Original Article

The impact of critical event checklists on medical management and teamwork during simulated crises in a surgical daycare facility*

T. C. Everett,[1] P. J. Morgan,[2] R. Brydges,[3] M. Kurrek,[4] D. Tregunno,[5] L. Cunningham,[6] A. Chan,[7] D. Forde,[8] and J. Tarshis,[9]

1 Assistant Professor and Staff Anesthesiologist, The Hospital for Sick Children, University of Toronto, Ontario, Canada
2 Professor and Staff Anesthesiologist, 6 Research Assistant, 8 Senior Nurse, Women's College Hospital, University of Toronto, Toronto, Ontario, Canada
3 Assistant Professor and Scientist, 4 Associate Professor and Staff Anesthesiologist, University of Toronto, Toronto, Ontario, Canada
5 Associate Professor, School of Nursing, Queen's University, Kingston, Ontario, Canada
7 Anesthesia Assistant, Toronto Western Hospital, Toronto, Ontario, Canada
9 Associate Professor and Staff Anesthesiologist, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada

Summary

Although the incidence of major adverse events in surgical daycare centres is low, these critical events may not be managed optimally due to the absence of resources that exist in larger hospitals. We aimed to study the impact of operating theatre critical event checklists on medical management and teamwork during whole-team operating theatre crisis simulations staged in a surgical daycare facility. We studied 56 simulation encounters (without and with a checklist available) divided between an initial session and then a retention session several months later. Medical management and teamwork were quantified via percentage adherence to key processes and the Team Emergency Assessment Measure, respectively. In the initial session, medical management was not improved by the presence of a checklist (56% without checklist vs. 62% with checklist; $p = 0.50$). In the retention session, teams performed significantly worse without the checklists (36% without checklist vs. 60% with checklist; $p = 0.04$). We did not observe a change in non-technical skills in the presence of a checklist in either the initial or retention sessions (68% without checklist vs. 69% with checklist ($p = 0.94$) and 69% without checklist vs. 65% with checklist ($p = 0.36$), respectively). Critical events checklists do not improve medical management or teamwork during simulated operating theatre crises in an ambulatory surgical daycare setting.

Correspondence to: T. C. Everett
Email: tobias.everett@sickkids.ca
Accepted: 19 August 2016
Keywords: cognitive aids; critical events; simulation
See Appendix for list of co-investigators.
**Presented at the Royal College of Physicians and Surgeons of Canada Simulation Summit, Banff, Alberta, Canada, November 2015.*
This article is accompanied by an editorial by Marshall, Anaesthesia 2016; 71: doi: 10.1111/anae.13707 and an article by Harvey et al., Anaesthesia 2016, 71: doi: 10.1111/anae.13643.

Introduction

In recent years there has been an increasing research focus on cognitive aids and checklists in peri-operative care. Although there is clearly some overlap in these resources, a checklist is a sequence of definitive action points for ticking off in a given situation, whereas a cognitive aid is broader and may include decision branch-points and prompts for a practitioner's consideration/discretion. Multiple other terms for standardised practices are in common use [1]. Checklists may be considered in two categories: those that are used electively and routinely (e.g. WHO Surgical Safety Checklist [2]); and those that are used in emergencies and rarely (e.g. malignant hyperthermia crisis safety guideline [3]). The seminal paper by Haynes et al. provided evidence of the patient safety benefit of implementation of the surgical safety checklist into the pre-induction routine [4], although interestingly this was not reproduced in a subsequent study in a different jurisdiction [5]. Evidence for cognitive aids used in a peri-operative crisis is similarly mixed [1], although the availability of a cognitive aid has been shown to improve adherence to best-practice guidelines and completion of critical tasks in an emergency [6–11]. Teamwork and non-technical skills have also been shown to improve where a cognitive aid is available [9, 12]. However, other research reveals concerning deficits in clinical performance despite the availability of relevant cognitive aids [13–18]. Clearly there are unanswered questions regarding the contextual and implementation factors that are associated with the maximal positive impact of crisis checklists.

The use of critical event checklists (hereafter referred to as 'checklists') in tertiary-level hospitals during simulated crises has been shown to result in superior adherence to key processes [10, 11]. We sought to investigate the impact of the same checklists in a different context, namely a surgical daycare facility. In these facilities, serious incidents and deaths are rare (5.3 and 0.78 per 100,000 procedures, respectively [19]), and unlike large teaching hospitals, they have lean clinical teams and no dedicated rapid response teams. With events being rare and without specialised resuscitation teams, we theorised that management of crises would stand to benefit greatly from the introduction of checklists.

