Effect of Using a Safety Checklist on Patient Complications after Surgery: A Systematic Review and Meta-Analysis

1Brigid M. Gillespie Ph.D., R.N., National Health & Medical Research Council Research Centre for Excellence in Nursing, Research Centre for Health Practice Innovation, Griffith Health Institute, Griffith University, Gold Coast Campus, Queensland 4222, Australia

2Wendy Chaboyer, Ph.D., R.N., National Health & Medical Research Council Research Centre of Research Excellence in Nursing, Research Centre for Health Practice Innovation, Griffith Health Institute, Griffith University. Queensland, Australia

3Lukman Thalib, Ph.D., B.Sc., Department of Community Medicine, Faculty of Medicine, Kuwait University, Safat, Kuwait

4Melinda John, B.Sc., Research Centre for Health Practice Innovation, Griffith Health Institute, Griffith University, Queensland, Australia

5Nicole Fairweather, F.A.N.Z.C.A., M.D., Department of Anaesthetics, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia

6Kellee Slater, F.R.A.C.S., M.D., Department of Hepatobiliary, Liver Transplant and General Surgery, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia
Corresponding author:

Brigid M. Gillespie, Ph.D., R.N.

Telephone: 61-7 5552-9718

Fax: 61-7-5552-8526

Email: B.Gillespie@griffith.edu.au

Funding Acknowledgment: Brigid Gillespie acknowledges the financial assistance of the Australian Research Council, Early Career Discovery Fellowship Scheme, Canberra, ACT, Australia.

Conflict of Interest Statement: The authors declare no competing interests.

Running Header: Checklists in surgery improve patient outcomes (46 characters with spaces)

TOC Statement: The investigators conducted a meta-analysis which included 7 nonrandomized checklist studies with a total of 37,339 patients. Use of a checklist significantly reduced complications, wound infections, and blood loss, but not mortality.
ABSTRACT

Background: Previous before-and-after studies indicate that the use of safety checklists in surgery reduces complication rates in patients.

Methods: A systematic review of studies was undertaken using MEDLINE, CINAHL, Proquest, and the Cochrane Library were utilised to identify studies that evaluated checklists in surgery on complication rates. Study quality was assessed using the Methodological Index for Nonrandomized Studies (MINORS). The pooled risk ratio (RR) was estimated using both fixed and random effects models. For each outcome, number needed to treat (NNT) and the absolute risk reduction (ARR) was also computed.

Results: Of the 207 intervention studies identified, 7 representing 37,339 patients were included in meta-analyses and all were cohort studies. Results indicated that the use of checklists in surgery compared with standard practice led to a reduction in any complication (RR = 0.63, 95% CI: 0.58-0.72, p < 0.0001, ARR = 3.7%, NNT = 27), wound infection (RR = 0.54, 95% CI: 0.40-0.72, p = 0.0001, ARR = 2.9%, NNT = 34), and reduction in blood loss (RR = 0.56, 95% CI: 0.45-0.70, p = 0.0001, ARR = 3.8%, NNT = 33). There were no significant reductions in mortality (RR = 0.79, 95% CI: 0.57-1.11, p = 0.191, ARR = 0.44%, NNT = 229), pneumonia (RR = 1.03, 95% CI: 0.73-1.4, p = 0.857, ARR = 0.04%, NNT = 2,512) or unplanned return to operating room (RR = 0.75, 95% CI: 0.56-1.02, p = 0.068, ARR = 0.52%, NNT = 192).

