Review

ANNALS OF SURGERY OPEN

OPEN

Maximizing Surgical Success by Aligning Interventions to Outcomes: A Systematic Review

Bonnie A. Armstrong, PhD,*† Arthur Tung, BSc,† Lisha Lo, MPH,‡ Spencer S. Abssy, BSc,§ Maham Zulfiqar, BSc,‖ Juliette van Oost, BSc,§ Julie Wong, BSc,¶ Josh Janevski, BSc,§ Julia Martyniuk, MI,# and Patricia Trbovich, PhD*†**

Objective: This study aimed to identify common intraoperative interventions in surgery and evaluate their effectiveness in improving surgical outcomes.

Background: Despite decades of efforts, surgical adverse events remain stubbornly high. There are concerns that too much responsibility is placed on individuals to create change (ie, person-based interventions) rather than adapting systems to support human performance (ie, system-based interventions). This focus may be due to our limited understanding of which interventions most effectively improve outcomes.

Methods: A 2-step search was conducted. Systematic and meta-analytic reviews of Medline, CINAHL, Embase, PsycINFO, Scopus, Cochrane Reviews, Cochrane Protocols and Cochrane Trials were identified, and individual studies within these reviews were selected. Qualitative content analysis categorized intervention and outcome types using inductive and deductive methods. Intervention details and directional findings for all outcomes were extracted.

Results: A total of 575 studies were included in the final analysis comprising 5,288,513 cases, 25,435 providers and patients, 2608 hospitals, across 50 countries, with 1221 outcomes extracted. Overall, the most common interventions were person-based, including education (38%) and policy (19%). Person-based interventions were more likely to improve interpersonal outcomes such as culture, professional development, and resilience. In contrast, system-based interventions, such as technology (15%), cognitive aids (11%), equipment (11%), standardization (4%), and environment redesign (2%), though less frequently implemented, were effective across all outcome types.

Conclusions: Although person-based interventions are widely implemented, system-based interventions generally have a greater impact on surgical outcomes. These results offer valuable insights for optimizing the alignment of interventions to outcomes.

Keywords: intervention effectiveness, intraoperative interventions, person-based interventions, surgical outcomes, system-based interventions

INTRODUCTION

Unintentional harm to patients undergoing surgery, such as retained foreign objects or wrong site surgery, is a leading contributor to preventable injury and death. In fact, preventable intraoperative adverse events have been on the rise despite the implementation of mitigation strategies, which can only partly be attributed to improved reporting of surgical errors.[1,2] Although there is more opportunity to err due to the increase in surgical volumes and the complexity of surgical care, adverse events also stem from the implementation of interventions that do not adequately address the root causes of problems in surgery and instead introduce new risks.[3,4] Further, the methodologies used to investigate surgical error and the interventions chosen to address their causes vary widely,[5,6] resulting in extensive variation in intervention effectiveness.[7]

There are 2 main approaches to improving patient safety in surgery: person-based interventions and system-based interventions.[4] Person-based interventions focus on the individual, emphasizing responsibility for remembering and executing tasks, typically through training or reinforcing policies.[3,4] In contrast, system-based interventions are designed to support or guide individuals by optimizing the broader healthcare infrastructure. These may involve standardized protocols, operating room (OR) design, and ergonomic improvements in surgical equipment.[3,4] Such interventions provide a nudge or prompt to ensure tasks are completed correctly. While these 2 approaches are frequently used together, the distinction between them remains useful for understanding their different functions. Person-based interventions target individual's role in task execution, whereas system-based interventions create external prompts or structures to support correct behavior. For example, a surgical safety

From the *Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada; †Humanera, Office of Research and Innovation, North York General Hospital, Toronto, Ontario, Canada; ‡Centre for Quality Improvement and Patient Safety, University of Toronto, Toronto, Ontario, Canada; §Human Biology, University of Toronto, Toronto, Ontario, Canada; ‖Department of Psychology, University of Toronto, Scarborough, Ontario, Canada; ¶Department of Physical and Environmental Sciences, University of Toronto, Toronto, Ontario, Canada; #Gerstein Science Information Centre, University of Toronto, Toronto, Ontario, Canada; and **Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada.

This review was funded by the Canadian Institutes of Health Research (CIHR) Grant (#168021 2020/10).

The authors declare that they have nothing to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

PROSPERO registration number: CRD42023355257.

