr>
<tr>
<td>Huseman, 2012</td>
<td>Quasi-experimental, single-sample pretest/posttest</td>
<td>Multidisciplinary (direct care providers), N = 178</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Unvalidated (chart review)</td>
<td>PS (M)—statistically significant improvement in some area, but not in other areas, also not sustained over time</td>
<td>Not addressed</td>
</tr>
<tr>
<td>Jankouskas, 2007</td>
<td>None experimental</td>
<td>Multidisciplinary (practicing), N = 40</td>
<td>Theory with HFS with pretest/posttest</td>
<td>Validated (ANTS)</td>
<td>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</td>
<td>Element of situation awareness was difficult to visualize on videotaped simulations, making it difficult to evaluate</td>
</tr>
<tr>
<td>Jankouskas, 2011</td>
<td>Experimental, randomized, controlled, pretest/posttest experimental</td>
<td>Multidisciplinary (students), N = 96</td>
<td>HFS vs. alternate</td>
<td>Validated (ANTS)</td>
<td>SA (I)—significant improvements in situational awareness RU (I)—significant increase in task management and teamwork measures C (I)—a statistically significant improvement L (I)—significant improvement</td>
<td>Limited generalizability; diffusion of treatment over time, history effect, potential bias from unblinded principle investigator, use of inappropriate tool</td>
</tr>
<tr>
<td>Johnson, 2013</td>
<td>Quasi-experimental</td>
<td>Nursing only</td>
<td>HFS vs. Web-manikin group</td>
<td>Unvalidated</td>
<td>PS (I), RU (I)—manikin group</td>
<td>Small sample, varied</td>
</tr>
<tr>
<td>First Author, Year</td>
<td>Methods</td>
<td>Sample</td>
<td>Educational Intervention</td>
<td>Assessment Measures</td>
<td>CRM Skill and Major Findings</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>--------</td>
<td>--------------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Kesten, 2015</td>
<td>Repeated measures pilot study</td>
<td>Nursing only (practicing), N = unknown</td>
<td>Repeated HFS intervention with evaluation over time</td>
<td>Validated (APRN EVAL tool)</td>
<td>PS (I), SA (I), C (I), L (I)—statistically significant improvement in scores over time, continued improvement from time 3-4 in leadership domain</td>
<td>Small sample, lack of standardization of tools for population, differences with HFS scenarios, previous experience affecting CRM learning</td>
</tr>
<tr>
<td>Lasater, 2005</td>
<td>Mixed methods</td>
<td>Nursing only (student nurses), N = 48</td>
<td>Theory course with complementary HFS sessions</td>
<td>Unvalidated (LCJSR, LCJPS) with validated (CCTDI)</td>
<td>PS (I)—statistically significant outcomes in critical thinking</td>
<td>Involvement of a single centre, lack of comparable baseline data, and lack of sample variation</td>
</tr>
<tr>
<td>Lavigne Fadale, 2014</td>
<td>Quasi-experimental pretest/posttest design</td>
<td>Critical care or emergency room nurses, N = 16</td>
<td>Repeated intervention with repeated testing over time</td>
<td>Unvalidated (time to task, number of interventions [vasopressor titrations])</td>
<td>PS (M)—trend toward significance for most points with only statistically significant improvement in number of vasopressor titrations</td>
<td>Small sample, potential for performance bias, equipment malfunction</td>
</tr>
<tr>
<td>Liaw, 2011</td>
<td>Experimental, prospective, randomized, control trial with a pretest/posttest design</td>
<td>Nursing only (students), N = 31</td>
<td>Pretest/posttest with HFS intervention vs. control without HFS intervention</td>
<td>Unvalidated (RAPIDS tool)</td>
<td>PS (I), C (I)—significant improvement over control group</td>
<td>Not generalizable, unknown long-term retention of skills, single-scenario exposure, lack of applicability to actual clinical setting</td>
</tr>
<tr>
<td>Maneval, 2012</td>
<td>Experimental, randomized, controlled, pretest/posttest experimental</td>
<td>Nursing only (novice nurse) N = 26</td>
<td>HFS vs. alternate</td>
<td>Validated (HSRT)</td>
<td>PS (NI)—nonstatistically significant improvements in mean posttest scores</td>
<td>Small sample size, higher than national average pretest scores, and a lack of management support affecting participation</td>
</tr>
<tr>
<td>Meurling, 2013</td>
<td>Quasi-experimental exploratory</td>
<td>Medical students, N = 54</td>
<td>3 HFS scenarios with evaluation over time</td>
<td>Unvalidated (a TEAM programme, i.e., mix of time-to-task, frequency of behaviours)</td>
<td>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</td>
<td>Sample not generalizable</td>
</tr>
<tr>
<td>First Author, Year</td>
<td>Methods</td>
<td>Sample</td>
<td>Educational Intervention</td>
<td>Assessment Measures</td>
<td>CRM Skill and Major Findings</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------------------</td>
<td>---------------------------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>-------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Morgan, 2011</td>
<td>Experimental, randomized, controlled, blinded</td>
<td>Physician only (anaesthesiologists), N = 59</td>
<td>HFS alone vs. HFS with debrief</td>
