Simulation-Based Trial of Surgical-Crisis Checklists


ABSTRACT

BACKGROUND

Operating-room crises (e.g., cardiac arrest and massive hemorrhage) are common events in large hospitals but can be rare for individual clinicians. Successful management is difficult and complex. We sought to evaluate a tool to improve adherence to evidence-based best practices during such events.

METHODS

Operating-room teams from three institutions (one academic medical center and two community hospitals) participated in a series of surgical-crisis scenarios in a simulated operating room. Each team was randomly assigned to manage half the scenarios with a set of crisis checklists and the remaining scenarios from memory alone. The primary outcome measure was failure to adhere to critical processes of care. Participants were also surveyed regarding their perceptions of the usefulness and clinical relevance of the checklists.

RESULTS

A total of 17 operating-room teams participated in 106 simulated surgical-crisis scenarios. Failure to adhere to lifesaving processes of care was less common during simulations when checklists were available (6% of steps missed when checklists were available vs. 23% when they were unavailable, P<0.001). The results were similar in a multivariate model that accounted for clustering within teams, with adjustment for institution, scenario, and learning and fatigue effects (adjusted relative risk, 0.28; 95% confidence interval, 0.18 to 0.42; P<0.001). Every team performed better when the crisis checklists were available than when they were not. A total of 97% of the participants reported that if one of these crises occurred while they were undergoing an operation, they would want the checklist used.

CONCLUSIONS

In a high-fidelity simulation study, checklist use was associated with significant improvement in the management of operating-room crises. These findings suggest that checklists for use during operating-room crises have the potential to improve surgical care. (Funded by the Agency for Healthcare Research and Quality.)
Operating-room crises (e.g., massive hemorrhage and cardiac arrest) are high-risk, stressful events that require rapid and coordinated care in a time-critical setting. The reported incidence may be rare for an individual practitioner, but the aggregate incidence for a hospital with 10,000 operations a year is estimated to be approximately 145 such events annually. These are situations in which the way the team cares for a patient will make the difference between life and death. Failure to effectively manage life-threatening complications in surgical patients has been recognized as the largest source of variation in surgical mortality among hospitals.

Checklists have long been accepted in other high-risk industries (e.g., aviation and nuclear power) as a tool to aid performance during rare and unpredictable critical events. In the field of medicine, the use of surgical safety checklists during routine operative care has been associated with significant reductions in morbidity and mortality and is rapidly becoming a standard of care. Crisis-related cognitive aids (of which checklists are a subset) and manuals exist for the operating room. However, the effect of crisis checklists on performance during intraoperative crises has been largely untested, particularly in studies that include seasoned clinicians (not just trainees) or that include more than one institution.

We therefore sought to study the effect of a crisis-checklist intervention. Owing to the relative infrequency of these events, a live clinical trial was not feasible. We used high-fidelity medical simulation to facilitate a structured observation of these unpredictable events. In previous work, we developed a set of intraoperative critical-event checklists and confirmed their usability in a pilot study with two operating-room teams. We hypothesized that crisis checklists would significantly improve adherence to best practices in a randomized, controlled trial involving teams from both academic and community hospitals.

**METHODS**

**STUDY PARTICIPANTS**
Participants were recruited from three hospitals (one academic medical center and two community hospitals) in the Boston area over a 17-month period (August 2010 through December 2011). Teams consisted of anesthesia staff (attending physicians, residents, and certified registered nurse anesthetists), operating-room nurses, surgical technologists, and a mock surgeon participant (except in a small number of cases in which volunteer surgeons and surgical residents were available). Each team spent a 6-hour day in a high-fidelity simulated operating room where they were presented with a series of crisis scenarios and tested for their adherence to critical, evidence-based practices. In half the scenarios, randomly chosen, the team had access to a previously developed set of checklists for crisis events. In the other half, they worked from memory, as in usual care.

Staff members were enrolled by means of sign-up sheets and random selection of staff members who were already scheduled to work on the study dates. The hospital departments allowed staff members to be scheduled to the simulator in lieu of a standard work day. During the initial study sessions, there was limited participation by surgeons and surgical residents in the processes of care being studied. Given this factor as well as scheduling and other difficulties, volunteer surgeons were welcome, but their participation was not mandatory. All anesthesia staff (attending physicians, residents, and certified registered nurse anesthetists) were required by their departmental or hospital policy to hold current certification in advanced cardiac life support. All study participants attended one and only one study session. Approval from the local institutional review board was obtained before the start of the study, and all study participants gave written informed consent.

