The Impact of Checklists on Inpatient Safety Outcomes: A Systematic Review of Randomized Controlled Trials

Jamie M Boyd, BA1-3, Guosong Wu, MSc1, Henry T Stelfox, MD, PhD1-4*

1Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada; 2O’Brien Institute for Public Health, University of Calgary, Calgary, Alberta, Canada; 3Department of Critical Care Medicine, University of Calgary, Calgary, Alberta, Canada; 4Alberta Health Services, Alberta, Canada.

BACKGROUND: Systematic reviews of non-randomized controlled trials (RCTs) suggest that using a checklist results in fewer medical errors and adverse events, but these evaluations are at risk of bias.

OBJECTIVE: To conduct a systematic review of RCTs of checklists to determine their effectiveness in improving patient safety outcomes in hospitalized patients.

METHODS: Ovid EMBASE, Ovid MEDLINE, PubMed, and the Cochrane Central Register of Controlled Trials were searched from inception until December 8, 2016. The search was restricted to RCTs. Included studies reported patient safety outcomes of a checklist intervention. Data extracted included the study characteristics, setting, population, intervention, outcomes measures, and sample size.

MEASUREMENTS AND MAIN RESULTS: 11,225 citations were identified, of which 9 (16,987 patients) satisfied the inclusion criteria. Citations reported evaluations of checklists designed to improve surgical safety, prescription of medications, heart failure management, pain control, infection control precautions, and physician handover. Studies reported significant reductions in postoperative complications and medication-related problems and improved compliance with evidence-based prescribing of medications, infection control precautions, and patient handover procedures. 30-day mortality was reported in 3 studies and was significantly lower among patients allocated to the checklist group (odds ratio 0.60, 95% confidence interval, 0.41-0.89, \(P = 0.01\), \(I^2 = 0.0\%), \(P = 0.573\)). Methodological quality of the studies was moderate.

CONCLUSION: A small number of citations report RCT evaluations of the impact of checklists on patient safety. There is an urgent need for high-quality evaluations of the effectiveness of patient safety checklists in inpatient healthcare settings to substantiate their perceived benefits. Journal of Hospital Medicine 2017;12:675-682. © 2017 Society of Hospital Medicine

In response to widely publicized reports highlighting the challenges of suboptimal quality of healthcare, improving patient safety has been a leading healthcare initiative for more than 10 years.\(^1\)\(^-\)\(^4\) Numerous strategies to improve patient safety have been proposed,\(^5\)\(^-\)\(^9\) but improvements have been limited, which raises questions about whether the right approaches are being employed.\(^10\)\(^11\)

Checklists have served as a foundation for the standardization and safety of aviation and nuclear power\(^12\)\(^-\)\(^13\) and are advocated as simple and effective instruments for ensuring safe care.\(^2\)\(^4\)\(^15\) Systematic reviews of observational studies suggest that checklists can reduce medical errors and adverse events,\(^15\)\(^-\)\(^19\) but these reviews are at risk of bias due to the limitations of observational methods. Furthermore, discordant results of recent high-profile evaluations of the World Health Organization (WHO) Surgical Safety Checklist highlight the need for checklist evaluations using rigorous study designs.\(^20\)\(^-\)\(^22\) Therefore, we sought to conduct a systematic review of RCTs (randomized controlled trials) to determine whether checklists, as a type of decision-support tool, are effective at improving patient safety outcomes in hospitalized patients.

METHODS

The study protocol was registered with the PROSPERO Register of Systematic Reviews (registration number: CRD42016037441) and developed according to the Preferred Reporting Items in Systematic Reviews and Meta-analyses (PRISMA) statement.\(^23\)

Search Strategy

On December 8, 2016, we systematically searched Ovid MEDLINE, Ovid EMBASE, PubMed, and the Cochrane Central Register of Controlled Trials. The search was performed using no language or publication date restrictions and included 2 groups of terms (key words with similar characteristics): ‘checklists’ and ‘patient outcomes assessment’. We restricted our search to patient outcomes because these are more patient-oriented than the proximal processes of care that may not translate into outcomes. The search was restricted to RCTs using the Cochrane Highly Sensitive Search Strategy for Identifying Randomized Trials from the Cochrane Collaborative.\(^24\) The MEDLINE search strategy is
depicted in Appendix I (Supplementary File 1). Reference lists of included articles were manually searched for additional publications. The search strategy was designed with the help of an information scientist (DL). EndNote X7 (Thomas Reuters, Philadelphia, PA, USA) was the reference software used for the management of citations.