<td>Validated (ANTS)</td>
<td>PS (NI)—no improvement in nontechnical skill performance</td>
<td>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)</td>
</tr>
<tr>
<td>Przybyl, 2015</td>
<td>Quality improvement, pretest/posttest</td>
<td>Nursing only (practicing), N = 93</td>
<td>CRRT course with addition of HFS education</td>
<td>Unvalidated (knowledge questionnaire)</td>
<td>PS (no comment on significance)—increase in understanding of CRRT principles and critical thinking related to operation of CRRT machine</td>
<td>Challenges with recruitment</td>
</tr>
<tr>
<td>Schubert, 2012</td>
<td>Quasi-experimental, within-sample pretest with staged posttest</td>
<td>Nursing only (practicing), N = 58</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Validated (multiple choice) with LTT</td>
<td>PS (I)—statistically significant improvement in problem-solving skills</td>
<td>Untested tool, small sample, high attrition rate, participants working previous night shift may have affected performance</td>
</tr>
<tr>
<td>Shapiro, 2004</td>
<td>Quasi-experimental, prospective observational</td>
<td>Multidisciplinary (practicing), N = 20</td>
<td>HFS vs. alternate</td>
<td>Validated (team dimensions rating form)</td>
<td>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</td>
<td>Small sample size</td>
</tr>
<tr>
<td>Shinnick, 2013</td>
<td>Quasi-experimental, one-group pretest/posttest</td>
<td>Prelicensure nursing students, N = 154</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Validated (HSRT)</td>
<td>PS (NI)—gains in knowledge without statistically significant gains in critical thinking</td>
<td>Previous experience with scenario content and simulation environment</td>
</tr>
<tr>
<td>Singer, 2013</td>
<td>Quasi-experimental, prospective cohort</td>
<td>Physician only (students), N = 67</td>
<td>HFS vs. alternate</td>
<td>Adapted validated checklist</td>
<td>PS (I)—first-year medical residents outperformed</td>
<td>Single-centre study with limited numbers,</td>
</tr>
<tr>
<td>First Author, Year</td>
<td>Methods</td>
<td>Sample</td>
<td>Educational Intervention</td>
<td>Assessment Measures</td>
<td>CRM Skill and Major Findings</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>--------</td>
<td>-------------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Sittner, 2009</td>
<td>Quasi-experimental, within-sample pretest with staged posttest</td>
<td>Nursing only (practicing), N = 11</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Validated (checklist or questionnaire)</td>
<td>third-year residents after completing an HFS intervention</td>
<td>not blinded and limited number of competencies assessed</td>
</tr>
<tr>
<td>Straka, 2012</td>
<td>Quasi-experimental, single-sample pretest/posttest</td>
<td>Nursing only (novice nurse), N = 26</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Unvalidated (checklist or questionnaire)</td>
<td>PS (I) — significant improvement in problem-solving skills</td>
<td>Small sample with no control group, participants previous experience potentially affecting pretest scores</td>
</tr>
<tr>
<td>Sullivan-Mann, 2009</td>
<td>Experimental, two group by two times mixed model</td>
<td>Nursing only (student nurse), N = 53</td>
<td>3 HFS sessions vs. 5 HFS sessions</td>
<td>Validated (HSRT)</td>
<td>PS (I) — statistically significant improvement in problem-solving skills</td>
<td>Time frame, small sample, lack of standardized tool</td>
</tr>
<tr>
<td>Ten Eyck, 2010</td>
<td>Experimental, randomized, controlled, blinded</td>
<td>Physician only (students), n = 68</td>
<td>HFS vs. alternate</td>
<td>Unvalidated (time to task)</td>
<td>PS (M) — statistically significant improvement in times in 4 of 8 critical tasks L (I) — showed more significant improvement as the program progressed</td>
<td>Small sample, differences in instruction between groups, lack of no intervention control, may not be generalizable</td>
</tr>
<tr>
<td>Wolf, 2008</td>
<td>Quasi-experimental, single-sample pretest/posttest</td>
<td>Nursing only (practicing), n = 6</td>
<td>Pretest/posttest with HFS intervention</td>
<td>Unvalidated (chart review)</td>
<td>PS (I) — improvement from 40% accurate triage to 70%-100% postsimulation intervention</td>
<td>Lack of generalizability to “real” setting, other institutions, and different situations; lack of validated instrument; single evaluator of cases</td>
</tr>
<tr>
<td>Wunder, 2016</td>
<td>Quasi-experimental, pretest/posttest</td>
<td>Nursing only (students), n = 32</td>
<td>Repeated HFS with repeated testing and CRM lecture</td>
<td>Validated (ANTS)</td>
<td>PS (I), SA (I), RU (I), C (I), L (I) — increase in total score overall, breakdown not given</td>
<td>Lack of interrater reliability and rater familiarity with students</td>
</tr>
</tbody>
</table>

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; PEAR = 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, Hupecey, Kolanowski, & Murray, 2011; Liaw, Rethans, Scherpier, & Piyane, 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


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