SDC *Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.annalsofsurgery.com).*

Reprints: Bonnie A. Armstrong, PhD, University of Toronto, Institute of Health Policy, Management and Evaluation, 155 College Street, Suite 425, Toronto, ON M5T 3M6, Canada. E-mail: bonniearmstrong03@gmail.com.

Annals of Surgery Open (2025) 1:e558

Received: 26 September 2024; Accepted 2 February 2025

Published online 7 March 2025

DOI: 10.1097/AS9.0000000000000558

checklist is a system-based intervention that reminds the surgical team of critical tasks, but it often requires person-based components such as training or policy enforcement to be effective. Thus, although interventions may combine both approaches, recognizing their distinct roles helps clarify their individual contributions to improving safety. Each approach addresses different aspects of the problem—individual responsibility versus system-level support—and understanding their differences can guide more targeted and effective interventions.

When recurring issues arise in surgery, there is a strong tendency to quickly implement a person-based intervention without fully understanding the system in which the errors occur.[8,9] According to the hierarchy of effectiveness framework, person-based interventions are considered weaker corrective actions compared with system-based interventions because they fail to remove the hazard and focus instead on warning individuals about the hazard.[3] For example, Kellogg et al[10] found that over an 8-year period at an American academic medical center, person-based interventions such as education, training, process changes, and policy reinforcement were the most commonly implemented interventions in 302 root-cause analyses. Despite these efforts, critical incidents such as retained foreign objects continued to recur, suggesting the inappropriate application of these interventions to resolve these types of adverse events. Narrowly focusing on person-based corrective actions overlooks the many contributing safety threats deeply engrained within the sociotechnical system. The higher frequency of person- over system-based interventions is largely due to their lower cost and the ease with which they can be implemented. Developing system-based interventions such as automation and forcing functions tends to be costly and requires active monitoring for implementation issues.[11]

System-based interventions generally address underlying causal factors more effectively than person-based interventions, such as education and policy reinforcement. However, certain surgical outcomes can still be significantly improved by person-based interventions. For instance, trainees who are monitored by senior surgeons and given real-time feedback either in person or via telementoring,[12] as well as video-based education approaches have shown to improve a variety of technical surgical skills.[13] In another study, Englemann et al[14] showed that a new policy that instructed surgeons to take a 5-minute unstructured break after every 25-minute work period was associated with a significant decrease in the surgeon's stress levels and adverse events for complex laparoscopic general surgery procedures in children. In light of this research, it may be an oversimplification to associate person-based approaches with ineffective improvement strategies. Rather, both person- and system-based interventions may be effective if aligned to the appropriate surgical outcomes in need of improvement.

Although previous research has described the effectiveness of surgical interventions on clinical outcomes, these investigations primarily assess the effectiveness of a single or a small subset of interventions and outcomes.[15,16] Existing research has yet to comprehensively map the range of intraoperative surgical interventions to the specific outcomes they are most likely to improve. Therefore, the objectives of this systematic review were to (1) identify common types of intraoperative interventions and determine their prevalence and (2) evaluate the impact of these intervention types on surgical outcomes.

METHODS

To identify common surgical interventions, we reviewed relevant studies from published systematic reviews and meta-analyses. This approach was chosen because conducting a typical systematic literature search for all interventions enhancing intraoperative performance would be impractical and would include interventions with very minimal research that falls outside the scope of this review. This approach to study identification is reported in the Cochrane Handbook for Systematic Reviews of Interventions for similar purposes.[17]

Protocol and Registration

This systematic review was conducted in accordance with the Cochrane Handbook for Systematic Reviews of Interventions[17] and reported in accordance with the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement[18] and the 'Enhancing transparency in reporting the synthesis of qualitative research' statement[19] (seeSupplemental Digital Content, http://links.lww.com/AOSO/A480, for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses and Enhancing transparency in reporting the synthesis of qualitative research checklists). It was a preplanned, comprehensive search of published literature within systematic reviews and meta-analyses that described the effectiveness of intraoperative interventions. The review was prospectively registered and updated using PROSPERO (CRD42023355257).