Eligibility Criteria
We selected all studies reporting patient safety outcomes of a checklist intervention, using the following inclusion criteria: 1) acute care hospital inpatient population, 2) checklist intervention, 3) contain a control group (ie, no checklist), 4) report one or more patient safety outcome, as defined by the authors (eg, medical errors, adverse events, mortality), and 5) RCT design. We restricted our focus to inpatient populations given the heterogeneity of illness and patient care between acute and community settings. We defined a checklist as a tool that details the essential steps of a task, requiring the target provider to indicate whether an item was completed or not.1,2 Tools that included only 1 item (eg, electronic prompts) or did not require acknowledgement of the items (eg, guidelines) were excluded. We defined patient safety outcomes as the authors’ definition of patient safety (eg, medical error, adverse event, provider compliance with safety regulations).

Study Selection
Two reviewers (JMB, GW) independently, and in duplicate, reviewed the titles and abstracts of the retrieved citations against the eligibility criteria. The same 2 reviewers subsequently reviewed the full text of relevant articles for inclusion. Eligibility disagreements were resolved by consensus. A Kappa statistic was calculated for reviewer agreement of full-text screening.2,25 Reviewers were not blinded to author or journal names.2,6

Data Extraction
The structured data extraction form was calibrated using the first 2 articles. The 2 reviewers (JMB, GW) independently, and in duplicate, extracted data from included studies on the study characteristics, setting, study population, sample size, intervention used, outcomes examined, analytic method, and study quality. The data extraction form is depicted in Appendix II (Supplementary File 2). Coding discrepancies were resolved by consensus.

Quality Assessment
The 2 reviewers (JMB, GW) extracted data on study quality independently and in duplicate using 2 approaches. First, reviewers assessed study quality using a component method derived from the Cochrane Collaboration criteria.2,4 For each included study, the reviewers documented if the authors had adequately described inclusion/exclusion criteria, randomization, allocation concealment, blinding of participants/outcome assessors, attrition, cross over, baseline characteristics, and power calculation. Second, the reviewers calculated and reported the Jadad score for each included study, a validated assessment scale that assigns points (1 to 5) based on randomization, blinding, and attrition.2,7

Analysis
Owing to the heterogeneity of the data and the small number of studies that satisfied the inclusion criteria, the data were analyzed using guidelines for the narrative synthesis of a systematic review.2,8 Descriptive statistical findings from each included study were reported. The DerSimonian and Laird method for random-effects models was used to calculate a pooled estimate of 30-day all-cause mortality from the raw data available from a subset of studies (number of events, study population).2,29 Stata SE version 13.1 (Stata Corp, LP, College Station, TX) was used to perform the statistical analyses.

RESULTS
The literature search identified 11,225 unique citations from which 83 abstracts were eligible for full-text review. We identified 9 full-text articles for inclusion in the review (Figure 1 [Supplementary File 3]). The main reasons for citation exclusion during the full-text review were that the study design was not an RCT (39%) or there was no checklist intervention (34%). Inter-rater agreement for full-text inclusion was fair (K=0.660, 95% confidence interval[CI],0.414-0.828).

Study Characteristics
Characteristics of the included studies are summarized in Tables 1 and 2. Six of the studies were conducted in at least 1 teaching hospital.2,10-32 The studies varied in target populations for both the checklist user and patients. The outcomes reported varied; 3 studies examined 30-day mortality,2,21,30,36 4 studies examined hospital length of stay,2,21,30,33,36 and 2 studies reported user compliance with the checklist.2,3,31 Five of the studies reported patient outcomes,2,21,30,33,35,36 and 5 studies reported provider-level outcomes related to patient safety (eg, compliance with checklist items such as communication of medications, isolation precautions, etc.).3,1,3-4,3,37

Description of Checklists
Supplementary File 4 (Table 3) provides a detailed breakdown of the checklists’ purpose and components. Six of the checklists were designed to directly reduce patient safety errors,2,21,30,33,35-37 whereas 3 of the checklists were designed to indirectly reduce patient safety events by increasing compliance with processes of care.2,3,2,34 Six checklists were constructed and pilot tested by the research team conducting the RCT,2,35 and the 3 remaining studies used modified versions of previously validated checklists.2,3,36,3,37 The number of items included in the checklist ranged from 2 to 54.