**STUDY DESIGN**
A schematic of team participation is provided in the Supplementary Appendix, available with the full text of this article at NEJM.org. Each team was exposed to a series of simulated intraoperative crises (including air embolism, anaphylaxis,
asystolic cardiac arrest, hemorrhage followed by ventricular fibrillation, malignant hyperthermia, unexplained hypotension and hypoxemia followed by unstable bradycardia, and unstable tachycardia. Each team was randomly assigned to manage half the scenarios with the checklists available and the remaining scenarios by memory alone. Members of the teams participating in the pilot study (which included surgeons and surgical residents) also underwent randomization.

The checklists were provided in booklet form in two locations: adjacent to the anesthesia machine and adjacent to the circulating nurse’s work area. A detailed description of the development of the crisis checklists was published previously,26 and a brief description is provided in the Supplementary Appendix, along with the surgical safety checklists. All authors vouch for the completeness and accuracy of the data and analyses as presented and for the fidelity of the study to the protocol, available at NEJM.org.

The primary outcome measure was failure to adhere to life-saving processes of care for each crisis. There were a total of 47 key processes across the scenarios tested (see the Supplementary Appendix). These were derived from evidence-based guidelines and scored in a binary fashion (yes or no). All simulation sessions were recorded as multiscreen synchronized videos, and three physician reviewers, all of whom were authors, observed and scored the sessions. For the first 200 key processes, the reviewers worked in pairs, and interrater reliability was assessed. A kappa score greater than 0.90 for each reviewer pair was established before the remaining processes were reviewed by a single physician reviewer. Any case of disagreement or uncertainty among reviewers regarding adherence by the team to a key process was decided by means of expert review (assessment of the video by a senior surgeon, a senior anesthesiologist, or a senior physician who was an expert in guidelines for advanced cardiac life support). A random 15% sample of the data was rereviewed by an outside physician reviewer who was unaware of both the study design and the hypothesis being tested. Study participants were surveyed regarding their perceptions of the usefulness and clinical relevance of the checklists, with the use of Likert scales ranging from 1 (disagree strongly) to 5 (agree strongly). Items from the full survey are provided in Tables S1 through S4 in the Supplementary Appendix.

### Table 1. Professional Characteristics of the Participants.

<table>
<thead>
<tr>
<th>Position</th>
<th>Participants (N = 67)</th>
<th>Years of Experience in Specialty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. (%)</td>
<td>1 to 5</td>
</tr>
<tr>
<td>Anesthesia attending physician</td>
<td>17 (25)</td>
<td>0</td>
</tr>
<tr>
<td>Surgical attending physician</td>
<td>2 (3)</td>
<td>0</td>
</tr>
<tr>
<td>Anesthesia resident*</td>
<td>10 (15)</td>
<td>0</td>
</tr>
<tr>
<td>Surgical resident*</td>
<td>2 (3)</td>
<td>0</td>
</tr>
<tr>
<td>Operating-room nurse</td>
<td>20 (30)</td>
<td>0</td>
</tr>
<tr>
<td>Surgical technologist</td>
<td>9 (13)</td>
<td>0</td>
</tr>
<tr>
<td>Certified registered nurse anesthetist</td>
<td>7 (10)</td>
<td>29</td>
</tr>
</tbody>
</table>

* One anesthesia resident who participated was a first-year anesthesia resident at the end of the first year of clinical anesthesia training (second postgraduate year). The remaining anesthesia residents were in their second or third year of clinical anesthesia training, and the surgical residents were in their second or third postgraduate year of training.

### Statistical Analysis

Data were analyzed with the use of SAS software, version 9.3 (SAS Institute). All reported P values are two-sided, and P values of less than 0.05 were considered to indicate statistical significance. Agreement between two physician reviewers was assessed with the use of Cohen’s kappa. Bonferroni adjustment of P values was performed to account for the number of P values reported and the analysis of the data from the two study dates reviewed in the pilot usability trial. Multivariate relative-risk regression for dichotomous outcomes was used to compare failure rates with and with-
out the checklist while accounting for clustering of results within a team, with adjustment for institution, scenario, and time of day. Time of day was included in the model to address the possible confounding effects of learning the simulator processes or becoming fatigued over the course of the day. An interaction term for institution type was included in a second model to assess whether the effect of the intervention differed according to whether it was conducted in an academic medical center or a community hospital. A stratified analysis was performed to compare the effect of the checklist according to scenario type.

