Effect of Using a Safety Checklist on Patient Complications after Surgery

A Systematic Review and Meta-analysis

Brigid M. Gillespie, Ph.D., R.N., Wendy Chaboyer, Ph.D., R.N., Lukman Thalib, Ph.D., B.Sc., Melinda John, B.Sc., Nicole Fairweather, F.A.N.Z.C.A., M.D., Kellee Slater, F.R.A.C.S., M.D.

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 to identify studies that evaluated the effects of checklist use in surgery on complication rates. Study quality was assessed using the Methodological Index for Nonrandomized Studies. The pooled risk ratio (RR) was estimated using both fixed and random effects models. For each outcome, the number needed to treat (NNT) and the absolute risk reduction (ARR) were 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 to 0.72; $P < 0.0001$; ARR, 3.7%; NNT, 27) and wound infection (RR, 0.54; 95% CI, 0.40 to 0.72; $P = 0.0001$; ARR, 2.9%; NNT, 34) and also reduction in blood loss (RR, 0.56; 95% CI, 0.45 to 0.70; $P = 0.0001$; ARR, 3.8%; NNT, 33). There were no significant reductions in mortality (RR, 0.79; 95% CI, 0.57 to 1.11; $P = 0.191$; ARR, 0.44%; NNT, 229), pneumonia (RR, 1.03; 95% CI, 0.73 to 1.4; $P = 0.857$; ARR, 0.04%; NNT, 2,512), or unplanned return to operating room (RR, 0.75; 95% CI, 0.56 to 1.02; $P = 0.068$; ARR, 0.52%; NNT, 192).

Conclusion: Notwithstanding the lack of randomized controlled trials, synthesis of the existing body of evidence suggests a relationship between checklist use in surgery and fewer postoperative complications. **(ANESTHESIOLOGY 2014; 120:1380-9)**

OVER the last decade, checklists have become commonplace in healthcare practice as a strategy to improve patient safety. Surgical checklists emphasize several salient components of patient safety: safe anesthesia 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 (SSC) has heralded a legislative mandate for the implementation of checklists in operating rooms (ORs) in over 122 countries.† Thus, checklists have become synonymous as best practice in high-risk areas such as surgery.[2–4]

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 characterizes 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

What We Already Know about This Topic

- Before-and-after studies suggest that use of perioperative safety checklists reduces complication rates

What This Article Tells Us That Is New

- The investigators conducted a meta-analysis that included seven 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

This article is featured in "This Month in Anesthesiology," page 1A.

Submitted for publication July 29, 2013. Accepted for publication December 17, 2013. From the National Health and Medical Research Council Research Centre of Research Excellence in Nursing, Research Centre for Health Practice Innovation, Griffith Health Institute, Griffith University, Queensland, Australia (B.M.G., W.C.); Department of Community Medicine, Faculty of Medicine, Kuwait University, Safat, Kuwait (L.T.); Research Centre for Health Practice Innovation, Griffith Health Institute, Griffith University, Queensland, Australia (M.J.); Department of Anaesthetics, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia (N.F.); and Department of Hepatobiliary, Liver Transplant and General Surgery, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia (K.S.).

* Press Release, World Health health Organization (WHO): New checklist to help make surgery safer. Available at: http://www.who.int/mediacentre/news/releases/2008/pr20/en/. Accessed April 24, 2013.

† World Health Organization Patient Safety Web Map. Available at: http://maps.cga.harvard.edu:8080/Hospital/. Accessed April 24, 2013.

 Anesthesiology 2014; 120:1380-9

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 data.[1,8,11] 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 the 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 database functions such as "apply related words" and "explosion" were used to maximize the search. Reference lists of retrieved articles 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 articles if they were (1) not written in English and (2) used a checklist that was not perioperative-specific or encompassed other clinical settings (*e.g.*, SURgical Patient Safety System).‡

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 were 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."[13] In this review, definitions of the primary outcome, namely, "any major complication" were informed by the American College of Surgeons National Quality Improvement Program and/or the Clavien Classification system[14] or however defined by the primary study authors. Secondary outcomes included individual complications such as mortality, surgical site infection (SSI), pneumonia, wrong-site surgery, unplanned return to OR, and blood loss more than 500 ml (*i.e.*, preoperative blood loss of 500 ml more than expected for a given case)[15] or however defined by the primary study authors. A surgical checklist was defined as a cognitive tool for delineating team tasks according to preinduction, before skin incision, and before leaving the OR and used to prompt interdisciplinary verification of surgical and patient details to increase teamwork.[2,4,9] Standard care was defined as the usual practices undertaken by the interdisciplinary surgical teams before 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.[16] The original version of the Methodological Index for Nonrandomized Studies is a 12-item index. Each item is scored as 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, that is, 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 (B.M.G. and M.J.) independently assessed study quality, and the proportion of agreement was measured using the intraclass correlation coefficient. A coefficient of 0.70 or greater 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*-squared (I^2)

‡ The SURPASS checklist is far more comprehensive and interdisciplinary; however, it is also more complex to implement and often requires reorganization of care processes.