Impact of the Checklist
Table 4 summarizes the adverse events, medical errors, resource utilization and/or compliance reported for each checklist. Chaudhary et al. reported significant decreases in Grade III (requiring intervention)3,3 and IV (life-threatening)3,3 postopera-
TABLE 1. Characteristics of Included Studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Setting</th>
<th>Type of RCT</th>
<th>Sample size</th>
<th>Study Period</th>
<th>Target population</th>
<th>Follow-up duration</th>
<th>Target provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing Group for the CHECKLIST-ICU, 2016</td>
<td>39 academic hospital ICUs &amp; 70 non-academic hospital ICUs, Brazil</td>
<td>Cluster</td>
<td>6,761</td>
<td>6 months</td>
<td>Intensive care unit</td>
<td>Until hospital discharge or death, truncated at 60 days</td>
<td>Multidisciplinary ICU staff (n=6,370 team members across 118 ICUs)</td>
</tr>
<tr>
<td>Gentili M et al, 2016</td>
<td>82 major hospitals, Italy</td>
<td>Cluster</td>
<td>3,520</td>
<td>4 months</td>
<td>Oncology</td>
<td>21 days</td>
<td>Oncology healthcare providers (n=82 oncology units)</td>
</tr>
<tr>
<td>Salzwedel C et al, 2016</td>
<td>University hospital, Germany</td>
<td>Cross-over</td>
<td>134</td>
<td>5 months</td>
<td>Surgical</td>
<td>None</td>
<td>Anesthesiology residents or board-certified anesthesiologists (n=50)</td>
</tr>
<tr>
<td>Chaudhary N et al, 2015</td>
<td>Hospital in India</td>
<td>Parallel</td>
<td>700</td>
<td>15 months</td>
<td>Surgical</td>
<td>Until hospital discharge</td>
<td>Surgical residents (n=NR)</td>
</tr>
<tr>
<td>Haugen AS et al, 2015</td>
<td>Tertiary teaching hospital (1100 beds) &amp; community hospital (300 beds), Norway</td>
<td>Stepped Wedge Cluster</td>
<td>5, 205</td>
<td>10 months</td>
<td>Surgical</td>
<td>30 days postoperatively</td>
<td>Surgical specialties (n=5)</td>
</tr>
<tr>
<td>Bassor A et al, 2013</td>
<td>443-bed community teaching hospital, USA</td>
<td>Parallel</td>
<td>96</td>
<td>15 months</td>
<td>Cardiovascular</td>
<td>6 months after discharge</td>
<td>House staff (n=9 teams)</td>
</tr>
<tr>
<td>Masson SC et al, 2013</td>
<td>2 community &amp; 2 tertiary referral hospitals, Canada</td>
<td>Parallel</td>
<td>61</td>
<td>5 months</td>
<td>Intensive care unit</td>
<td>None</td>
<td>Pharmacy residents (n=6)</td>
</tr>
<tr>
<td>Ong MS et al, 2013</td>
<td>440-bed metropolitan teaching hospital, Australia</td>
<td>Cross-over</td>
<td>300</td>
<td>4 months</td>
<td>Radiology</td>
<td>None</td>
<td>Radiology porters (n=11)</td>
</tr>
<tr>
<td>Salzwedel C et al, 2013</td>
<td>University hospital, Germany</td>
<td>Cross-over</td>
<td>120</td>
<td>2 months</td>
<td>Surgical</td>
<td>None</td>
<td>Anesthesiology residents (n=41)</td>
</tr>
</tbody>
</table>