Search Strategy

A comprehensive search strategy was developed by an information specialist and health science librarian (J.M.) in consultation with the review team to identify relevant peer-reviewed literature. The search strategy was peer-reviewed following the Peer Review of Electronic Search Strategies. Subject headings and text words related to the following concepts were searched: surgery, quality improvement initiatives, and systematic reviews. The Scottish Intercollegiate Guidelines Network's search filter for systematic reviews was adapted for the systematic review search concept.[20] The search was developed and finalized in Ovid MEDLINE, before being translated to Ovid Embase, Ovid PsycInfo, EBSCO CINAHL, Wiley Cochrane, and Scopus. The search was limited to exclude animal-only studies and restricted to articles published after 2019. No language limit was applied to the search. The searches were run on January 12, 2023. The search results were exported into Endnote to deduplicate records, and then imported into Covidence for screening.

Eligibility Criteria

Inclusion criteria were (1) OR team members (eg, surgeons, anesthesiologists, physician assistants, perfusionists, scrub and circulating nurses, or interventional radiologists), (2) at the staff or trainee level, (3) who treat human patients, (4) who are participating in any type of intraoperative intervention or quality improvement initiative (or a control version of an intervention) aimed at improving any surgical outcome, and (5) with any type of study design examined. Exclusion criteria were (1) animal studies, (2) commentary/opinion papers, (3) letters to the editor, (4) non-OR setting, (5) nonintervention studies, (6) nonsurgical populations, (7) no reporting of surgical outcomes, and (8) studies focused solely on improving technical skills through simulations were excluded due to the well-established link between simulations and technical skill enhancement.[21–23] However, studies that used simulations to improve both technical skills and other surgical outcomes, such as nontechnical skills (NTSs), were included, because they assessed a broader range of outcomes.

Selection Process

Four rounds of screening across 2 phases were conducted, in which relevant reviews were identified. Individual studies that met our inclusion criteria were then identified within those reviews.

Review-Level Screening

Three pairs of screeners (L.L., S.A., M.Z., J.O., J.W., and J.J.) independently reviewed titles/abstracts and full texts in 2 separate rounds of study selection and then inductively extracted the data in duplicate from the included reviews into a preformatted Excel worksheet (Microsoft, Redmond, Washington, USA, V.16.0.14131.20278). Discrepancies between pairs of screeners were resolved through arbitration by a third screener.

Study-Level Screening

Subsequent to review-level screening, all studies within each included review were then screened by the same 3 pairs of screeners (L.L., S.A., M.Z., J.O., J.W., and J.J.). The title/abstract and full text for each study were screened in 2 separate rounds of study selection and then screeners inductively extracted the data in duplicate from the included studies into a preformatted Excel worksheet (Microsoft, Redmond, Washington, USA, V.16.0.14131.20278). Discrepancies between pairs of screeners were resolved through arbitration by a third screener. Deduplication at the study level was manually performed during screening.

Data Extraction

Study characteristics, intervention descriptions, and surgical outcomes were extracted from all included studies. If a study implemented multiple interventions (eg, introducing a new surgical safety checklist [intervention 1] alongside a policy requiring all surgical staff to use it [intervention 2]), we categorized the intervention based on its dominant component, in line with the study's primary goals. For example, if the goal was to ensure that all necessary safety checks were completed in a consistent manner, the checklist would be categorized as the main intervention, with the policy serving as a supporting measure to ensure compliance. Study characteristics included: author name, year of publication, title, sample size, participant gender, method of data collection, study setting, and intervention description, surgical access type, provider roles (eg, surgeon, anesthesiologist, and nurse), surgical risk, expertise level, intervention design (ie, bundle vs. single intervention), reports of institutional support, and location income level. All outcomes reported in studies were extracted (ie, multiple outcomes could be extracted from 1 study, and therefore, the number of outcomes surpasses the number of studies included in this review). Outcome data were inductively categorized into technical skill, NTS, patient outcome, culture, compliance and protocol quality, ergonomics, professional development, and resilience. Outcomes were given 1 of 3 ratings: improved, did not improve, and mixed. Studies that examined multiple outcomes of the same outcome category may report "mixed" results; whereby 2 or more outcomes within the same outcome type reported differing ratings (eg, one improved and one did not improve). These outcomes were analyzed and reported separately from those with nonmixed results because they do not fit neatly into either "improved" or "did not improve" outcome ratings. For example, consider a study that reports on 2 outcomes: average blood loss and average postoperative complications. The study finds that average blood loss improved postintervention, whereas average postoperative complications did not. While these outcomes fall under the "patient outcomes" category ("complications" subcategory), the study would receive a "mixed" rating for this category due to the divergent results. Separating each outcome individually, rather than grouping them by type, would introduce an unnecessary level of detail that detracts from our broader, high-level analysis. Maintaining the "mixed" rating allows us to present a more balanced and comprehensive overview of the intervention's impact without oversimplifying or overcomplicating the findings.