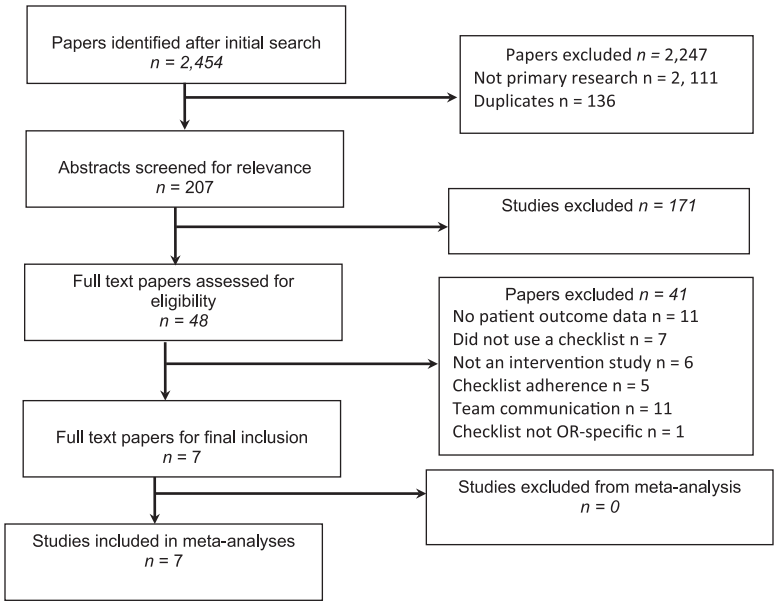

Fig. 1. Preferred reporting items for systematic reviews and meta-analyses; flow diagram[19] of literature search. OR = operating room. Adapted from Moher *et al.* PLoS Med 2009; 6:e1000097. Adaptations are themselves works protected by copyright. So in order to publish this adaptation, authorization must be obtained both from the owner of the copyright in the original work and from the owner of copyright in the translation or adaptation.

and Cochrane *Q* statistic (chi-square test) for heterogeneity. The significance level of the chi-square test was set at *P* value less than 0.05. The degree of statistical heterogeneity was determined if the I^2 was 50% or greater. In cases of statistical heterogeneity, only random effects models were used. Results of the meta-analysis were presented using risk ratio and corresponding 95% CI and the *z* test for pooled effect size estimates. For each outcome, the number needed to treat (NNT) and absolute risk reduction (ARR) were calculated from the pooled risk ratio and the event rates to further place risk estimates into context. Forrest plots were used for graphical display of results.

To assess the robustness of the results, sensitivity analyses were performed.[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[12] based on the primary outcome, that is, 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, seven (3.4%) studies[20–26] with a total of 37,339 patients met eligibility criteria. The included studies were published between 2009 and 2012. The 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 the study by van Klei *et al.*[24] who used a retrospective cohort design. Most studies (six of seven, 86%) used the WHO Surgical Safety Checklist[27]; however, the study by Bliss *et al.*[21] used the *Association of peri-Operative Registered Nurses Comprehensive Surgical Checklist*,§ which includes items based on the WHO Checklist* and the Joint Commission Universal Protocol. The studies by Haynes *et al.*[22] and Weiser *et al.*,[25] although based on the same multinational study, focused 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 reported by Haynes *et al.*[22] The study by Weiser *et al.*[25] was included in the analysis of only one outcome, expected blood loss more than 500 ml (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 to 0.96; $P = 0.008$). Out of a possible score of 20 on the modified 10-item Methodological Index for Nonrandomized Studies tool,[16] the mean score for the seven included studies was 12.6 (±1.3; range, 11 to 15). Four studies (57%) achieved a score of 13/20 or greater.

§ Association of periOperative Registered Nurses (AORN). Comprehensive Checklist 2013. Available at: http://www.aorn.org/Clinical_Practice/ToolKits/Correct_Site_Surgery_Tool_Kit/Comprehensive_checklist.aspx. Accessed March 3, 2013.