Additional post hoc analyses were done, all of which are reported here. A secondary analysis was performed to compare the effect of the checklist according to whether a surgeon was present or absent. Additional secondary analyses were performed to adjust for whether the team included an anesthesia resident or a certified registered nurse anesthetist, whether a senior (third-year) clinical anesthesiology resident was present or absent, whether a senior anesthesiologist or nurse (defined as a person with ≥15 years of experience in his or her specialty) was present or absent, and whether the team was randomly assigned first to scenarios in which checklists were available or scenarios in which they were unavailable.

**RESULTS**

**CHARACTERISTICS OF THE PARTICIPANTS**

A total of 17 operating-room teams participated in 106 simulated surgical-crisis scenarios. The professional characteristics of the participants are shown in Table 1. Participants had a wide range of years of experience in their specialty. Attending surgeons (together with surgical residents) were available for two of the study dates. The remaining teams contained mock surgeons and surgical assistants (i.e., senior surgical residents and scrub nurses, respectively, who attended to the operative field without participating in decision making or completing surveys; these stand-in staff members were not counted as participants).

**VIDEO ANALYSIS**

Each pair of independent reviewers had excellent interrater reliability (kappa ≥0.92 for all reviewer pairs). Given that the key processes were hard end points (e.g., calling for help within 1 minute after ventricular fibrillation), consensus with the expert reviewer was easily achieved on video replay for any instance of initial disagreement or uncertainty among the physician reviewers regarding adherence to a key process. Of the 750 key processes reviewed, 10 (1%) required expert review for this purpose, and there was full agreement among all reviewers immediately after video replay. For the 15% of the data randomly selected for rereview by an outside physician, there was excellent interrater reliability with respect to the outcome assessments (kappa = 0.91).

**ANALYSES OF CHECKLIST USE**

Checklist use during operating-room crises resulted in nearly a 75% reduction in failure to adhere to critical steps in management. Of 371 critical steps in the management of surgical crises, 24 (6%) were missed when the checklists were available, as compared with 89 of 379 steps (23%) missed when the checklists were not available. I bars indicate 95% confidence intervals.

![Figure 1. Association between Use or Nonuse of Operating-Room Crisis Checklists and Failure to Adhere to Critical Steps in Management.](image-url)

The use of checklists during operating-room crises resulted in nearly a 75% reduction in failure to adhere to critical steps in management. Of 371 critical steps in the management of surgical crises, 24 (6%) were missed when the checklists were available, compared with 89 of 379 steps (23%) missed when the checklists were not available. I bars indicate 95% confidence intervals.
The significant effect of checklist use was also seen when we stratified the results according to scenario (Table 2). The scenarios were grouped into three categories to provide samples large enough for analysis: scenarios that were directly related to algorithms for advanced cardiac life support (asystolic cardiac arrest, ventricular fibrillation, and unstable tachycardia); ACLS scenarios preceded by a hemodynamically unstable condition included clinically significant hypoxemia and hypotension followed by unstable bradycardia, and hemorrhage followed by ventricular fibrillation; and other crisis scenarios included malignant hyperthermia, anaphylaxis, hemorrhage, and air embolism. The failure rate was calculated as the number of critical steps that were not adhered to in the management of the scenario.

<table>
<thead>
<tr>
<th>Scenario Type*</th>
<th>Failure Rate†</th>
<th>P Value‡</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With Checklists</td>
<td>Without Checklists</td>
</tr>
<tr>
<td></td>
<td>no. /total no. (%)</td>
<td>no. /total no. (%)</td>
</tr>
<tr>
<td>ACLS scenario</td>
<td>7/100 (7)</td>
<td>15/89 (17)</td>
</tr>
<tr>
<td>ACLS scenario preceded by hemodynamically unstable condition</td>
<td>14/154 (9)</td>
<td>46/172 (27)</td>
</tr>
<tr>
<td>Other crisis scenario</td>
<td>3/117 (3)</td>
<td>28/118 (24)</td>
</tr>
</tbody>
</table>

* Scenario types were as follows: advanced cardiac life support (ACLS) included asystolic cardiac arrest, ventricular fibrillation, and unstable tachycardia; ACLS scenarios preceded by a hemodynamically unstable condition included clinically significant hypoxemia and hypotension followed by unstable bradycardia, and hemorrhage followed by ventricular fibrillation; and other crisis scenarios included malignant hyperthermia, anaphylaxis, hemorrhage, and air embolism.
† The failure rate was calculated as the number of critical steps that were not adhered to in the management of the scenario.
‡ P values were calculated in a model that accounted for clustering by team, with adjustment for time of day and institution.