Our hypothesis was that in a surgical daycare facility, completion of key processes during a simulated crisis would be improved when using a checklist when compared with management without a checklist available. Our secondary hypothesis was that teamwork would be improved where the checklist was available, consistent with other work in similar contexts [9, 12]. Finally, we planned to investigate the retention or attrition of any benefits gained from the use of a checklist over time.

Methods

After Research Ethics Board approval, all staff working in an ambulatory surgical facility were invited to participate in the study. The facility contains six operating theatres and conducts around 4400 adult general anaesthetics per year. Each theatre team comprised a surgeon, anaesthetist and three operating theatre nurses. Practitioners were randomly assigned to teams in order to eliminate the potentially confounding effect of certain combinations of practitioners being more accustomed to working together. With respect to sample size, previous work demonstrated that in a tertiary-level teaching hospital, the failure to adhere to key processes moved from 23% without checklists to 6% with checklists available [10]. In an initial pilot of these checklists, these figures were 24% and 4%, respectively [11]. Using these measures of effect and associated statistics, we made our preliminary sample size calculations. We allowed for the possibility of a smaller effect size in our context and increased our sample size accordingly. Ultimately we studied a total of 56 simulation encounters.

At each session we ran peri-operative in situ crisis simulations with and without checklists available. We used an A-B-A-B experimental design, where each team represented a 'case' and served as its own control [20–22]. In the A-B-A-B design, 'A' refers to baseline conditions (without checklist), and 'B' refers to the intervention condition (with checklist). Using this design, the repeated baseline measurements established the pattern of scores that we expected the intervention condition to change [23]. We designed eight scenarios

which we ordered randomly for each team, with four simulations in an initial session and four in a later, retention session (Fig. 1). This randomisation distributed evenly the possibility that experience in particular scenarios might confer an advantage in subsequent scenarios. The random order also avoided the situation where certain scenarios always had a checklist available, whereas other scenarios would not. The eight scenarios were: myocardial infarction leading to pulseless activity; venous air embolus; unstable bradycardia (third degree heart block); malignant hyperpyrexia; anaphylaxis; fire in the operating theatre; unstable atrial fibrillation; and tension pneumothorax. Patients' characteristics and history were typical for a daycare facility and the operation was customised to the specialty of the surgeon participant. Scenarios were piloted in advance and their difficulty assessed, adjusted and standardised by expert clinician/simulation educators.

The scenarios took place outside normal working hours in the operating theatre, pre-operative area or recovery room. Participants received a standardised orientation to the objectives of the session and manikin. The participants were not blinded to the objective of the study and were given a full orientation to the checklists, including a point-by-point orientation to a sample checklist (for an event that would not be used during that session). We coached the teams with respect to how best to use the checklists and the team had the opportunity to review the whole binder of 12 checklists [11]. We did not direct the teams regarding who should be either the team leader or checklist reader. Before each simulation scenario started, we read participants a stem consisting of case history and stage of procedure. All team members were familiar with the surgical equipment and the environment where the simulations were conducted. Real surgical equipment was available matched to the procedure in the scenario. Sham anesthesia medications were prepared in syringes sized consistent with local practice.

In addition to the computerised manikin (SimMan® 380000; Laerdal Medical Canada Ltd, Toronto, ON, Canada), we used hybrid simulation to augment physical resemblance and functional task alignment [24]. This involved mounting a task trainer matched to the surgical case over the manikin to provide a realistic cognitive load for the surgeon