Qualitative Content Analysis

Qualitative content analysis encompasses a suite of established analytical techniques for interpreting textual data.[24] This review uses both deductive (using categories from previously published frameworks) and inductive (newly created categories and subcategories) approaches to identify categories pertaining to intervention and outcome types.

A deductive approach was used for NTS outcomes. These outcome categories were derived from the NTSs for surgeons validated measures that describe the 5 NTSs: situation awareness, decision-making, leadership, communication, and teamwork.[25]

An inductive approach was primarily chosen to determine intervention and outcome types, with the exception of NTS outcomes (eg, interventions seeking to implement new policies on clinical practice and processes of surgical site infection prevention were categorized as "policy" with the subcategory of "process"; outcomes that capture adherence to policy and practice were categorized as "compliance and protocol quality" outcomes). Four investigators (B.A.A., A.T., L.L., and P.T.) inductively generated categories from the data through discussion.

Quality Assessment

Four quality assessment tools were used to assess 6 study types. Case–control and cohort studies were assessed using the case–control and cohort versions of the "Newcastle-Ottawa Scale".[26] Pre–post and controlled intervention studies were assessed using the "National Institute of Health Study Quality Assessment Tools".[27] Qualitative studies were assessed with the "Nagpal Quality Assessment Tool".[28] Cross-over studies were assessed with the "Cross-Over Quality Assessment Standard".[29]

Each study was rated independently by 2 reviewers using the appropriate quality assessment tool. All conflicts were resolved through discussion. In line with previous reviews, due to the heterogeneity of quality assessment tools used, categories describing low quality (high risk of bias), medium quality (medium risk of bias), and high quality (low risk of bias) were developed to compare research quality across all included studies.[30]

See Supplemental Digital Contents 1 to 8, http://links.lww.com/AOSO/A480, for more details on the protocol, search strategy, selection process, data extraction, and quality assessment.

RESULTS

Study Selection

A total of 35,801 reviews were identified. After duplicates were removed and reviews were screened, 65 reviews were retained (Fig. 1A). The 65 reviews contained a total of 1355 studies. After duplicates were removed across reviews and studies were screened, 575 studies were included in the final analysis (Fig. 1B). The interrater reliability (IRR; average agreement) for review-level screening was 99% for titles and abstracts and 86% for full text. IRR for study-level screening was 87% for titles and abstracts, 96% for full text, and 90% for data extraction.

Study Characteristics

A total of 5,288,513 cases, 1102 simulations, 25,435 participants (ie, providers and patients), and 1189 teams, from more than 2608 hospitals, across 50 countries were included in this analysis. The following surgical specialties collectively accounted for the majority (50%) of studies: general (35%), orthopedic (9%), and urology (6%). The data primarily focused on the outcomes of interventions that involve surgeons (49%) and surgical teams (44%), with less emphasis on anesthesiologists (6%) and nurses (1%). Access types varied among the studies, with scopes being the most common (32%), followed