**Abbreviations:** NR, not reported.

tive complications (23% vs. 33%, P = 0.04) and 30-day mortality (5.7% vs 10.0%, P = 0.04) for patients assigned to the Modified WHO Surgical Safety Checklist compared to controls. Conversely, Haugen et al. reported a nonsignificant reduction in 30-day mortality between the WHO Surgical Safety Checklist group and controls (1.0% vs 1.6%, P = 0.151). Bassor et al. reported no significant difference in 30-day hospital readmission for uncompensated heart failure due to the heart failure discharge checklist group when compared to controls (6% vs. 4%, P = NS); however, an exploratory analysis that excluded patients who died during the follow-up period found a significant difference in 30-day readmission rates (2% vs. 20%, P = 0.02). Gentili et al. reported a higher proportion of patients with pain control in the checklist group compared to the controls (67.6% vs. 54.8%), as well as fewer incidents of analgesic therapy–related uncontrolled adverse events (25.9% vs. 49.9%); however, the statistical significance of these differences were not reported. The Writing Group for CHECKLIST-ICU reported no significant difference for in-hospital mortality between the checklist and control groups (adjusted odds ratio [AOR] 1.02, 95% CI, 0.82–1.26, P = 0.88), nor for the secondary clinical outcomes examined (Table 4). However, there was a significant difference between the checklist group and control group for 3 of the 7 outcomes related to processes of patient care, including a reduction in the use of both urinary catheters (adjusted rate ratio [ARR] 0.86, 95% CI, 0.80–0.93, P < 0.001) and central venous catheters (ARR 0.90, 95% CI 0.83–0.98, P = 0.02). Mason et al. reported that when using the FASTHUG-MAIDENS checklist, more drug-related problems were identified by pharmacy residents (in relation to the number identified by the ICU pharmacist) both per patient encounter (P = 0.008) and overall (P < 0.001). Ong et al. reported higher rates of compliance with isolation precautions for infectious diseases in the checklist group (71% vs. 38%, P < 0.01); however, compliance with the checklist was low (40%) and qualitative analyses found participants were dissatisfied with the checklist. Salzwedel et al. reported the number of items handed over by anesthesia residents postoperatively to be higher in the checklist group than the control group (48.7% vs. 32.4%, P < 0.001). In a more recent study, Salzwedel et al. reported that proportion of items deemed by the attending anesthesiologist as “must be handed over” were more often actually handed over by the anesthesia residents assigned to the checklist group when compared to controls (87.1% vs. 75.0%, P = 0.005).

**30-day Mortality**
A random-effects model pooling data from the 3 studies that reported data for 30-day all cause mortality suggested a significant reduction with use of a checklist (OR 0.60, 95% CI, 0.41–0.89; P = 0.01, I² = 0.0%, P = 0.573).

**Study Quality**
Supplementary File 5 (Table 5) summarizes the quality assessment of the 9 studies. The clarity of description for each intervention varied. All studies reported inclusion/exclusion criteria and randomization procedures. Three studies indicated that outcome assessors were blinded to intervention allocation, while this was unclear in 2 studies.
Three studies reported baseline characteristics.  Two studies reported power calculations; however, one study had a sample size that was less than that required to achieve the target power. The Jadad scores ranged from 1 to 5.

**DISCUSSION**

This systematic review identified 9 RCTs that examined the impact of a checklist on patient safety outcomes in hospitalized patients. The studies employed checklists with different purposes and elements and measured different patient safety outcomes. The methodological quality of the included studies was moderate. In aggregate, the results suggest that checklists may be effective at improving patient safety outcomes, but the small number of moderate quality studies and the heterogeneity of interventions and outcome measures suggests that there is an urgent need for further evaluation.
The most important observation from our systematic review is the paucity of high-quality evidence evaluating checklists’ impact on patient safety outcomes in acute inpatient care. The implementation of checklists is increasingly common as they are relatively low cost to develop and implement, and intuitively make sense. This is particularly true in an era of increasing efforts to standardize care as a means for improving quality and minimizing cost (ie, previous systematic reviews cite 38 unique studies).39 However, implementation of an inadequately tested checklist risks unintended consequences (eg, inefficient resource utilization).18 The small number of RCTs identified might be owing to quality improvement efforts traditionally focusing on ‘real life’ applicability over rigorous research methodology.40 The translation of evidence into clinical practice is known to be slow;41 however, these more rigorous methodologies reduce the risk of biases and generate high-quality evidence, which help to fulfill the necessity to identify best practices while avoiding these unintended consequences.