Conclusions: Notwithstanding the lack of randomized controlled trials, our synthesis of the existing body of evidence suggests a relationship between checklist use in surgery and fewer postoperative complications.
Introduction

Over the last decade, checklists have become commonplace in healthcare practice as a strategy to improve patient safety. Surgical checklists emphasise several salient components of patient safety; safe anaesthesia and airway function, correct surgical site/side, infection prevention, and effective teamwork.\(^1,2\) The intent of a checklist as a safety tool is to standardize and make more predictable, team performance across a diverse range of individuals, situations and clinical environments.\(^2\) The fervent introduction of checklists such as the World Health Organization’s (WHO) Surgical Safety Checklist (WHO SSC) has heralded a legislative mandate for their implementation in operating rooms (OR) in over 122 countries.\(^†\)

Thus, checklists have become synonymous as best practice in high-risk areas such as surgery.\(^2,3,4\)

Checklists hold the promise of reducing catastrophic errors such as wrong site/wrong patient surgery\(^5\), improving interprofessional communications\(^6,7\), enhancing work satisfaction\(^8,9\), and flattening the hierarchy that often characterises the culture of surgical teams.\(^10\) A recent systematic review on the impacts and implementation of checklists suggests that checklist use was associated with increased detection of potential safety hazards.\(^4\) However, the results of many of the studies included reflect patterns of practice on a local or regional level rather than amassed on a wider scale. Thus, evaluating the impact of new clinical practice initiatives on the outcomes of care is problematic in the absence of accurate, large scale patient outcome

---


Further examination to evaluate to what extent checklists improve clinical outcomes is needed. To this end, we conducted a systematic review and meta-analysis of available studies that tested effects of surgical safety checklists on complication rates in surgical patients.
Materials and Methods

Literature Search and Eligibility Criteria

The study protocol was informed by guidelines developed by groups such as the Cochrane Collaboration\(^{12}\) and the Institute of Health Improvement. A systematic literature search was undertaken using MEDLINE, CINAHL (EBECOhost source), Proquest (Nursing and Allied Health Source), and the Cochrane Library. Publications dated from January 2000 to May 2013 were included. MeSH search terms, and their combination included: “randomized controlled trial”; “checklist”; “mortality”; “surgery”; “morbidity”; and, “intervention”, and “complication”. Specific data base functions such as “apply related words” and “explosion” were used to maximize the search. Reference lists of retrieved papers were further screened for additional publications. Studies meeting the following criteria were included: Design – randomized controlled trial (RCT), prospective and retrospective cohort, quasi-experimental and interrupted time series; Population – patients undergoing elective or emergency surgical procedures; Intervention – a single OR-specific surgical checklist; Comparator – control group where a surgical checklist was not used; Outcome – postoperative complications, however defined by the primary study authors. We excluded papers if they were: a) not written in English; and b) used a checklist that was not perioperative-specific or encompassed other clinical settings (e.g., SURgical Patient Safety System [SURPASS]).\(^{\dagger}\)

\(^{\dagger}\) The SURPASS checklist is far more comprehensive and interdisciplinary; however, it is also more complex to implement and often requires reorganization of care processes.
Data Extraction

Study-specific descriptive information included country, study design, type of patients, surgical procedures, specific nature of checklist intervention, and quantitative patient outcomes. Patient morbidity and mortality rates were extracted. Data extraction was independently performed by two authors and discrepancies adjudicated by a third review author. Authors of primary studies were contacted where necessary (Personal Written Communication: Mehrdad Askarian, M.D., M.P.H., Department of Community Medicine, Shiraz University, Shiraz, Iran, February 27, 2013; Thomas Weiser, M.D., M.P.H., Department of Surgery, General Hospital, Boston, Massachusetts, February 29, 2013) to obtain additional information on published data.

A postoperative complication was broadly defined as “an undesirable, unintended event and would not have occurred had the operation gone as could be reasonably hoped”. In this review, definitions of the primary outcome ‘any major complication’ were informed by the American College of Surgeons National Quality Improvement Program and/or the Clavien Classification system or however defined by the primary study authors. Secondary outcomes included individual complications; mortality, surgical site infection (SSI), pneumonia, wrong site surgery, unplanned return to OR, and blood loss > 500 mL (i.e., preoperative blood loss of 500mL more than expected for a given case), or however defined by the primary study authors. A surgical checklist was defined as a cognitive tool delineating team tasks according to preinduction, before skin incision, and before leaving the OR, used to prompt interdisciplinary verification of surgical and patient details to increase teamwork. Standard care was defined
as the usual practices undertaken by the interdisciplinary surgical teams prior to checklist implementation or however defined by primary study authors.