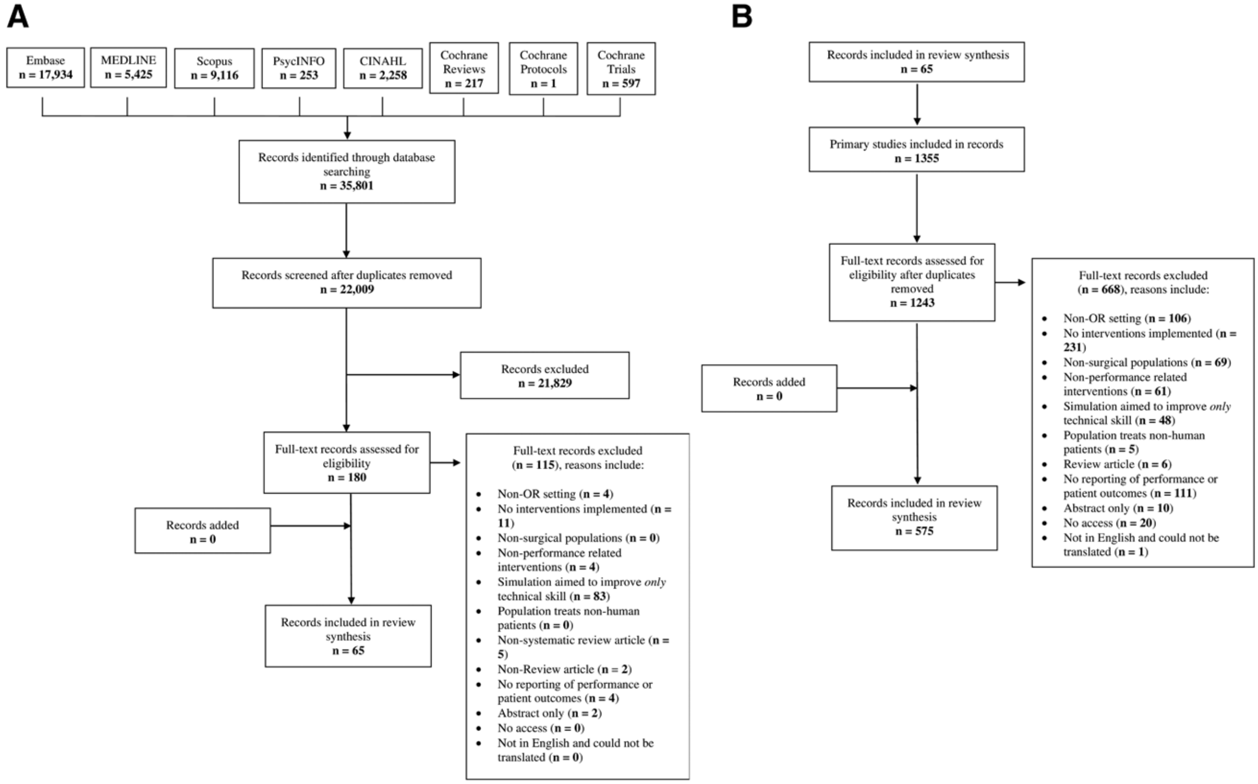

FIGURE 1. PRISMA flow diagram. (A) Review-level selection. (B) Study-level selection. PRISMA indicates Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

by open procedures (26%). Some studies reported mixed access types (10%), whereas others used robotics (4%). A significant portion of studies (28%) did not specify the access type. Most studies focused on surgeries with an intermediate (34%) or low (27%) surgical risk and were conducted in real OR settings (57%). The majority of studies were conducted in high-income locations (93%). Most studies did not report participant gender (75%), whether they received institutional support for their intervention (85%), or the duration of intervention effects postimplementation (91%). See Supplemental Table 1, http://links.lww.com/AOSO/A480, for more details about study characteristics, Supplemental Table 2, http://links.lww.com/AOSO/A480, for the data extracted from each study, and Supplemental Digital Contents 9 and 10, http://links.lww.com/AOSO/A480, for the complete list of reviews (n = 65) and studies (n = 575) that were included in this analysis.

Study Quality

The IRR (ie, average agreement) across all quality assessment tools by investigators was 76% (438/575). Of the 575 studies included in this review, 129 (22%) were rated high in quality (ie, low risk of bias), 259 (45%) were rated medium in quality (ie, medium risk of bias), and 187 (33%) were rated low in quality (ie, high risk of bias). High-quality studies were characterized by good alignment of research questions and results reported/interpreted, comprehensive information on the study design and methodology, including reliable measurements, large representative sample sizes, and appropriate randomization techniques to minimize allocation, measurement bias, and confounding factors. Studies with medium-quality scores reflected good alignment of research questions and results reported/interpreted, reliable measurements, medium-to-large sample sizes, and use of appropriate randomization methods. Low-quality studies were characterized by inadequate information regarding study design and methodology, use of unvalidated tools or measurements, poor alignment between research question and reported/interpreted results, smaller sample sizes, and critical flaws such as high dropout rates in randomized controlled trials or inappropriate statistical analyses. The improvement of surgical outcomes (indicating a positive intervention effect) did not depend on the quality of the study; there was only an 8% difference in outcome improvement between low-quality and high-quality studies. Specifically, 80% of low-quality studies reported improvement, compared with 79% of medium-quality studies, and 72% of high-quality studies. See Supplemental Tables 3 to 7, http://links.lww.com/AOSO/A480, for quality assessment details for each study.