The studies varied both in the approaches used to develop checklists and in the number of items included (ranging from 2 to 54). What is the optimal method for developing a checklist and how does this impact their effectiveness?42 The answers to these questions are not known. However, this review highlights some important issues to consider when developing a checklist. As the number of items or complexity of a task increases, our ability to efficiently perform the task without aid decreases.43-45 As such, a well-designed checklist should detail explicit instructions on the what, where, when, and how of a given task in a fashion that ensures a consistent accuracy for completing the work.5 It is recommended that construction of a checklist follow the principles of human factors engineering: engage stakeholders and human factors experts in the design; are developed based on user needs and realities; list items in order of importance; are concise and ensure usability and evaluate potential negative consequences (eg time to complete); are pilot tested and validated before implementation; are updated as needed based on the

### TABLE 3. Characteristics of the Checklists Used

<table>
<thead>
<tr>
<th>Reference</th>
<th>Name of Checklist</th>
<th>Purpose of Checklist</th>
<th>No. of Checklist Items</th>
<th>Checklist Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing Group for the CHECKLIST-ICU, 2016</td>
<td>Daily Rounds Checklist</td>
<td>Reduce in-hospital mortality and improve care processes in daily patient care rounds</td>
<td>11</td>
<td>Items related to venous thromboembolism, ventilator-associated pneumonia, central line–associated bloodstream infections, urinary tract infections, nutrition, analgesia, sedation, readiness for extubation, severe sepsis, acute respiratory distress syndrome, antibiotics, tidal volume</td>
</tr>
<tr>
<td>Gentili M et al, 2016</td>
<td>38Checkpain</td>
<td>Improve pain control in oncological patients</td>
<td>7</td>
<td>Items related to intensity of pain, frequency of pain assessment, modification of therapy of pain ≥3 intensity, presence of adverse events, therapy management for adverse events, presence of pain acutization events, therapy for treatment of pain acutization events</td>
</tr>
<tr>
<td>Sahuwedel C et al, 2016</td>
<td>Checklist for patient handover</td>
<td>Improve quality of patient handover and continuity of essential information</td>
<td>54</td>
<td>Patient identification, underlying disease and surgical intervention, pre-existing medical conditions, allergies and medication, anesthesia induction, intraoperative course, cardiovascular system, fluids &amp; blood products, respiratory system, gastrointestinal system, kidney, neurology, postoperative instructions, other</td>
</tr>
<tr>
<td>Chaudhary N et al, 2015</td>
<td>Modified WHO Surgical Safety Checklist</td>
<td>Reduce postoperative complications* and mortality</td>
<td>24</td>
<td>Checklist includes preoperative (sign in), intra-operative (time out), and postoperative (sign out) periods &amp; modified to add review of imaging studies and use of DVT prophylaxis</td>
</tr>
<tr>
<td>Haugen AS et al, 2015</td>
<td>Modified WHO Surgical Safety Checklist</td>
<td>Reduce postoperative complications† and mortality</td>
<td>20</td>
<td>Checklist includes preoperative (sign in), intra-operative (time out), and postoperative (sign out) periods</td>
</tr>
<tr>
<td>Basoor A et al, 2013</td>
<td>Heart failure discharge checklist</td>
<td>Reduce hospital readmission and improve quality of care</td>
<td>31</td>
<td>Documentation regarding medication use, appropriate dose up titration, relevant education and counseling, and follow-up instructions</td>
</tr>
<tr>
<td>Masson SC et al, 2013</td>
<td>FASTHUG-MAIDENS</td>
<td>Reduce the number of drug-related problems‡</td>
<td>14</td>
<td>Modified FASTHUG, Feeding, analgesia, sedation, thromboembolic prophylaxis, hypovolemic or hyperbate delirium, stress ulcer prophylaxis, glucose control, medication reconciliation, antibiotics or anti-infectives, indications for medications, drug dosing, electrolytes, hematology and other lab tests, no drug interactions, allergies, duplicates, side effects, stop dates</td>
</tr>
<tr>
<td>Ong MS et al, 2013</td>
<td>NR</td>
<td>Increase adherence to safety precautions during patient transfer</td>
<td>2</td>
<td>Guideline-based infection control precautions &amp; clinical escort requirements to be handed off to ward nurse</td>
</tr>
<tr>
<td>Sahuwedel C et al, 2013</td>
<td>NR</td>
<td>Increase the number of patient items handed over and improve communication quality</td>
<td>37</td>
<td>Patient demographics, comorbidities, allergies, analgesics, vitals, labs, lines &amp; tubes, transfusions, ventilation, pending investigations</td>
</tr>
</tbody>
</table>