**Study Quality**

Systematic review and meta-analysis of studies other than RCTs is challenging because combining observational studies of heterogeneous quality may be highly biased.\(^{12,13,16}\) Therefore, we required a quality assessment tool that allowed us to compare observational studies with wider heterogeneity. For this reason, quality was assessed using a modified version of the previously validated *Methodological Index for Nonrandomized Studies* (MINORS).\(^{16}\) The original version of the MINORS is a 12-item index. Each item is scored as either: 0 = not reported; 1 = reported but inadequate; or, 2 = reported and adequate, with the combined score ranging from 0 to 24 for observational studies.\(^{16}\) For the purposes of this review, we made modifications to the index. Specifically, items 7 and 10\(^{16}\) were not considered as these items are better suited to evaluate studies using contemporaneous groups, *i.e.*, control and intervention groups that are managed during the same time period (no historical controls). Therefore, in this study, the highest possible score was 20. Two review authors (Gillespie and John) independently assessed study quality and the proportion of agreement was measured using the intraclass correlation coefficient. A coefficient of ≥0.70 was considered adequate.\(^{17}\)
**Statistical Analysis**

The software package *Comprehensive Meta-Analysis Version 2.0* (Biostat Solutions Inc., Englewood, NY) was used to estimate the overall pooled effect size with fixed or random effects models for each patient outcome. Statistical heterogeneity among the studies was assessed using $I^2$-squared ($I^2$) and Cochrane’s $Q$ statistic ($\chi^2$ test) for heterogeneity. The significance level of the $\chi^2$ test was set at $p < 0.05$. The degree of statistical heterogeneity was determined if the $I^2$ was $\geq 50\%$. In cases of statistical heterogeneity, only random effects models were used.

Results of the meta-analysis are presented using risk ratio and corresponding 95% confidence intervals [CI] and the $z$ test for pooled effect size estimates. For each outcome, the number needed to treat (NNT) and absolute risk reduction (ARR) was calculated from the pooled risk ratio and the event rates to further place risk estimates into context. Forrest plots are used for graphical display of results.

To assess the robustness of the results, sensitivity analyses were performed.\textsuperscript{12, 18} We estimated patient outcome measures after excluding studies with lower methodological quality to check whether the results had changed. If the results did not change significantly after excluding low quality studies, then they were considered to be robust. If the results had changed or the conclusions differed, then they had low stability. We assessed potential publication bias by performing informal visual inspection of funnel plot symmetry\textsuperscript{12} based on the primary outcome, any major complication.
Results

Figure 1 summarizes the number of observational studies identified in the systematic review and meta-analysis according to the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* statement.\(^{19}\) Our search yielded 207 articles. Of these, 7 (3.4%) studies\(^{20-26}\) with a total of 37,339 patients met eligibility criteria. The included studies were published between 2009 and 2012. Characteristics of each study are summarized in table 1. All studies were observational in nature, and most employed a prospective cohort design using historical controls, with the exception of van Klei *et al.*\(^{24}\) who used a retrospective cohort design. Most studies (6/7, 86%) used the WHO Surgical Safety Checklist\(^{27}\), however the study by Bliss *et al.*\(^{21}\) used the *Association of peri-Operative Registered Nurses (AORN) Comprehensive Surgical Checklist*\(^{\S}\) (refer to the Supplemental Digital Content 1 which details items on the AORN Checklist) which includes items based on the WHO Checklist\(^{*}\) and the Joint Commission Universal Protocol (JCUP). The studies by Haynes *et al.*\(^{22}\) and Weiser *et al.*\(^{25}\) although based on the same multinational study, focussed on different subanalyses of patients. Weiser *et al.*\(^{25}\) examined checklist use in urgent surgeries whereas Haynes *et al.*\(^{22}\) included all patient groups in their analyses. Consequently, where both of these studies examined the same patient outcome, we analyzed those outcomes (*i.e.*, any major complication, mortality, SSI) based on the results of Haynes *et al.*\(^{22}\) The study by Weiser *et al.*\(^{25}\) was included in the analysis of only one outcome, expected blood loss > 500mL (an outcome not measured by Haynes *et al.*\(^{22}\)).