The examples provided in Table 3, which involve failure by highly qualified teams to adhere to recommended procedures because of key details missed at time-critical moments, are probably not unique to the teams studied. Failures such as these were observed for every type of scenario. Analogous failures occurred in the aviation industry in the era before checklist use was widely accepted for routine and critical operations. Current aviation-accident reports have been

**DISCUSSION**

Critical events in high-risk professions call for rapid, coordinated, and accurate maneuvers, despite stress and increased task load, as a requisite for successful management. In this high-fidelity simulation-based study, we found that the use of crisis checklists was associated with a significant improvement in adherence to recommended procedures for the most common intraoperative emergencies. After participation, 97% of the participants agreed that they would use the checklists if presented with these operative emergencies in real life. A total of 97% of the participants agreed (i.e., gave a score of 4 or higher) with the statement “If I were having an operation and experienced this intraoperative emergency, I would want the checklist to be used.” All participants rated the overall quality of the session as above average or excellent (score of 4 or 5, respectively).
noted to include checklist solutions as part of their recommendations.\textsuperscript{11,28} Team training has long been embraced by the aviation industry and other high-reliability organizations.\textsuperscript{29-31} The integration of checklist use with simultaneous team training may augment our observed effect. There are an increasing number of multicenter programs that focus on both the training of operating-room teams and the use of surgical safety checklists.\textsuperscript{15}

A limitation of the study may be the absence of surgeons as participants in most of the simulations. Their participation would have been preferable, but it was difficult to enlist volunteers. However, we found no evidence that the presence of a surgeon reduced the benefit of the checklist intervention. The key processes tracked for this study, which were developed by a multidisciplinary panel that included surgeons, were primarily the responsibility of nursing and anesthesiology staff. Although the number of teams participating in the study was too small to draw any conclusion about a difference in the effect of the intervention according to whether a surgeon was present or absent, the stratified analysis suggested that, if there was any difference, the presence of a surgeon improved the benefit of the intervention. We believe many clinicians would perceive the value of having checklists available for the management of events in the operating room that are relatively infrequent, often unpredictable, and capable of happening at any time during any operation.

Our findings should be interpreted in the context of the study design; we studied checklists in a simulated operating room rather than in actual operating rooms with real patients. Accepting the results of interventions tested by means of simulation invariably carries risks. Events undoubtedly occur in more varied circumstances than we could simulate, and it is unclear whether the effect on adherence to recommended processes would increase or decrease. High-fidelity simulation has become increasingly accepted in medicine as a means of training and evaluation.\textsuperscript{26,32} Lessons from other fields — such as aviation and nuclear power — indicate that testing in a well-structured simulation setting can be an efficient aid in assessing the value of safety protocols and that deferring the adoption of such protocols because of the infeasibility of large-scale trials is imprudent.

The results of this study suggest that hospitals and ambulatory surgical centers should consider implementation of checklists to increase the safety of surgical care. Monitoring

| Table 3. Examples of the Effect of Crisis Checklists on Adherence to Critical Processes of Care. |
|---|---|---|
| **Example** | **With Checklists** | **Without Checklists** |
| 1 | Unstable tachycardia: synchronized cardioversion delivered promptly and all shocks synchronized | Unstable bradycardia: “The pacer is not working”; >10-min delay to transcutaneous pacing because the setting selected by the provider was below the energy level necessary to enable pacing of the heart |
| 2 | Unstable bradycardia: prompt transcutaneous pacing | Unstable tachycardia: not a single shock synchronized |
| 3 | Malignant hyperthermia: all seven critical care processes completed (including dantrolene administration, cooling, treatment of hyperkalemia, and discontinuation of volatile anesthetic agents) | Anaphylaxis: three of six critical care processes not completed (e.g., help never called for and insufficient fluid resuscitation) |
| 4 | Asystole: chest compressions initiated within 17 seconds | Ventricular fibrillation: >1.5 min before chest compressions started |

| Table 4. Participants’ Perceptions of Crisis Checklists, with Responses across All Checklist Scenarios.\textsuperscript{a} |
|---|---|
| **Survey Statement** | **Response Score** |
| The checklist helped me feel better prepared during the emergency scenario | 4.4±0.81 |
| The checklist was easy to use | 4.3±0.84 |
| I would use this checklist if I were presented with this operative emergency in real life | 4.5±0.76 |
| If I were having an operation and experienced this intraoperative emergency, I would want the checklist to be used | 4.7±0.60 |