Table 1. Characteristics and Results of Review Studies

Author/Year/ Country	Study Design	Type of Checklist	n/Type of Procedure	Mean Age	Sex, M/F	Outcome Measures	Results: Pre/Post, %, *P* Value
Askarian *et al.* (2011)[20], Iran	Prospective cohort using historical controls	WHO (19 items)	294/General procedures	Pre: 45 ± 10.9; Post: 45.7 ± 11.8	73/71, 58/92	SSI, pneumonia, acute renal failure, any complication	Pre: 10.4, Post: 5.3, $P < 0.1$; Pre: 7.6, Post: 3.3, $P = 0.17$; Pre: 4.9, Post: 2.6, $P = 0.17$; Pre: 22.0, Post: 10.0, $P = 0.03$
Bliss *et al.* (2012)[21], United States	Prospective cohort using historical controls	AORN adaption of WHO and JCUP checklists (29 items)	2,398/Specific high-risk procedures	Pre: 54.5 ± 17.6; Post: 54.8 ± 15.7; Control: 54.6 ± 17.7	99/147, 30/40, 827/1,248	SSI, pneumonia, bleeding, DVT/PE, ventilator >48 h, UTI, any complication	Pre: 5.3, Post: 5.5, $P = 0.845$; Pre: 2.8, Post: 0.0, $P = 0.362$; Pre: 4.9, Post: 2.7, $P = 0.392$; Pre: 2.0, Post: 0.0, $P = 0.074$; Pre: 3.3, Post: 0.0, $P = 0.362$; Pre: 2.8, Post: 2.7, $P = 0.972$; Pre: 15.9, Post: 8.2, $P = 0.000$
Haynes *et al.* (2009)[22], Multinational	Prospective cohort using historical controls	WHO (19 items)	7,688/All procedures	Pre: 46.8 ± 18.1; Post: 46.7 ± 17.9	1,635/2,098, 1,677/2,278	SSI, pneumonia, unplanned return to OR, mortality, any complication	Pre: 6.2, Post: 3.4, $P < 0.001$; Pre: 1.1, Post: 1.3, $P = 0.46$; Pre: 2.4, Post: 1.8, $P = 0.047$; Pre: 1.5, Post: 0.8, $P = 0.003$; Pre: 11.0, Post: 7.0, $P < 0.001$
Sewell *et al.* (2011)[23], United Kingdom	Prospective cohort using historical controls	WHO (19 items)	965/Orthopedic procedures	Pre: 56.7; Post: 53.2	227/253, 248/237	SSI, pneumonia, unplanned return to OR, mortality, any complication	Pre: 4.4, Post: 3.5, P = NR; Pre: 2.1, Post: 2.5, P = NR; Pre: 1.0, Post: 1.0, P = NR; Pre: 1.9, Post: 1.6, P = NR; Pre: 8.5, Post: 7.6, P = NR
van Klei *et al.* (2012)[24], The Netherlands	Retrospective cohort	WHO (22 items)	25,513/Nonday case procedures	Pre: 53.7 ± 17.9; Post: 54.2 ± 17.7	7,380/6,982, 5,861/5,290	Mortality (crude)	Pre: 3.1, Post: 2.8, $P = 0.19$
Weiser *et al.* (2010)[25], Multinational*	Prospective cohort using historical controls	WHO (19 items)	1,750/Urgent procedures	Pre: 37.0 ± 17.9; Post: 36.3 ± 17.5	322/520, 333/575	Expected blood loss, SSI, mortality, any complication	Pre: 11.2, Post: 6.6, $P < 0.0008$; Pre: 20.2, Post: 13.2, $P < 0.0001$; Pre: 3.7, Post: 1.4, $P < 0.0067$; Pre: 18.4, Post: 11.7, $P < 0.0001$
Yuan *et al.* (2012)[26], Liberia	Prospective cohort using historical controls	WHO (19 items)	481/All procedures	Pre: NR; Pre: NR	142/90, 170/79	SSI, mortality, any complication	Pre: 28.6, Post: 9.9, $P < 0.0001$; Pre: 2.2, Post: 2.8, $P < 0.334$; Pre: 32.95, Post: 19.1, $P < 0.005$

* Subset of data reported in relation to urgent procedures included in the study by Haynes *et al*. The study by Weiser *et al.* was included in the meta-analysis of one outcome not measured by Haynes *et al*., that is, expected blood loss of ≥500 ml.

AORN = Association of periOperative Registered Nurses; DVT = deep vein thrombosis; JCUP = joint universal protocol; NR = not reported; OR = operating room; PE = pulmonary embolism; post = postoperative; pre = preoperative; SSI = surgical site infection; UTI = urinary tract infection; WHO = World Health Organization.

Primary Outcome

Any Major Complication. Five published studies[20–23,26] with 9,747 patients compared checklist use relative to this outcome that combined single complications that included up to 18 individual complications (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, and acute renal failure). As there was no significant heterogeneity among studies (chi-square = 4.37, *df* = 4, *P* = 0.357, I^2 = 8.6%), we used the fixed effect model (fig. 2). Pooling of individual studies indicated that 381 of 4,922 (7.7%) patients in the intervention group and 592 of 4,835 (12.2%) patients in the control group developed any complication. The ARR was 3.7% (95% CI, 2.6 to 4.8%), and the NNT was 27 (95% CI, 21 to 38). That is, one less major complication may be prevented for every 27 patients when 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 of 15,840 (2.3%) patients in the intervention group and 520 of 18,807 (2.8%) patients in the control group died. As there was significant heterogeneity (chi-square = 5.58, *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 to 0.76%), and the NNT was 229 (95% CI, 131 to 867), suggesting that one less death may be prevented for every 229 patients when the checklist