The quality assessments of the studies by two independent reviewers were in good agreement (intraclass correlation coefficient = 0.80; CI = 0.27-0.96; \( p = 0.008 \)). Out of a possible score of 20 on the modified 10-item MINORS tool\(^\text{16}\), the mean score for the 7 included studies was 12.6 (± 1.3, range 11-15). Four studies (57%) achieved a score of ≥ 13/20.

**Primary Outcome**

*Any Major Complication:* Five published studies\(^\text{20-23, 26}\) with 9,747 patients compared checklist use relative to this outcome which combined single complications that included up to 18 individual complications (e.g., pulmonary embolism, deep vein thrombosis, SSI, pneumonia, unplanned return to OR, blood loss, death, wound dehiscence, cerebrovascular accident, myocardial infarction, vascular graft failure, coma, sepsis, unplanned intubation, systematic inflammatory response syndrome, septic shock, cardiac arrest, acute renal failure). As there was no significant heterogeneity among studies (\( \chi^2 = 4.37, \text{df} = 4, p = 0.357, I^2 = 8.6\% \)), we used the fixed effect model (fig. 2). Pooling of individual studies indicated that 381/4,922 (7.7%) patients in the intervention group and 592/4,835 (12.2%) patients in the control group developed any complication. The ARR was 3.7% (95% CI 2.6%-4.8%) and the NNT was 27 (95% CI 21-38). That is, one less major complication may be prevented for every 27 patients where the checklist is used.

**Secondary Outcomes**
**Mortality:** Four published studies\(^{22-24, 26}\) with 34,642 patients compared checklist use in relation to this outcome. Pooled analysis revealed that 356/15,840 (2.3%) patients in the intervention group and 520/18,807 (2.8%) patients in the control group died. As there was significant heterogeneity (\(\chi^2 = 5.58, \text{df} = 3, p = 0.039, I^2 = 46.3\%\)) among studies, a random effects model was used (fig. 3). The ARR was 0.44\% (95\% CI 0.12\%-0.76\%) and the NNT was 229 (95\% CI 131-867), suggesting that one less death may be prevented for every 229 patients where the checklist is used. However, in 3/4 included studies, the confidence intervals cross the line of no difference, which suggests that the use of a checklist is not associated with a reduction in mortality.

**SSI:** Five published studies\(^{20-23, 26}\) with 9,747 patients assessed this outcome. As there was significant heterogeneity between studies (\(\chi^2 = 6.50, \text{df} = 4, p = 0.0001, I^2 = 38.4\%\)), a random effects model was used (fig. 4). Pooling of individual studies showed that 188/4,912 (3.8\%) patients in the intervention group and 344/4,835 (7.1\%) patients in the control group developed a SSI. The ARR was 2.9\% (95\% CI 2.1\%-3.8\%) and the NNT was 34 (95\% CI 26-47). That is, one less SSI may be prevented for every 34 patients where the checklist is used.

**Pneumonia:** Four published studies\(^{20-23}\) with 9,266 patients evaluated this outcome. As there was no significant heterogeneity between studies (\(\chi^2 = 11.4, \text{df} = 4, p = 0.022, I^2 = 64.3\%\), fig. 5), a fixed effect model was used. Pooled analysis revealed that 69/4,663 (1.5\%) patients in the intervention group and 70/4,603 (1.5\%) patients in the control group developed pneumonia. The ARR was 0.04\% (95\% CI -0.4\%-5.0\%) and the NNT was 2,512. In this instance, the number of patients required in the checklist group is greater than 190, compared to the
standard care group, which exceeds 224 (95% confidence). In these four studies, the confidence intervals cross the line of no difference, suggesting that the use of a checklist is not associated with a reduction in pneumonia.