**ABBREVIATIONS:** DVT, deep vein thrombosis; NR, not reported; WHO, World Health Organization.

* Complications defined using the Clavien-Dindo classification.
† Complications defined using the WHO’s International Classification of Diseases, Tenth Revision (ICD-10).
‡ Number of drug-related problems identified by the pharmacy resident relative to the total number of problems determined for each patient encounter.
## TABLE 4. Adverse events, medical errors, resource utilization and checklist compliance

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcome</th>
<th>Checklist Group</th>
<th>Control Group</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing Group for the CHECKLIST-ICU, 2016</td>
<td>In-hospital mortality</td>
<td>32.9%</td>
<td>34.6%</td>
<td>AOR: 1.02 (0.82 to 1.26); P=.88</td>
</tr>
<tr>
<td></td>
<td>ICU mortality</td>
<td>26.3%</td>
<td>25.4%</td>
<td>AOR: 1.17 (0.93 to 1.47); P=.19</td>
</tr>
<tr>
<td></td>
<td>Catheter-related bloodstream infections*</td>
<td>8.3%</td>
<td>8.3%</td>
<td>ARR: 1.03 (0.73 to 1.45); P=.88</td>
</tr>
<tr>
<td></td>
<td>Ventilator-associated pneumonia</td>
<td>5.4%</td>
<td>4.8%</td>
<td>ARR: 1.04 (0.68 to 1.58); P=.87</td>
</tr>
<tr>
<td></td>
<td>Urinary tract infection</td>
<td>10.6%</td>
<td>8.0%</td>
<td>ARR: 1.28 (0.91 to 1.81); P=.16</td>
</tr>
<tr>
<td></td>
<td>Ventilator-free days in 28-day period, days (95% CI)</td>
<td>11.2 (12.9 to 13.0)</td>
<td>13.4 (13.2 to 13.6)</td>
<td>Mean diff: -2.2 (-2.6 to 4.0); P=.36</td>
</tr>
<tr>
<td></td>
<td>ICU length of stay, days (95% CI)</td>
<td>11.0 (10.6 to 10.7)</td>
<td>10.4 (10.1 to 10.7)</td>
<td>Mean diff: -0.6 (-0.7 to 0.1); P=.28</td>
</tr>
<tr>
<td></td>
<td>Hospital length of stay, days (95% CI)</td>
<td>18.7 (18.4 to 19.2)</td>
<td>20.0 (19.4 to 20.6)</td>
<td>Mean diff: -1.3 (-1.6 to 0.8); P=.31</td>
</tr>
<tr>
<td></td>
<td>Tidal volume ≤8mL/kg of predicted body weight</td>
<td>67.5%</td>
<td>58.9%</td>
<td>AAR: 1.14 (1.03 to 1.26); P=.01</td>
</tr>
<tr>
<td></td>
<td>Moderate sedation to alert and calm</td>
<td>40.5%</td>
<td>35.0%</td>
<td>A R: 1.19 (1.00 to 1.42); P=.05</td>
</tr>
<tr>
<td></td>
<td>Central venous catheter use</td>
<td>72.4%</td>
<td>72.9%</td>
<td>A R: 0.90 (0.83 to 0.98); P=.02</td>
</tr>
<tr>
<td></td>
<td>Urinary catheter use</td>
<td>62.8%</td>
<td>74.8%</td>
<td>AAR: 0.86 (0.80 to 0.93); P=.001</td>
</tr>
<tr>
<td></td>
<td>Overall patient items handed over, median (IQR)</td>
<td>116.5% (252.5 to 341.6)</td>
<td>120.5 (80.5 to 170.5)</td>
<td>60.0% (36.7 to 251.0)</td>
</tr>
<tr>
<td></td>
<td>Compliance with checklist</td>
<td>79.2%</td>
<td>76.4%</td>
<td>A AR: 1.03 (0.89 to 1.20); P=.65</td>
</tr>
</tbody>
</table>

**Gentili M et al, 2016**

- Patients with controlled pain: 67.6% vs. 54.8%, NR
- Intensity of pain, mean: 2.7 vs. 2.7, NR
- Uncontrolled adverse events related to analgesic therapy: 25.9% vs. 49.9%, NR
- No. breakthrough pain events: 1.9 vs. 1.9, NR