\textsuperscript{a} Plus–minus values are means ±SD. Data included 196 responses from 67 participants. Response scores were on a Likert scale and ranged from 1 (disagree strongly) to 5 (agree strongly).
during implementation would be necessary. Future work should include determining the best medium and user interface (e.g., paper or electronic, and if electronic, a laptop or tablet version), a method to update the checklists as evidence evolves, and appropriate mechanisms for training and implementation. The persons implementing the checklists must customize them to their institution. A shift in the medical culture may be necessary if health care providers are expected to pull out a cognitive aid during an intraoperative (or any other) emergency. As noted above, checklists and other cognitive aids for managing critical events both inside and outside the operating room have existed for decades.6-10 Yet there is no standard for crisis checklists of any kind to be available or used in medical care.

Experts have long recognized the potential for human fallibility in complex systems.24-26 As Berry points out, it has been nearly 100 years since the surgeon W. Wayne Babcock called for emergency protocols to be rehearsed and “posted on the walls of every operating room” after observing that readily accessible cognitive aids would have improved the treatment of his own patients.37,38 Although some have advocated not waiting for a randomized, controlled trial in the case of interventions with strong face validity for improved outcomes,39 we believe that testing is valuable and necessary and that, with the use of simulation, it can be done efficiently. Our study shows a significant and substantial value in the use of carefully designed crisis checklists for the operating room.

The views expressed in this article are those of the authors and do not necessarily represent the official views of the Agency for Healthcare Research and Quality. Supported by a grant from the Agency for Healthcare Research and Quality (1R18 HS018537-01).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Yelena Kuznetsov, Olivia Nelson, Joan Vitello, Ashish Jha, Elizabeth Edmondson, Elizabeth Morse, Paul Preston, Connie Lopez, Stephen Poole, Andrew Camerato, Gloria Hicks, Seth Jones, Robin Smeland-Wagman, Richard Kaufman, Marc Rubin, Mitchell Scott Rein, Glynnie David Stanley, Charlene Scarpa, Jacob Joffe, Susan Duffy Smith, Barbara Jo Ryan, Anna Dasilva, Mark Hershey, Ann Mullen, Steven Nelson, Sarah Peyre, Rebecca Kim, Dante Foster, Alvin Kwok, Aini Kallase, William Perry, Elizabeth Guzy, Shawn Rangell, Sunil Eappen, Luke Funk, Marcus Semel, Marta McCrum, Thomas Weiser, Alex Haynes, Christopher Dodgion, Lyn Hu, Samuel Rice-Townsend, Luise Pernar, Jonathan Spector, Susan Block, Robin Reenberg, Grace Galvin, Vikas Yadav, Mazza Shaikh, Kendra Bowman, James Sachetta, Beenweatte Baldeo, Michael Trioli, Emily Hinchev, Nicole Kissane, Steven Yule, Thomas Lennom, Robert Slane, and the members of the Intraoperative Critical Event Working Group for their expertise and help with this project.

REFERENCES


Copyright © 2013 Massachusetts Medical Society.

ICMJE SEEKING TWO NEW MEMBER JOURNALS
The International Committee of Medical Journal Editors (ICMJE) is seeking two new member journals to be represented by their editors-in-chief. Information about the ICMJE is available at www.icmje.org. Candidate journals should meet the following criteria:

- be a peer-reviewed general medical journal that publishes original research involving humans
- have a governance structure that ensures editorial independence
- have an editor with experience in the position who expects to continue in the position for at least another 3 years
- be financially able to support the editor’s participation in ICMJE activities

In considering candidates, the ICMJE may seek to improve the balance of geographic areas and publishing models among its membership.

To apply, editors-in-chief of interested journals should submit the following materials to the ICMJE (at icmje@acponline.org):

- brief curriculum vitae
- cover letter describing the journal, including but not necessarily limited to details of the journal’s history, sponsor or publisher, governance structure, publishing model (e.g., subscription, author-pays open access), target audience, print circulation and online traffic, number of manuscript submissions per year, processes used to select material for publication, acceptance rate, databases where indexed, website address, and guidelines for authors
- statement on how the journal might benefit from ICMJE membership and how the ICMJE might benefit from the journal’s membership (should not exceed 1000 words)

The deadline for applications is January 31, 2013.