**Expected blood loss ≥ 500mL:** Two published studies\(^21,25\) with 2,069 patients assessed this outcome. As there was no significant heterogeneity between studies (\(\chi^2 = 9.38, df = 1, p = ., I^2 = 0.0\%\), fig. 6), a fixed effect model was used. Pooling of individual studies showed that 122/981 (12.4\%) patients in the intervention group and 182/1,088 (16.7\%) patients in the control group had experienced bleeding requiring ≥ 4 units of blood within 72 h after surgery. The ARR was 3.8\% (95\% CI 0.6\%-5.9\%) and the NNT was 33 (95\% CI 17-168). Therefore, one less patient may have blood loss > 500mL for every 31 patients where the checklist is used.

**Unplanned return to OR:** Two published studies\(^22,23\) with 8,653 patients assessed this outcome. As there was no significant heterogeneity between studies (\(\chi^2 = 0.193, df = 1, p = 0.661, I^2 = 0.0\%\), fig. 7), we used a fixed effect model. Pooling of individual studies showed that 76/4,440 (1.7\%) patients in the intervention group and 95/4,213 (2.3\%) patients in the control group had an unplanned return to the OR within 72 h after surgery. The ARR was 0.52\% (95\% CI -0.5\%-1.1\%) and the NNT was 192, suggesting that unplanned return to OR may be avoided for every 192 patients where the checklist is used. The number of patients required in the checklist group exceeds 91, compared to the standard care group, which exceeds 1,825 (95\% confidence). In the two included studies for this outcome, the confidence intervals cross the line of no difference, suggesting that the use of a checklist is not associated with a reduction in unplanned return to OR.
**Reporting Bias**

The funnel plot (fig. 8) for the primary outcome, any major complication, was relatively symmetrical suggesting that publication bias was not present. Begg's test ($z = 0.939$, $p = 0.347$) and Egger's test ($p = 0.452$) did not support publication bias.
Discussion

We found seven observational studies that examined the effect of using surgical checklists on any major postoperative complication, and individual complications (i.e., mortality, SSI, pneumonia, blood loss and unplanned return to OR). The main finding from this analysis is that checklists appear to be associated with the reduction in the risk of ‘any major complication’ which was measured as a composite variable. Our results are similar to those of an earlier meta-analysis by Borchard et al.\textsuperscript{28} which included three studies that examined the outcome ‘any complication’. Yet, given the observational nature of the research, we are unable to establish causality. In both ours and Borchard’s\textsuperscript{28} reviews, the primary authors measured ‘any major complication’ based on previous definitions of the American College of Surgeons National Quality Improvement Program and/or the Clavien Classification system.\textsuperscript{14} That the authors\textsuperscript{14-17, 20} combined all of the complications they measured in the primary studies as a single outcome (i.e., a composite measure) more likely demonstrated significant before-and-after changes in this outcome.

Notably, there were significant risk reductions for the outcomes of SSI\textsuperscript{20-23, 26} and expected blood loss.\textsuperscript{12, 16} These results may reflect the specific items that target these aspects of intraoperative care. For example, item 7 (sign-in) of the WHO SSC\textsuperscript{27} addresses the risk of blood loss while items 12/13 (time out) require confirmation of equipment availability / instrument sterility and antibiotic prophylaxis (refer to the Supplemental Digital Content 1 which details WHO SSC items). Hospital-acquired infections are considered preventable—and therefore human error may contribute to SSI.\textsuperscript{4} Using a checklist in surgery is considered
effective in assisting individuals and teams to remember key information or actions that would otherwise be overlooked, reducing the potential for human error.\textsuperscript{24, 29}