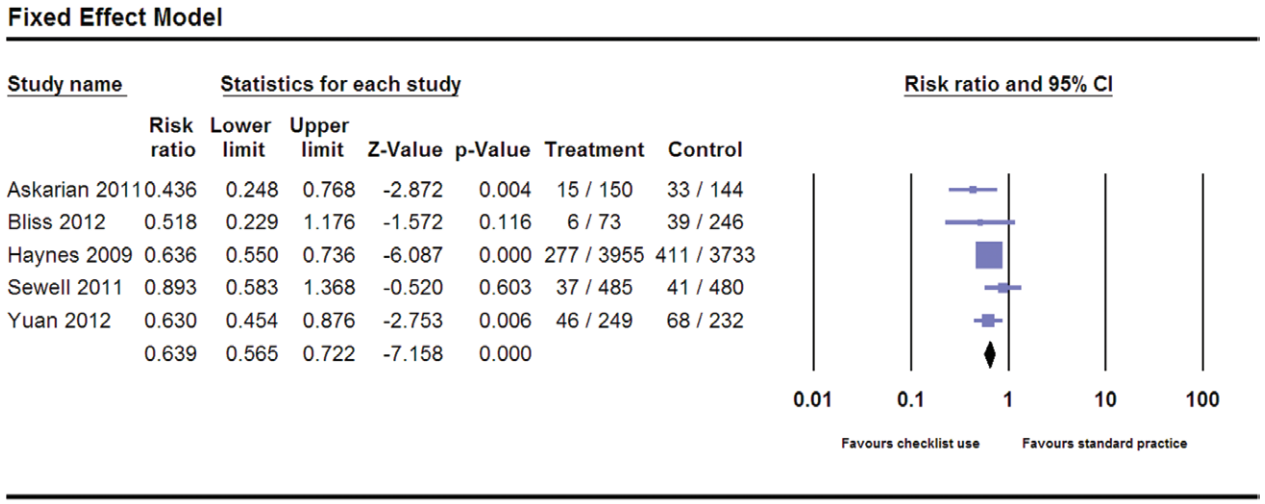

Fig. 2. Forrest plot comparing checklist use with standard practice for any major complication; five studies[20–23,26] have been included (risk ratio, 0.63 [95% CI, 0.56–0.72], *z* = –7.16, *P* < 0.0001).

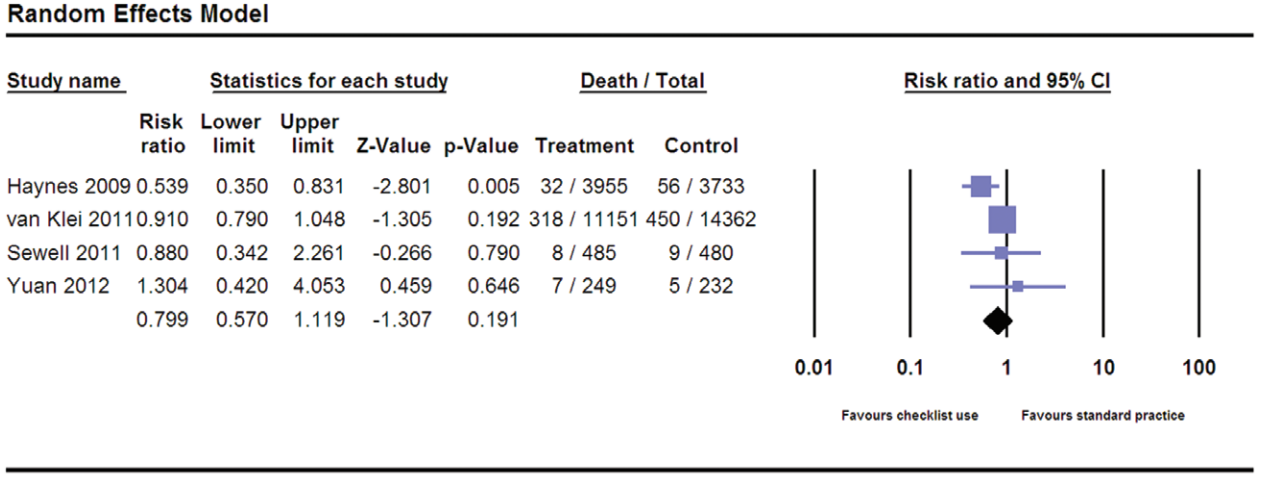

Fig. 3. Forrest plot comparing checklist use with standard practice for patient mortality; four studies[22–24,26] have been included (risk ratio, 0.79 [95% CI, 0.57–1.11], *z* = –1.30, *P* = 0.191).

is used. However, in three of four included studies, the CIs cross the line of no difference, which suggests that the use of a checklist is not associated with a reduction in mortality.

Surgical Site Infection. Five published studies[20–23,26] with 9,747 patients assessed this outcome. As there was significant heterogeneity between studies (chi-square = 6.50, $df = 4$, $P = 0.0001$, $I^2 = 38.4\%$), a random effects model was used (fig. 4). Pooling of individual studies showed that 188 of 4,912 (3.8%) patients in the intervention group and 344 of 4,835 (7.1%) patients in the control group developed a SSI. The ARR was 2.9% (95% CI, 2.1 to 3.8%), and the NNT was 34 (95% CI, 26 to 47). That is, one less SSI may be prevented for every 34 patients when 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-square = 11.4, $df = 4$, $P = 0.022$, $I^2 = 64.3\%$; fig. 5), a fixed effects model was used. Pooled analysis revealed that 69 of 4,663 (1.5%) patients in the intervention group and 70 of 4,603 (1.5%) patients in the control group developed pneumonia. The ARR was 0.04% (95% CI, −0.4 to 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 with that in the standard care group, which exceeds 224 (95% confidence). In these four studies, the CIs cross the line of no difference, suggesting that the use of a checklist is not associated with a reduction in pneumonia.