In support of the next 2 sections below, see Supplemental Table 8, http://links.lww.com/AOSO/A480, which provides definitions, examples, and more details concerning the frequency of each intervention and outcome type.

Frequency of Intraoperative Intervention Types

Fifty-eight % (331/575) of studies in this review focused on person-based interventions (ie, education, policy), and 42% (244/575) of studies focused on system-based interventions (ie, technology, cognitive aids, equipment, standardization, and environment redesign). Specifically, the most common person-based intervention type was education (n = 220), most often in the form of learning through mentorship, followed by policy (n = 111) interventions that typically involved process changes to a workflow procedure or system. The most common system-based intervention was the use of technology (n = 86), with the majority aimed at improving the surgeon's visualization of the surgical site. The second most common system-based interventions were cognitive aids (n = 61) and equipment (n = 61). Cognitive aids refer to tools designed to

TABLE 1.

Effectiveness of Intervention Types on Outcomes

	Outcome Types																															
	Technical Skills				Nontechnical Skills				Patient Outcomes				Culture				Professional Development				Resilience				Compliance and Protocol Quality				Ergonomics			
Intervention Types	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%
Person-based interventions																																
Education																																
Mentorship	36	16	1	69	14	3	0	82	7	19	0	27	21	0	0	100	23	1	0	95	3	0	0	100	2	0	0	100	0	0	0	-
Psychological well-being training	0	0	0	-	9	0	0	100	0	0	0	-	2	0	0	100	5	0	0	100	2	0	0	100	0	0	0	-	0	0	0	-
Simulation	30	2	2	94	49	4	1	92	6	2	0	75	25	1	1	96	42	1	3	98	3	1	0	75	3	1	1	75	1	0	0	100
Didactics	11	2	0	85	21	0	1	100	6	12	1	33	11	0	0	100	9	0	1	100	1	0	0	100	6	0	0	100	4	0	0	100
Warm-ups	19	7	1	73	3	3	0	50	2	4	0	33	2	0	0	100	5	0	0	100	3	1	1	75	0	0	0	-	2	1	0	67
Weighted average	-	-	-	71	-	-	-	91	-	-	-	30	-	-	-	95	-	-	-	97	-	-	-	73	-	-	-	82	-	-	-	75
Policy																																
Process	5	4	1	56	2	4	0	33	84	45	10	65	7	3	2	70	2	1	0	67	2	2	0	50	22	6	2	79	3	0	0	100
Resourcing	4	2	0	67	0	0	0	-	12	1	4	92	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Weighted average	-	-	-	58	-	-	-	33	-	-	-	63	-	-	-	70	-	-	-	67	-	-	-	50	-	-	-	79	-	-	-	100
System-based interventions																																
Technology																																
Apps	2	1	0	67	2	0	0	100	0	0	0	-	3	0	0	100	4	3	0	57	1	0	0	100	0	0	0	-	0	0	0	-
Digital checklist	0	0	0	-	1	0	0	100	0	0	0	-	3	0	0	100	0	0	0	-	0	0	0	-	1	0	1	100	0	0	0	-
Vision	59	15	5	80	0	0	0	-	5	15	1	25	12	2	2	86	6	2	0	75	5	0	1	100	1	0	0	100	14	1	6	93
Other (eg, VR)	5	1	0	83	1	1	0	50	0	2	0	0	5	1	0	83	2	0	0	100	0	0	0	-	0	0	0	-	1	0	0	100
Weighted average	-	-	-	80	-	-	-	60	-	-	-	27	-	-	-	83	-	-	-	64	-	-	-	100	-	-	-	100	-	-	-	93
Cognitive aids																																
Average	5	3	0	63	31	0	0	100	9	4	0	69	19	2	1	90	5	1	0	83	0	0	0	-	36	1	2	97	0	0	0	-
Equipment																																
Assist with practice	0	1	0	0	0	0	0	-	8	2	0	80	4	0	0	100	2	1	0	67	0	0	0	-	13	0	0	100	0	0	0	-
Assist with surgery	17	9	0	65	0	0	0	-	31	30	5	51	0	2	0	0	0	0	0	-	0	0	0	-	0	0	0	-	1	0	0	100
Weighted average	-	-	-	67	-	-	-	-	-	-	-	48	-	-	-	67	-	-	-	67	-	-	-	-	-	-	-	100	-	-	-	100
Standardization																																
Average	5	0	0	100	2	0	0	100	22	4	2	85	0	0	0	-	1	0	0	100	0	0	0	-	12	2	0	86	0	0	0	-
Environment redesign																																
Average	6	3	4	67	1	2	0	33	0	1	0	0	1	0	0	100	0	0	0	-	1	1	0	50	0	0	0	-	0	0	0	-
Total	204	66	14	76	136	17	2	89	192	141	23	57	115	11	6	91	106	10	4	91	21	5	2	81	96	10	6	91	26	2	6	93

The numbers reflect the frequency each outcome was reported across studies for each outcome type.