In our analyses, the nonsignificant results relative to individual complications of pneumonia\textsuperscript{20-23}, patient mortality and unplanned return to OR\textsuperscript{22, 23} may be attributed to several factors: First, the numbers of patients who developed pneumonia or who died were relatively lower for each of these individual complications, representing less than 3\% patients who received the checklist intervention. Given the low percentages of patients that developed pneumonia or died, a significant reduction would have required a much larger number of patients, and unlikely given the potential effect size (\textit{i.e.}, low numbers, small effect size). Second, all studies included herein individually reported superiority of checklist use over standard practice—but ‘standard practice’ was never defined in any of these studies. Written correspondence with study authors (Mehrdad Askarian, M.D., M.P.H., Department of Community Medicine, Shiraz University, Shiraz, Iran, February 27, 2013; Thomas Weiser, M.D., M.P.H., Department of Surgery, General Hospital, Boston, Massachusetts, United States, February 29, 2013) indicated that “timeout” was used in some hospitals, but in most cases, there was no routine or standardized practice prior to checklist implementation. Thus, there was no stable baseline from which to draw comparisons. It may be that standard practice is naturally improving because of contemporary trends and organizational initiatives around patient safety.\textsuperscript{2, 3} There is little doubt that the culture of surgery is changing, and the importance of human factors is increasingly being recognized.\textsuperscript{30} Finally, there is a danger in treating the checklist as a ‘tick box’ exercise, rather than emphazizing the need for clinicians to engage in the process.\textsuperscript{4, 11, 29}
**Methodological Limitations**

We acknowledge some limitations. While meta-analysis is powerful—it is also controversial because small violations in meeting critical assumptions can give misleading results. Critique of observational studies is not an exact science, and the quality assessment of these studies is essentially subjective. However, we used a tool that has established validity and appropriate, and has demonstrated an acceptable level of agreement in the appraisal of the included studies. Our search methods were exhaustive and robust, but it is possible that we may have missed other important studies. An additional concern was the modest number studies available for inclusion.

A significant methodological limitation to the review studies was the lack of a control and a comparison group—none were RCTs. All included studies used a cohort design, thus selection and reporting biases cannot be eliminated. Before-and-after studies are problematic in estimating the efficacy of an intervention. In some of the reviewed studies, insufficient sample size, selection bias, surgical team and institutional factors were also potential confounders. Moreover there is the likely disparity in patient characteristics due to nonrandomization which may have potentially increased variability in observed heterogeneity effects among studies. Where appropriate we used random-effects models to account for the presence of other unpublished studies, not included herein. Patient factors that may have confounded the results of the primary studies include comorbidities, procedural complexity, whether the surgery was elective or emergency—all of which impact on postoperative
outcomes. Patients needing emergency surgery are at greater risk of developing postoperative complications.\(^3\) Further, organizational factors, differences in hospital sites and temporal effects (i.e., organizational initiatives, cycles of surgical training) may have been responsible for reduced post operative complications, rather than intervention effects. Finally, team culture and communication practices may have led to differences in checklist implementation and the level of adherence.

In all of the reviewed studies, the WHO SSC or adapted version was used (see Supplemental Digital Content 1\(^3\)), and there was relative uniformity in the number of checklist items, with either 19\(^{20,22,23,26}\) or 22\(^{24}\) items. Yet across studies there was disparity in adherence in relation to; i) the number of occasions the checklist was used (per patient) after its introduction, and ii) the number of checklist items that were fully completed. For instance, of the review studies that examined checklist adherence post implementation, overall rates varied from 18.6\(^{26}\) to 96.9\%.\(^{23}\) The earlier systematic review by Borchard et al.\(^{28}\) examined checklist implementation as their major outcome. In that study, the mean overall compliance rate was 75\%, and the compliance for the ‘timeout’ component of the checklist was much higher, with an average of 91\%. In our review, variability in compliance rates may have confounded the results of these studies. Follow up periods in the review studies were up to 30 postoperative days or discharge, whichever came first.\(^{20-24,26}\) The brief duration of the follow-up period may have only captured the initial flush of enthusiasm following the introduction of a new project, biasing the results. As data collection was limited to this 30-day period, late complications may have been underestimated. However, it is also plausible that there were more events at the initiation of the study than during the entire study period.
Despite these limitations, the results of the current meta-analyses are meaningful because of the large number (≥40,000) and heterogeneity of patients analyzed. While we are cautionary in our judgement of external validity, the review studies included were conducted with patients drawn from global populations. As such our findings suggest that the WHO SSC has wide application.