Expected Blood Loss 500 ml or Greater. Two published studies[21,25] with 2,069 patients assessed this outcome. As there was no significant heterogeneity between studies (chi-square = 9.38, $df = 1$, $P =$, $I^2 = 0.0\%$; fig. 6), a fixed effects model was used. Pooling of individual studies showed

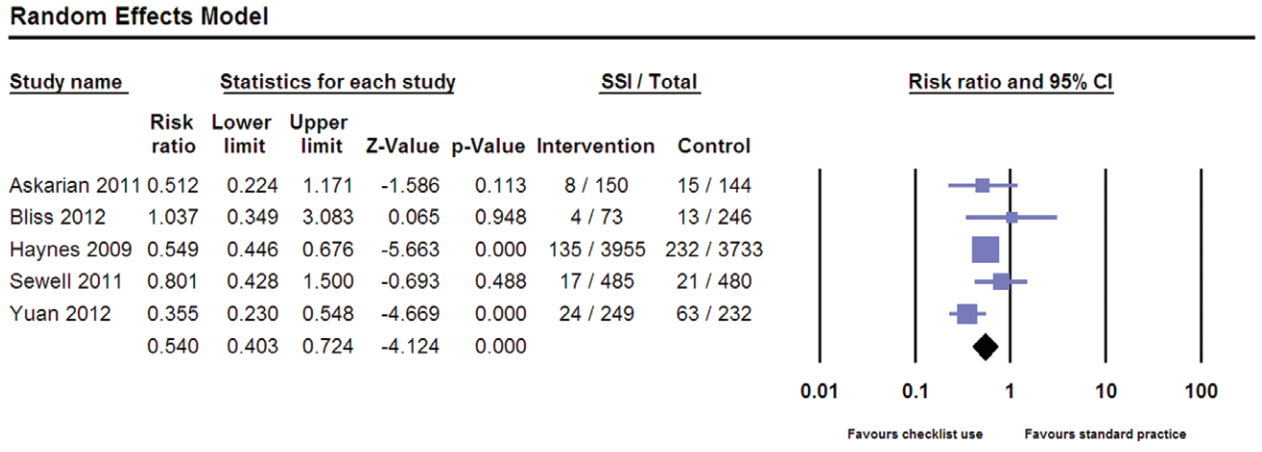

Fig. 4. Forrest plot comparing checklist use with standard practice for surgical site infection (SSI) complication; five studies[20–23,26] have been included (risk ratio, 0.54 [95% CI, 0.40–0.72], $z = -4.12$, $P < 0.0001$).

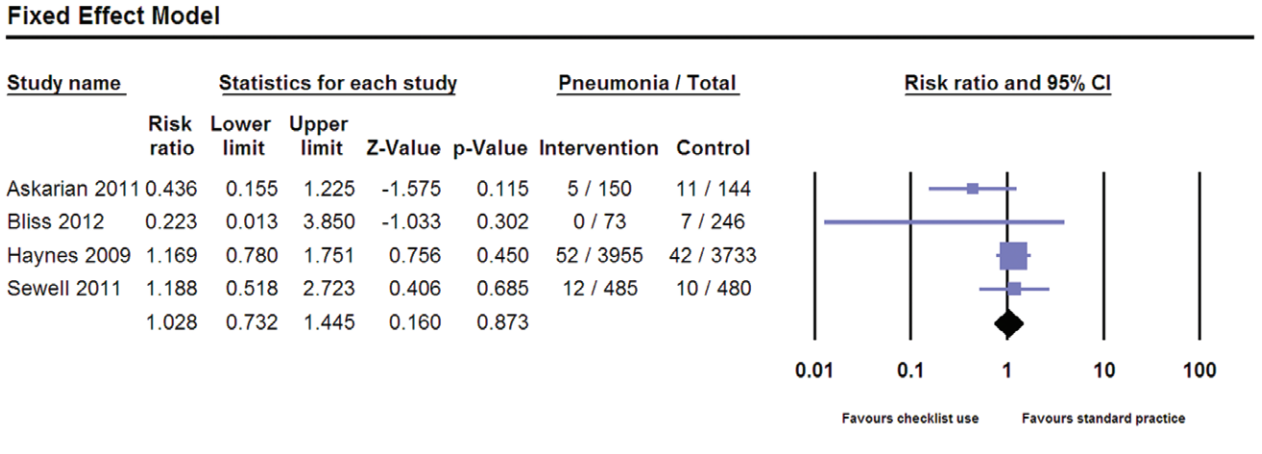

Fig. 5. Forrest plot comparing checklist use with standard practice for pneumonia complication; four studies[20–23] have been included (risk ratio, 1.03 [95% CI, 0.73–1.45], $z = 0.180$, $P = 0.857$).