**Salzwedel C et al, 2016**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Checklist Group</th>
<th>Control Group</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Must be”† items handed over, median (IQR)</td>
<td>87.1% (77.1 to 90.0)</td>
<td>75.0% (66.7 to 88.6)</td>
<td>.005</td>
</tr>
<tr>
<td>“Should be”‡ items handed over, median (IQR)</td>
<td>60.0% (36.7 to 101)</td>
<td>50.0% (33.3 to 69.0)</td>
<td>NS</td>
</tr>
<tr>
<td>Length of handover, seconds, median (IQR)</td>
<td>208 (142 to 276)</td>
<td>174 (115 to 255)</td>
<td>.201</td>
</tr>
</tbody>
</table>

**Chaudhary N et al, 2015**

- Compliance with checklist: 85%
- Patients experiencing one or more complications‡: 48% vs. 52%, .15
- Postoperative complications‡ per patient: 0.80 vs. 0.97, .06
- No. high grade IV/V postoperative complications‡ per patient: 0.23 vs. 0.33, .004
- Length of hospital stay, days, median: 9 vs. 9, .54
- 30 day mortality: 5.7% vs. 10.0%, .04

**Haugen AS et al, 2015**

- Complications§: 11.5% vs. 19.9%, <.001
- Length of hospital stay, days, mean: 7.0 vs. 7.8, .022
- 30-day hospital mortality: 1.0% vs. 1.6%, .151

**Basoor A et al, 2013**

- Evidence-based prescribing of medications: 83% vs. 48%, <.001
- Evidence-based medication dosing: 44% vs. 8%, <.001
- Length of hospital stay, days, mean: 6.4 vs. 4.6, NS
- 30-day readmission: 6% vs. 19%, NS
- 30-day readmission (excluding deaths): 2% vs. 20%, .02
- 30-day mortality: 6% vs. 4%, NS
- 6-month readmission for uncompensated heart failure: 23% vs. 42%, .045
- 6-month readmission for any reason: 31% vs. 60%, .003

**Masson SC et al, 2013**

- Drug-related problems identified per patient encounter: 73.2% vs. 52.4%, .008
- Overall drug-related problems capture rate: 77.1% vs. 52.5%, <.001

**Ong MS et al, 2013**

- Compliance with checklist: 40%
- Compliance with infection control precautions: 71% vs. 38%, <.01

**Salzwedel C et al, 2013**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Checklist Group</th>
<th>Control Group</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall patient items handed over, median (IQR)</td>
<td>48.7% (37.8 to 70.9)</td>
<td>32.4% (27.0 to 40.5)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Length of handover, seconds, median (IQR)</td>
<td>120.5 (80.5 to 170.5)</td>
<td>85.5 (85.3 to 106.0)</td>
<td>.003</td>
</tr>
</tbody>
</table>

**Abbreviations:** AOR, adjusted odds ratio; ARR, adjusted rate ratio; IQR, interquartile range; NR, not reported; NS, not significant.

* Number per 1,000 patient-days
† Items identified by the attending anesthesiologist on a patient-to-patient basis.
‡ Complications defined using the Clavien-Dindo classification.
§ Complications defined using the WHO’s International Classification of Diseases, Tenth Revision (ICD-10).
‖ Number of drug-related problems identified by the pharmacy resident relative to the total number of problems determined for each patient encounter.
generation of new findings or changes in operational procedures. These general principles of human factors engineering provide a practical approach for the development and evaluation of a checklist. In addition, standardization of operational definitions (ie, process, outcome, compliance) is important for study replication and robust meta-analyses.