RCT methods are likely to have increased internal validity but they cannot answer all important questions for a given intervention.\textsuperscript{31} The “sterility” offered by RCTs is likely untenable in assessing the effectiveness of checklist use in real world clinical environments, where pragmatic approaches to practice improvements are often required.\textsuperscript{29, 32} Checklists facilitate an exchange or clarification of information among team members to avert sentinel events in surgery.\textsuperscript{1, 3, 5} Consequently, checklist use supports a safety culture\textsuperscript{33}, and is viewed as best practice.\textsuperscript{2, 4, 29} As such, it would be inappropriate to withhold their use in clinical practice rendering RCT methods, unfeasible. While statistical significance is important when considering the efficacy of an intervention; the intervention implemented must be practical, sustainable, cost effective and clinically important. In some instances, there were minor modifications made to the checklists to ensure that they were contextually responsive to the local hospitals where the review studies were conducted.

The strengths of this meta-analysis include the rigour used to identify and evaluate available studies. The \textit{Preferred Reporting Items for Systematic Reviews and Meta-Analyses} guidelines\textsuperscript{19} were adhered to and we utilized an appropriate, validated tool to assess study quality.\textsuperscript{16} The reviewed studies incorporated 11 countries (\textit{i.e.}, Canada, India, Jordan, New Zealand, Philippines, Tanzania, England, the United States, Iran, the Netherlands, and Liberia),
and varied in patient case-mix, consequently our results may have global applicability. There was reasonable homogeneity in the outcome measures and the intervention evaluated. Despite some methodological limitations and the caveats around interpretation of findings, the meta-analyses presented herein represent the best available current evidence.

Conclusions

These results suggest that checklists are associated with a reduction in overall complications in surgical patients. Surgical safety checklists provide a means to safeguard patients and minimize risk through increased team cohesion and coordination. Importantly checklists should be used to augment, and not replace, other initiatives that contribute to a safety culture.
References


Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Flow Diagram\textsuperscript{10} of literature search. Figure adapted with permission.

OR = operating room.
Figure 2: Forrest plot comparing checklist use with standard practice for any major complication, 5 studies\textsuperscript{20-23, 26} included (RR = 0.63 [95% CI: 0.56-0.72], z = -7.16, \( p < 0.0001 \)).

CI = confidence interval; RR = risk ratio.

Figure 3: Forrest plot comparing checklist use with standard practice for patient mortality, 4 studies\textsuperscript{22-24, 26} included (RR 0.79 [95% CI 0.57-1.11], z = -1.30, \( p = 0.191 \)).

CI = confidence interval; RR = risk ratio.
Figure 4: Forrest plot comparing checklist use with standard practice for SSI complication, 5 studies\textsuperscript{20-23, 26} included (RR 0.54 [95% CI 0.40-0.72], z = -4.12, \(p < 0.0001\)).

CI = confidence interval; RR = risk ratio; SSI = surgical site infection.

Figure 5: Forrest plot comparing checklist use with standard practice for pneumonia complication, 4 studies\textsuperscript{20-23} included (RR 1.03 [95% CI 0.73-1.45], z = 0.180, \(p = 0.857\)).

CI = confidence interval; RR = risk ratio.
Figure 6: Forrest plot comparing checklist use with standard practice for expected blood loss complication, 2 studies\textsuperscript{21, 25} included (RR 0.64 [95% CI 0.52-0.80], \( z = -4.07, p < 0.0001 \)).

CI = confidence interval; RR = risk ratio.

Figure 7: Forrest plot comparing checklist use with standard practice for unplanned return to operating room (OR), 2 studies\textsuperscript{22, 23} included (RR 0.75 [95% CI 0.56-1.02], \( z = -1.83, p = 0.068 \)).

CI = confidence interval; RR = risk ratio.
Figure 8: Funnel plot (observed and imputed values) for the primary outcome, Any Major Complication. Five studies²⁰-²³, ²⁶ included.