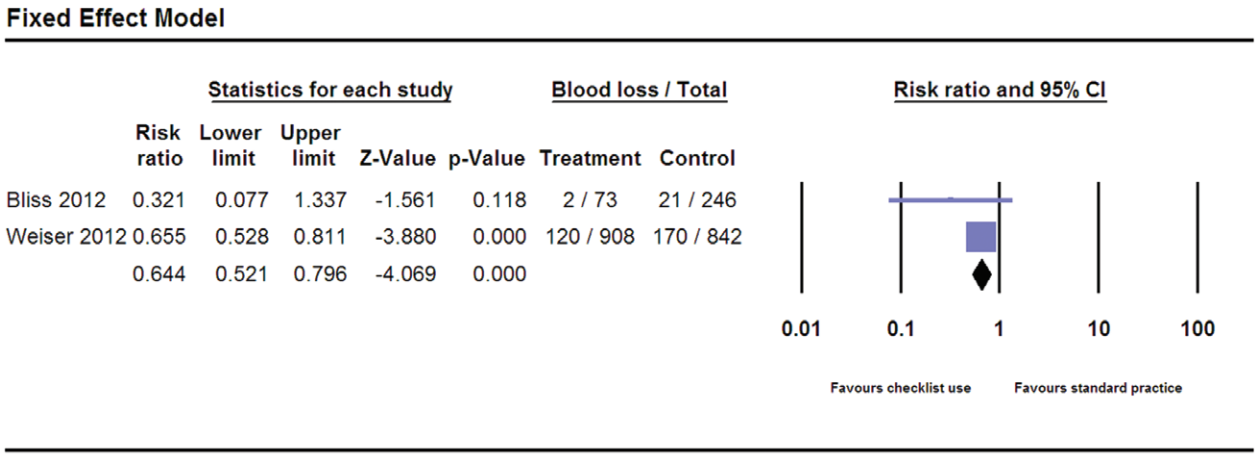

Fig. 6. Forrest plot comparing checklist use with standard practice for expected blood loss complication; two studies[21,25] have been included (risk ratio, 0.64 [95% CI, 0.52–0.80], $z = -4.07$, $P < 0.0001$).

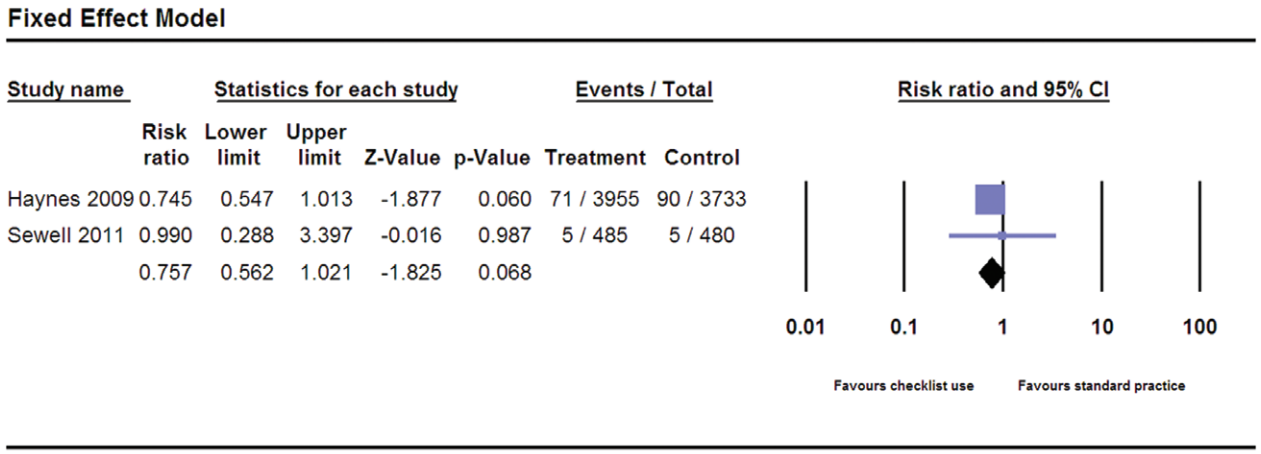

Fig. 7. Forrest plot comparing checklist use with standard practice for unplanned return to operating room (OR); two studies[22,23] have been included (risk ratio, 0.75 [95% CI, 0.56–1.02], $z = -1.83$, $P = 0.068$).

that 122 of 981 (12.4%) patients in the intervention group and 182 of 1,088 (16.7%) patients in the control group had experienced bleeding requiring 4 units of blood or greater within 72 h after surgery. The ARR was 3.8% (95% CI, 0.6 to 5.9%), and the NNT was 33 (95% CI, 17 to 168). Therefore, one less patient may have blood loss more than 500 ml for every 31 patients when 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-square = 0.193, $df = 1$, $P = 0.661$, $I^2 = 0.0\%$; fig. 7), we used a fixed effects model. Pooling of individual studies showed that 76 of 4,440 (1.7%) patients in the intervention group and 95 of 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 to 1.1%), and the NNT was 192, suggesting that unplanned return to OR may be avoided for every 192 patients when the checklist is used. The number of patients required in the checklist group exceeds 91, compared with that in the standard care group, which exceeds 1,825 (95% confidence). In the two included studies for this outcome, the CIs 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 test ($z = 0.939$, $P = 0.347$) and Egger 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,