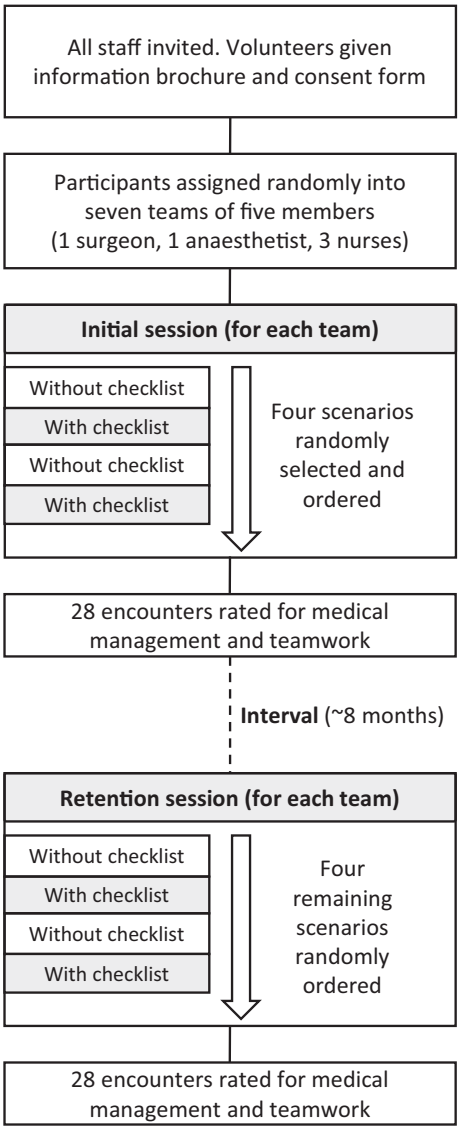

Figure 1 Participant flow diagram.

participant during the case. Task trainers included a laparoscopic suturing exercise, shoulder arthroscopic suturing exercise and open wound suturing task trainer, mounted in an anatomically appropriate location for the simulated surgery. For simulation of the operating room fire, we used a fog machine (Chauvet Lighting®, Sunrise, FL, USA) to fill the operating theatre with 'smoke'. We video recorded all the performances for subsequent performance rating.

Our primary outcome measure was adherence to the critical event key processes developed in the initial investigation of the checklists [11], expressed as the proportion of total number of key processes that were

satisfactorily completed. The key processes are specific, directly observable behaviours that do not require subjective judgement of an investigator and were recorded in real time. We also noted if checklists were used (when available), which checklist was used (i.e. correctly matched to the scenario), and which team member was responsible for holding/reading the checklist. We evaluated the non-technical components of the team performance using the Team Emergency Assessment Measure (TEAM), a measurement instrument developed for rating non-technical skills of medical emergency teamwork [25]. Arguments have been made for the validity and reliability of the TEAM in contexts similar to our own [26]. TEAM assesses the following: leadership (two items); team work (seven items); task management (two items); and an overall score from 1 to 10. We recruited four video raters (two doctors and two nurses), reflecting the professions of our interprofessional teams, to score the non-technical/behavioural performance of the teams using the TEAM. Three of the four reviewers had prior experience with non-technical skills assessments of recorded team simulations. Video raters were trained during a single day workshop, at which raters were oriented to the TEAM and shown a sequence of whole-team medical emergency simulations unrelated to the current study. Raters used the TEAM to rate sample performances and then discussed their scores, exploring the rationale behind the scores they chose. The activity was repeated with fresh sample videos until the raters scores began to converge and they demonstrated adequate inter-rater reliability (defined as an intraclass correlation coefficient greater than 0.8.) Once trained, raters then scored the study videos (all raters rated all the videos). Raters were blinded to participant identity and scored all sessions independently. Raters could see that a checklist was being used in some scenarios, but they were not informed of the study hypothesis or methodology in an effort to minimise any potential observer bias. Raters viewed the videos in a randomised order with respect to teams, scenarios and availability of checklists. The final score ascribed to each simulation encounter was an aggregate of TEAM scores across elements and raters and expressed as a proportion of maximum possible TEAM score. We also collected Kirkpatrick level one [27] data from the learners regarding their perceptions and reactions to the educational intervention, in the form of exit questionnaires using Likert scales.

We developed and standardised scenario-specific debriefing templates. The debriefing philosophies embedded in these templates were a blend of current models in debriefing, similar to the Promoting Excellence and Reflective Learning in Simulation (PEARLS) technique [28]. One of two debriefers watched the live performances of all teams and conducted the 45-min debriefing session, during which participants were encouraged to recall, analyse and reflect upon the scenario events and their reactions to them. If a deficit of knowledge or judgement had been noted during the scenario, the debriefer would explore the cognitive foundation of the issue and close the performance gap.

We expressed the impact of checklist availability on management using Hedges' g effect-size; that is, the standardised mean difference of each team's percentage adherence to key processes with and without the checklist available. We analysed the TEAM scores in the same way. Data were managed and analysed using Excel 2013 (Microsoft Corporation, Redmond, WA, USA) and Stata v12 (StataCorp LP, College Station, TX, USA).