Checklists used in aviation are perhaps best known and the evidence of their effectiveness is derived from the attribution of aviation errors to incomplete checklists. Although more recently implemented in medicine, checklists have the potential to guide the successful completion of complex tasks in healthcare. Systematic reviews of observational studies have been conducted for specific checklists (eg, WHO Surgical Safety Checklist) and for select patient populations (eg, surgical patients), and the number of included studies ranges from 7-27 (n = 38 unique studies). For example, Gillespie et al. in a systematic review and meta-analysis reported the implementation of Surgical Safety Checklists to be associated with a reduction in postoperative complications (relative risk [RR] 0.63, 95% CI, 0.58-0.72, P = < 0.001), but not mortality (RR 1.03, 95% CI, 0.73-1.4, P = 0.857). Similarly, Treadwell et al. reported in a systematic review of Surgical Safety Checklists that while data are promising, more evaluation of their impact on clinical outcomes is needed. These recommendations are nicely illustrated by Urbach et al.’s and O’Leary et al.’s evaluations of the mandatory adoption of Surgical Safety Checklists across all hospitals in Ontario, Canada, which respectively demonstrated no significant reductions in 30-day perioperatively complications for both adult (OR 0.97, 95% CI, 0.90-1.03, P = 0.29) and pediatric (AOR 1.01, 95% CI, 0.90-1.14, P = 0.9) patients. These data not only highlight the need for further evaluation of checklists but are also a reminder that checklists and their associated implementation strategies are complex interventions for which there may be important differences between the efficacy reported in clinical trials and the effectiveness reported in implementation studies. This all suggests that if checklists are to be effective in improving patient safety, process evaluations of implementation and realist reviews of published studies may be important to determine optimal approaches for implementation. We believe that, based on the limited currently available evidence, there is urgency for further robust evaluations of checklists before their widespread implementation. If effective, they should be widely implemented. If ineffective, they should be abandoned to minimize unintended consequences and inefficient use of resources.

There are 4 primary limitations to this review that should be considered when interpreting the findings. First, the RCT design is not the study design employed by most quality improvement initiatives. While some quality improvement experts may argue that an RCT design is insufficiently flexible for applied settings, it does minimize the risk of biased assessments of intervention effectiveness. Second, our search strategy included an RCT filter. The filter helped restrict the number of citations to be reviewed (n = 11,225) but could have resulted in improperly indexed studies being excluded.

To guard against this risk, we used the validated Cochrane Highly Sensitive Search Strategy for Identifying Randomized Trials, reviewed reference lists of citations included in the review, and solicited suggestions for missing studies from quality improvement experts. Third, our review was restricted to hospitalized patients. Although the studies evaluated commonly reported safety outcomes across patients with diverse clinical conditions, care settings, and providers that broadly reflect hospital-based care, evaluations of checklists in additional patient and provider groups are needed (eg, hospitalists). Furthermore, the effectiveness of checklists for improving patient safety outcomes in outpatients is important; however, the organizational and patient characteristics of these 2 settings (hospitalized vs outpatient) are sufficiently different to warrant separate systematic reviews. Finally, owing to the heterogeneity of the checklists used and outcomes measured, we were unable to perform a robust meta-analysis. Heterogeneity, combined with the small number of studies identified in our search, prevented us from applying statistical methods to assess for publication bias. This limitation of our systematic review highlights an important gap in the literature and emphasizes the importance of additional primary research to evaluate checklists.

In summary, we identified few RCTs that examined checklists designed to improve patient safety outcomes. The small number of existing studies suggests that checklists may improve patient safety outcomes; however, these observations were not reported for all outcomes examined and the studies were heterogeneous and of limited methodological quality. There is an urgent need for high-quality evaluations of the effectiveness of patient safety checklists in inpatient healthcare settings to substantiate their perceived benefits.

Acknowledgments
We would like to thank Diane Lorenzoetti for her help with the development of the search strategy.

Disclosure: The authors have no known conflicts of interest to declare.

Jamie Boyd was supported by a W21C – Alberta Innovates-Health Solutions (AIHS) Collaborative Research and Innovation Opportunities (CRIOS) Health Services Research graduate studentship. Guosong Wu was supported by a Western Regional Training Centre (WRTC) for Health Services Research graduate studentship. Dr. Stelfox was supported by a Population Health Investigator Award from Alberta Innovates Health Solutions.

Authors’ Contributions
HTS was responsible for the study’s conception. All 3 authors contributed to the study’s design and interpretation. JB and GW were responsible for searching the literature, reviewing abstracts, selecting full-text articles and critically appraising them. All 3 authors performed the analyses. JB drafted the manuscript and all 3 authors assisted in the successive revisions of the final manuscript. All authors have read and approved the final manuscript.

References


43. Miller GA. The magical number seven, plus or minus two. Some limits on our capacity for processing information. Psychol Rev. 1956; 63(2):81-97.