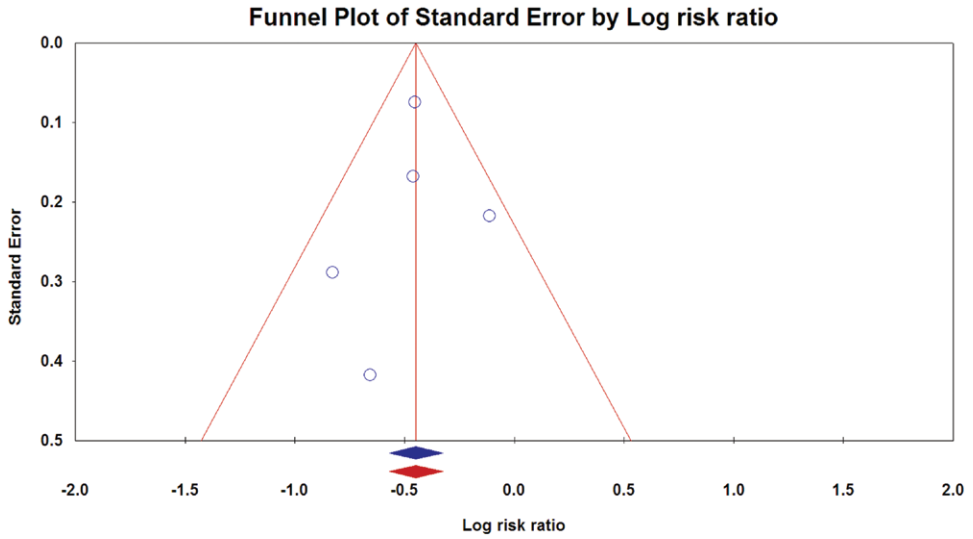

Fig. 8. Funnel plot (observed and imputed values) for the primary outcome, any major complication; five studies[20–23,26] have been included.

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.*[28] that 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 our and Borchard's[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.[14] That the authors[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[20–23,26] and expected blood loss.[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[27] addresses the risk of blood loss, whereas items 12 of 13 (time out) require confirmation of equipment availability/instrument sterility and antibiotic prophylaxis. Hospital-acquired infections are considered preventable—and therefore, human error may contribute to SSI.[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, thereby reducing the potential for human error.[24,29]

In our analyses, the nonsignificant results relative to individual complications of pneumonia,[20–23] patient mortality, and unplanned return to OR[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 who developed pneumonia or died, a significant reduction would have required a much larger number of patients, which is unlikely given the potential effect size (*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 before 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.[2,3] There is little doubt that the culture of surgery is changing, and the importance of human factors is increasingly being recognized.[30] Finally, there is a danger in treating the checklist as a "tick box" exercise, rather than emphasizing the need for clinicians to engage in the process.[4,11,29]

Methodological Limitations

We acknowledge some limitations. Although meta-analysis is powerful—it is also controversial because small violations in meeting critical assumptions can give misleading results.[31] 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[16] and that is 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 of 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.[31] In some of the reviewed studies,[20,26] insufficient sample size, selection bias, and 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.[12] 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, and 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 postoperative 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,§ 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 (1) the number of occasions the checklist was used (per patient) after its introduction, and (2) 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 judgment 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.

Randomized controlled trial methods are likely to have increased internal validity but they cannot answer all important questions for a given intervention.[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.[29,32] Checklists facilitate an exchange or clarification of information among team members to avert sentinel events in surgery.[1,3,5] Consequently, checklist use supports a safety culture,[33] and it is viewed as best practice.[2,4,29] As such, it would be inappropriate to withhold the use of checklists in clinical practice rendering RCT methods unfeasible. Although 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 rigor used to identify and evaluate available studies. The *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* guidelines[19] were adhered to, and we utilized an appropriate, validated tool to assess study quality.[16] The reviewed studies incorporated 11 countries (*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 the 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.

Acknowledgments

The authors thank the financial assistance of the Australian Research Council, Early Career Discovery Fellowship Scheme, Canberra, ACT, Australia.

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Gillespie: National Health and 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. b.gillespie@griffith.edu.au.