Results

Seven multidisciplinary teams each participated in eight scenarios, producing a total of 56 simulation encounters (28 with and 28 without checklists). The sessions were undertaken in two sessions separated by a mean (SD) of 7.9 (1.2) months. Participants were well matched in terms of baseline characteristics (Table 1).

With respect to medical management, in the initial sessions the teams did not perform better when the checklist was available with mean (SD) adherence to key processes 56% (26%) without the checklist vs. 62% (26%) with the checklist ($p = 0.50$, Hedges' $g = 0.25$, 95% CI −0.49 to 1.00). In the retention session a statistically significant difference in medical management was seen with the use of the checklists, with mean (SD) adherence to key processes 36% (26%) without the checklist vs. 59% (30%) with the checklist ($p = 0.04$, Hedges' $g = 0.79$, 95% CI 0.02–1.56) (Fig. 2).

Following the training workshop the four raters' inter-rater reliability using the TEAM instrument was excellent (intraclass correlation coefficient 0.89) [29]. As shown in Fig. 3, teamwork did not vary with session timing or availability of a checklist. At the initial session, mean (SD) teamwork was rated at 68% (9%) without the checklist available vs. 69% (11%), where the checklist was available ($p = 0.94$, Hedges' $g = 0.03$, 95% CI −0.71 to 0.77). At the second retention session mean (SD) teamwork was rated at 69% (10%) without the checklist vs. 65% (13%) with the checklist ($p = 0.36$, Hedges' $g = -0.34$, 95% CI −1.08 to 0.41). The inter-rater reliability for the study videos was fair (intraclass correlation coefficient 0.47) [29].

Response rate to the exit questionnaire was 100%. Some participant reactions to the simulations are illustrated in Fig. 4. Ninety-one percent of respondents agreed or strongly agreed that they would attend an annual simulation-based team training session. Ninety-four percent of respondents agreed or strongly agreed that every ambulatory/daycase surgical facility should have checklists immediately available.

With respect to the observed use of the checklists, the team elected to use the checklists in 28 (100%) of events where they were available. When they were used, the correct checklist was selected on only 21 (75%) of occasions, meaning that in scenarios where the checklists were available, the team was using the wrong checklist a quarter of the time. The checklist reader was most commonly a nurse (on 17 (61%) occasions), followed by the surgeon (on 9 (28%) occasions) and was the anaesthetist least frequently (2 (7%) occasions). In 17 (61%) of events where the checklist was available, the checklist reader role migrated between practitioners.

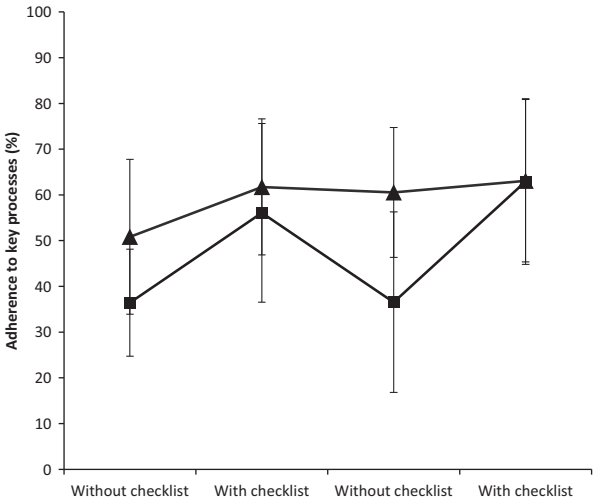

Figure 2 Medical management of scenarios without and with checklists available. Team Emergency Assessment Measure (TEAM) scores are shown for initial (▲) and retention session (■). Values are mean with error bars indicating standard deviation.

Discussion

We introduced critical event checklists in an ambulatory surgical facility and investigated their impact on medical management and teamwork in simulated operating room crises during two sessions separated by several months. In the initial session, we did not demonstrate a difference in medical management when the checklists were available. In a retention session, medical management was significantly improved by the presence of a checklist, but primarily because the teams performed worse without the checklists available. Teamwork did not appear to be affected by checklist use in either session.

This study was conducted in a surgical daycare facility, where critical events are rare and staff are

Table 1 Participant demographics of the seven inter-professional teams (comprising an anaesthetist, a surgeon and three nurses) involved in the simulated emergency scenarios. Values are number or mean.

	Team 1	Team 2	Team 3	Team 4	Team 5	Team 6	Team 7
Sex; male:female ratio	3:2	1:4	1:4	1:4	2:3	0:5	1:4
Age; years	53	44	42	50	46	39	45
Clinical experience as staff; years	25	16	14	15	15	8	18
Team members with previous experience of simulation-based team training	1	2	2	1	0	1	2

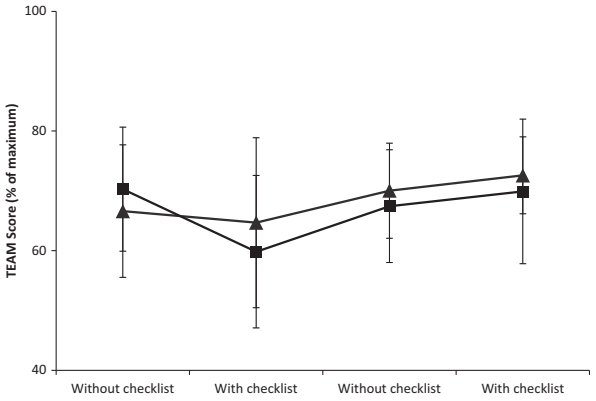

Figure 3 Quality of teamwork without and with checklists available. Team Emergency Assessment Measure (TEAM) scores are shown for initial (▲) and retention session (■). Values are mean with error bars indicating standard deviation.

infrequently involved in their management. By this rationale, we anticipated that the availability of a checklist acting as a cognitive aid would be advantageous. Our failure to demonstrate a meaningful impact of checklists was potentially troubling. In the initial study of the same checklists [11], two operating room teams managed eight simulation-based scenarios, four with and four without a checklist; the authors reported that teams were six times more likely to adhere to key processes when using the checklist. Similar to our study, teams in that study did not have previous exposure or practice using the checklists. Another study showed a 17% improvement in completing key processes when operating theatre teams used these checklists [10]. In our retention session, we observed a 23% mean improvement in adherence for teams performing with the checklists. This was the only comparison which reached statistical significance and the apparent positive impact was in fact secondary to teams underperforming without the checklists. We theorised that we had possibly induced a reliance on the checklists such that without them, the team was unable to perform to the level it had before they were introduced to the checklists. Similar to the previous study of these checklists [10], we used a multiple-scenario format. We randomised the order in which we delivered the scenarios to minimise any scenario-specific impact that might be a confounding factor in team performance but there remains the possibility that the complex interplay of team, timing, checklist availability and scenario resulted in too many combinations of variables that ultimately diluted or distorted our results. Other confounding factors may have diluted the observed