References

1. Pronovost PJ, Goeschel CA, Marsteller JA, Sexton JB, Pham JC, Berenholtz SM: Framework for patient safety research and improvement. Circulation 2009; 119:330–7
2. Weiser TG, Berry WR: Review article: Perioperative checklist methodologies. Can J Anaesth 2013; 60:136–42
3. Spiess BD: The use of checklists as a method to reduce human error in cardiac operating rooms. Int Anesthesiol Clin 2013; 51:179–94
4. Treadwell J, Lucas S, Tsou A: Surgical checklists: A systematic review of impacts and implementation. BMJ Qual Saf 2014; 23:299–318
5. Makary MA, Holzmueller CG, Thompson D, Rowen L, Heitmiller ES, Maley WR, Black JH, Stegner K, Freischlag JA, Ulatowski JA, Pronovost PJ: Operating room briefings: Working on the same page. Jt Comm J Qual Patient Saf 2006; 32:351–5
6. Lingard L, Espin S, Rubin B, Whyte S, Colmenares M, Baker GR, Doran D, Grober E, Orser B, Bohnen J, Reznick R: Getting teams to talk: Development and pilot implementation of a checklist to promote interprofessional communication in the OR. Qual Saf Health Care 2005; 14:340–6
7. Bleakley A, Boyden J, Hobbs A, Walsh L, Allard J: Improving teamwork climate in operating theatres: The shift from multiprofessionalism to interprofessionalism. J Interprof Care 2006; 20:461–70
8. Lingard L, Regehr G, Cartmill C, Orser B, Espin S, Bohnen J, Reznick R, Baker R, Rotstein L, Doran D: Evaluation of a preoperative team briefing: A new communication routine results in improved clinical practice. BMJ Qual Saf 2011; 20:475–82
9. Makary MA, Sexton JB, Freischlag JA, Holzmueller CG, Millman EA, Rowen L, Pronovost PJ: Operating room teamwork among physicians and nurses: Teamwork in the eye of the beholder. J Am Coll Surg 2006; 202:746–52
10. Gillespie BM, Chaboyer W, Longbottom P, Wallis M: The impact of organisational and individual factors on team communication in surgery: A qualitative study. Int J Nurs Stud 2010; 47:732–41
11. Fourcade A, Blache JL, Grenier C, Bourgain JL, Minvielle E: Barriers to staff adoption of a surgical safety checklist. BMJ Qual Saf 2012; 21:191–7
12. Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. Edited by Higgins J, Green S. Chichester, Wiley-Blackwell & Sons, 2011, pp 78–95
13. Sokol DK, Wilson J: What is a surgical complication? World J Surg 2008; 32:942–4
14. Clavien PA, Barkun J, de Oliveira ML, Vauthey JN, Dindo D, Schulick RD, de Santibañes E, Pekolj J, Slankamenac K, Bassi C, Graf R, Vonlanthen R, Padbury R, Cameron JL, Makuuchi M: The Clavien-Dindo classification of surgical complications: Five-year experience. Ann Surg 2009; 250:187–96
15. Liumbruno GM, Bennardello F, Lattanzio A, Piccoli P, Rossetti G; Italian Society of Transfusion Medicine and Immunohaematology (SIMTI) Working Party: Recommendations for the transfusion management of patients in the peri-operative period. II. The intra-operative period. Blood Transfus 2011; 9:189–217
16. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J: Methodological index for non-randomized studies (minors): Development and validation of a new instrument. ANZ J Surg 2003; 73:712–6
17. Polit D: Statistics and Data Analysis for Nursing Research, Second edition. Upper Saddle River: Pearson, 2010, pp 197–22
18. Mahid SS, Hornung CA, Minor KS, Turina M, Galandiuk S: Systematic reviews and meta-analysis for the surgeon scientist. Br J Surg 2006; 93:1315–24
19. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009; 6:e1000097
20. Askarian M, Kouchak F, Palenik CJ: Effect of surgical safety checklists on postoperative morbidity and mortality rates, Shiraz, Faghihy Hospital, a 1-year study. Qual Manag Health Care 2011; 20:293–7
21. Bliss LA, Ross-Richardson CB, Sanzari LJ, Shapiro DS, Lukianoff AE, Bernstein BA, Ellner SJ: Thirty-day outcomes support implementation of a surgical safety checklist. J Am Coll Surg 2012; 215:766–76
22. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, Herbosa T, Joseph S, Kibatala PL, Lapitan MC, Merry AF, Moorthy K, Reznick RK, Taylor B, Gawande AA; Safe Surgery Saves Lives Study Group: A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med 2009; 360:491–9
23. Sewell M, Adebibe M, Jayakumar P, Jowett C, Kong K, Vemulapalli K, Levack B: Use of the WHO surgical safety checklist in trauma and orthopaedic patients. Int Orthop 2011; 35:897–901
24. van Klei WA, Hoff RG, van Aarnhem EE, Simmermacher RK, Regli LP, Kappen TH, van Wolfswinkel L, Kalkman CJ, Buhre WF, Peelen LM: Effects of the introduction of the WHO "Surgical Safety Checklist" on in-hospital mortality: A cohort study. Ann Surg 2012; 255:44–9
25. Weiser TG, Haynes AB, Dziekan G, Berry WR, Lipsitz SR, Gawande AA; Safe Surgery Saves Lives Investigators and Study Group: Effect of a 19-item surgical safety checklist during urgent operations in a global patient population. Ann Surg 2010; 251:976–80
26. Yuan CT, Walsh D, Tomarken JL, Alpern R, Shakpeh J, Bradley EH: Incorporating the World Health Organization Surgical Safety Checklist into practice at two hospitals in Liberia. Jt Comm J Qual Patient Saf 2012; 38:254–60
27. World Health Organization (WHO): Implementation of the Surgical Safety Checklist. Geneva, WHO Press, 2008, pp 1–28
28. Borchard A, Schwappach DL, Barbir A, Bezzola P: A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. Ann Surg 2012; 256:925–33
29. Winters BD, Gurses AP, Lehmann H, Sexton JB, Rampersad CJ, Pronovost PJ: Clinical review: Checklists—Translating evidence into practice. Crit Care 2009; 13:210
30. Low D, Walker I, Heitmiller ES, Kurth D: Implementing checklists in the operating room. Paediatr Anaesth 2012; 22:1025–31
31. Walker E, Hernandez AV, Kattan MW: Meta-analysis: Its strengths and limitations. Cleve Clin J Med 2008; 75:431–9
32. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, Robertson N: Tailored interventions to overcome identified barriers to change: Effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2010:CD005470
33. Gillespie BM, Gwinner K, Chaboyer W, Fairweather N: Team communications in surgery—Creating a culture of safety. J Interprof Care 2013; 27:387–93