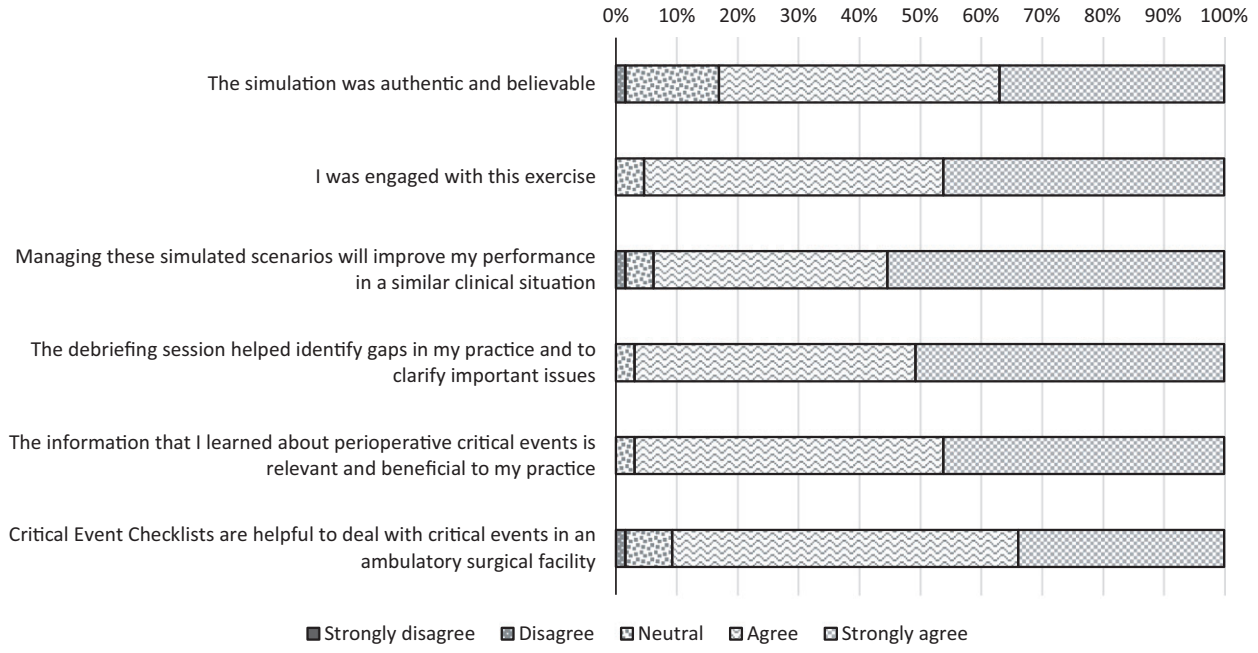

Figure 4 Participant reactions. A selection of questions and responses from the exit questionnaire

effect size. For example, we noted that a strong leader transformed the team performance regardless of the presence of a checklist. This is unsurprising as we know that leadership matters [30], but it is difficult to control for this in study design or adjust for this in the analysis. Despite the homogenous teams (in terms of baseline demographics), it is possible that some teams were more familiar with each other. We assert that in an ambulatory surgical facility of this nature, staff rotate such that all staff work with each other at some stage, so, if present at all, this would be a minor effect.

We reflected on our checklist orientation process for participants. Lack of familiarity with cognitive aids and checklists has been shown to be problematic [1, 31]. Although we provided the participants with an initial orientation to the checklists, we did not afford them much hands-on time with the binder before their scenarios, assuming that the individual checklists (once correctly selected) were self-explanatory. If familiarity was a factor, our findings might underline the importance of a comprehensive orientation during the implementation phase of checklists in the operating theatre. Regarding our participants choosing the wrong checklist a quarter of the time, we emphasise that a checklist is not a substitute for knowledge or judgement at diagnostics. Our study is not unique in showing that the wrong checklist can be selected [17]. We acknowledge that occasionally the limitations of simulation can result in diagnostic confusion or uncertainty but we balance that against the fact that a similar uncertainly may also arise in the real clinical arena.

To lend added insight beyond the other studies of these checklists [10, 11], we examined the non-technical aspects of performance using TEAM scores [25, 26]. We chose to study non-technical team performance because issues related to team function have been shown to affect negatively surgical performance and outcomes [32, 33] and cognitive aids are known to improve non-technical skills in the operating theatre [12, 34]. This is intuitive as removing the cognitive load of remembering stepwise medical management might mean that team members can give more attention to optimising teamwork. However, we did not demonstrate that the checklists improved teamwork. Although it may be that use of a checklist in our simulation scenarios had no effect on teamwork, an alternative explanation is that the checklist did improve team function, but our outcome measure (the combination of tool and raters) was not sufficiently sensitive to demonstrate the difference. A contributory factor may be the poorer inter-rater reliability for the rating of study videos than the training videos. The raters noted that the clinical intensity of the study scenarios was greater than in the sample scenarios used in the training workshop and that the study videos were consequently more difficult to rate. Alternatively, teamwork may not be the construct that checklists are meant to effect, and thus we might not expect any measure of team performance to be sensitive to checklist use, which is largely for enhancing medical management. Marshall et al. investigated the effect of cognitive aids on teamwork and showed that the design of the cognitive aid is influential and if complex and branched, cognitive aids were not associated with improvements in teamwork [35]. In our case, however, the checklists were of a linear design (not branched).

We were interested to observe the behaviour of the teams with respect to leadership, reading of the checklist and resource utilisation. The leader is conditioned to stand-back and be hands-off while directing the team management. The checklist reader (most commonly a nurse) was occupied reading the checklist and monitoring completion of tasks. That effectively eliminated two practitioners from contributing physical tasks to the intense clinical scenario where there are multiple urgent contemporaneous tasks. The nature of lean teams in a surgical daycare facility is such that the proportional reduction in active team members imposed by having a leader and a checklist reader out of action, stretched the team noticeably. We observed that as the teams felt this strain, they would switch the checklist reader to free the original reader to perform tasks. This only injected further confusion and potentially derailed the management. This gives weight to the argument that during implementation of checklists outside of teaching hospitals, where teams are lean, explicit instruction should be given to minimise this effect, including that the leader be more 'hands-on' and contribute physical tasks. This phenomenon of the variable impact of a code-reader has been investigated in similar contexts [14, 36].

The use of critical events checklists in an ambulatory surgical daycare had no effect on medical management or teamwork during simulated operating theatre crises. Our data align with other work suggesting checklists are not a panacea and their positive impact can only be assured where careful consideration is given to context-specific implementation and orientation [1]. We recommend the further investigation of those context-specific factors that maximise the benefit of checklists in the peri-operative period. Our findings would also support the simulation-based pilot of checklists to inform and optimise their effective implementation before use in the clinical environment.

Competing interests

This work was supported by funding from The Canadian Medical Protective Agency (CMPA), Ontario Simulation Network, SIM-one and an Ontario Ministry of Health and Long Term Care (MOHLTC) Innovation Fund Grant. No other external funding or competing interests declared.

References

1. Marshall S. The use of cognitive aids during emergencies in anesthesia: a review of the literature. *Anesthesia and Analgesia* 2013; **117**: 1162–71.
2. *The WHO Surgical Safety Checklist*. Geneva: World Health Organization, 2015.
3. The Association of Anaesthetists of Great Britain and Ireland. *Malignant Hyperthermia Crisis Safety Guideline*. London: The Association of Anaesthetists of Great Britain and Ireland, 2015.
4. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *New England Journal of Medicine* 2009; **360**: 491–9.
5. Urbach DR, Govindarajan A, Saskin R, Wilton AS, Baxter NN. Introduction of surgical safety checklists in Ontario, Canada. *New England Journal of Medicine* 2014; **370**: 1029–38.
6. Harrison TK, Manser T, Howard SK, Gaba DM. Use of cognitive aids in a simulated anesthetic crisis. *Anesthesia and Analgesia* 2006; **103**: 551–6.
7. Low D, Clark N, Soar J, et al. A randomised control trial to determine if use of the iResus(c) application on a smart phone improves the performance of an advanced life support provider in a simulated medical emergency. *Anaesthesia* 2011; **66**: 255–62.
8. McEvoy MD, Hand WR, Stoll WD, Furse CM, Nietert PJ. Adherence to guidelines for the management of local anesthetic systemic toxicity is improved by an electronic decision support tool and designated ''Reader''. *Regional Anesthesia and Pain Medicine* 2014; **39**: 299–305.
9. Neal JM, Hsiung RL, Mulroy MF, Halpern BB, Dragnich AD, Slee AE. ASRA checklist improves trainee performance during a simulated episode of local anesthetic systemic toxicity. *Regional Anesthesia and Pain Medicine* 2012; **37**: 8–15.
10. Arriaga AF, Bader AM, Wong JM, et al. Simulation-based trial of surgical-crisis checklists. *New England Journal of Medicine* 2013; **368**: 246–53.
11. Ziewacz EJ, Arriaga AF, Bader AM, et al. Crisis checklists for the operating room: development and pilot testing. *Journal of the American College of Surgeons* 2011; **213**: 212–7, e10.
12. Marshall SD, Mehra R. The effects of a displayed cognitive aid on non-technical skills in a simulated 'can't intubate, can't oxygenate' crisis. *Anaesthesia* 2014; **69**: 669–77.
13. Bould MD, Hayter MA, Campbell DM, Chandra DB, Joo HS, Naik VN. Cognitive aid for neonatal resuscitation: a prospective single-blinded randomized controlled trial. *British Journal of Anaesthesia* 2009; **103**: 570–5.
14. Burden AR, Carr ZJ, Staman GW, Littman JJ, Torjman MC. Does every code need a ''reader?'' improvement of rare event management with a cognitive aid ''reader'' during a simulated emergency: a pilot study. *Simulation in Healthcare* 2012; **7**: 1–9.
15. Mills PD, DeRosier JM, Neily J, McKnight SD, Weeks WB, Bagian JP. A cognitive aid for cardiac arrest: you can't use it if you don't know about it. *Joint Commission Journal on Quality and Saftey* 2004; **30**: 488–96.
16. Neily J, DeRosier JM, Mills PD, Bishop MJ, Weeks WB, Bagian JP. Awareness and use of a cognitive aid for anesthesiology. *Joint Commission Journal on Quality and Saftey* 2007; **33**: 502–11.
17. Nelson KL, Shilkofski NA, Haggerty JA, Saliski M, Hunt EA. The use of cognitive AIDS during simulated pediatric cardiopulmonary arrests. *Simulation in Healthcare* 2008; **3**: 138–45.
18. Catchpole K, Russ S. The problem with checklists. *British Medical Journal Quality and Safety* 2015; **24**: 545–9.
19. Vila Jr, Soto R, Cantor AB, Mackey D. Comparative outcomes analysis of procedures performed in physician offices and ambulatory surgery centers. *Archives of Surgery* 2003; **138**: 991–5.
20. Kazdin AE. *Single-case Research Designs: Methods for Clinical and Applied Settings*, 2nd edn. New York: Oxford University Press, 2010.
21. Janosky J. *An Overview of the Analysis of a Single-subject Design with Recommendations*. Chicago, IL: Annual Meeting of the American Educational Research Association, 1991.
22. Engel R, Schutt R. *The Practice of Research in Social Work*, 2nd edn. Thousand Oaks, CA: Sage Publications, 2008.
23. Weiner IS, Eisen RG. Clinical research: the case study and single-subject designs. *Journal of Allied Health* 1985; **14**: 191–201.
24. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. *Academic Medicine* 2014; **89**: 387–92.
25. Cooper S, Cant R, Porter J, et al. Rating medical emergency teamwork performance: development of the Team Emergency Assessment Measure (TEAM). *Resuscitation* 2010; **81**: 446–52.
26. Cooper SJ, Cant RP. Measuring non-technical skills of medical emergency teams: an update on the validity and reliability of the Team Emergency Assessment Measure. *Resuscitation* 2014; **85**: 31–3.
27. Kirkpatrick DL. *Evaluating Training Programs*. San Francisco, BK Berrett: Koehler Publishers Inc, 1998.
28. Eppich W, Cheng A. Promoting Excellence and Reflective Learning in Simulation (PEARLS): development and rationale for a blended approach to health care simulation debriefing. *Simulation in Healthcare* 2015; **10**: 106–15.

29. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* 1994; **6**: 284–90.
30. Salas E, Rosen M, King H. Managing teams managing crises: principles of teamwork to improve patient safety in the emergency room and beyond. *Theoretical Issues in Ergonomics Science* 2007; **8**: 381–94.
31. Goldhaber-Fiebert SN, Howard SK. Implementing emergency manuals: can cognitive aids help translate best practices for patient care during acute events? *Anesthesia and Analgesia* 2013; **117**: 1149–61.
32. Joint Commission. Communication: a critical component in delivering quality care. *Joint Commission: The Source* 2005; **3**: 1–10.
33. Lingard L, Espin S, Whyte S, et al. Communication failures in the operating room: an observational classification of recurrent types and effects. *Quality and Safety in Health Care* 2004; **13**: 330–4.
34. Russ S, Rout S, Sevdalis N, Moorthy K, Darzi A, Vincent C. Do safety checklists improve teamwork and communication in the operating room? A systematic review. *Annals of Surgery* 2013; **258**: 856–71.
35. Marshall SD, Sanderson P, McIntosh CA, Kolawole H. The effect of two cognitive aid designs on team functioning during intra-operative anaphylaxis emergencies: a multi-centre simulation study. *Anaesthesia* 2016; **71**: 389–404.
36. Power DJ, Boet S, Bould MD. Code Reader: a novel concept that warrants more research. *Simulation in Healthcare* 2012; **7**: 136–7; author reply 137.

Appendix

Other co-investigators in this study were: H. Carnahan, Dean, School of Human Kinetics and Recreation, Memorial University of Newfoundland, St. John's, Newfoundland, Canada; L. Murnaghan, Assistant Professor and Staff Orthopedic Surgeon, The Hospital for Sick Children, Toronto, Ontario, Canada; J. Semple, Professor and Surgeon-in-Chief, Women's College Hospital, Toronto, Ontario, Canada; and M. L. Ling, Senior Nurse, Women's College Hospital, Toronto, Ontario, Canada.