% indicates intervention effectiveness (improved/improved + did not improve ×100); D, did not improve; I, improved; M, mixed; VR, virtual reality; weighted average, accounts for the varying outcome frequencies across intervention subthemes to prevent the overall effectiveness from being biased by high-frequency outcomes.

support decision-making such as new surgical safety checklists. Equipment interventions involve physical tools used by the surgeon (eg, stapler) or by other OR staff (eg, nurses using radio frequency detection to prevent retained surgical sponges during counts).

Effects of Intervention Types on Outcomes in Surgery

From the 575 studies included in the review, 1221 outcomes were extracted. These outcomes were categorized into 8 types, with 3 of these types (ie, technical skills, NTS, and patient outcomes) further divided into subcategories. Below, we present the top 2 intervention types with the highest effectiveness scores for each outcome type and, where applicable, each outcome subcategory. Effectiveness scores were compared within each outcome type, not across different outcome types, to ensure a more valid and contextually meaningful comparison.

Technical Skills

Across all studies, 23% (284/1221) of outcomes reported were technical skills.

Standardization interventions showed the largest overall improvement on technical skills (100%; n = 5; Table 1). This intervention type improved both technical skill subcategories: proficiency and performance time (Table 2). For example, the formation of a standardized operating team and the development of standardized protocols for anesthetic management and patient preparation for posterior spinal fusion for scoliosis led to faster performance times and a reduction in OR time while maintaining quality of care.[31]

Technology interventions also improved technical skills (80%; n = 66; Table 1), with specific improvements observed in the proficiency subcategory of technical skills (Table 2). For instance, a 3-D visualization system in robotic surgery led to more accurate and faster completion of anastomosis compared with a 2-D visualization system.[32]

Nontechnical Skills

Of the 1221 outcomes extracted, 13% (155/1221) of these outcomes were NTS.

Cognitive aid interventions showed a high frequency of improvements in NTS (100%; n = 31; Table 1). Specific improvements

TABLE 2.

Effectiveness of Intervention Types on Technical Skill Subcategories

	Technical Skill Subcategory Outcomes							
	Proficiency				Performance Time			
Intervention Types	I	D	M	%	I	D	M	%
Person-based interventions								
Education								
Mentorship	26	6	1	81	7	10	0	41
Psychological well-being training	0	0	0	-	0	0	0	-
Simulation	22	1	2	96	8	1	0	88
Didactics	9	1	0	90	2	1	0	67
Warm-ups	15	5	1	75	4	2	0	67
Weighted average	-	-	-	81	-	-	-	47
Policy								
Process	3	1	0	75	2	3	1	40
Resourcing	0	0	0	-	4	2	0	67
Weighted average	-	-	-	75	-	-	-	56
System-based interventions								
Technology								
Apps	2	1	0	67	0	0	0	-
Digital checklist	0	0	0	-	0	0	0	-
Vision	31	5	2	86	28	10	3	74
Other (eg, VR)	3	0	0	100	2	1	0	67
Weighted average	-	-	-	86		-	-	74
Cognitive aids								
Average	2	0	0	100	3	3	0	50
Equipment								
Assist with practice	0	1	0	0	0	0	0	-
Assist with surgery	0	0	0	-	17	9	0	65
Weighted average	-	-	-	-	-	-	-	65
Standardization								
Average	1	0	0	100	4	0	0	100
Environment redesign								
Average	4	1	2	80	2	2	2	50
Total	120	22	8	85	84	44	6	66

The numbers reflect the frequency each outcome was reported across studies for each outcome type.

% indicates intervention effectiveness (improved/improved + did not improve ×100); D, did not improve; I, improved; M, mixed; VR, virtual reality; weighted average, accounts for the varying outcome frequencies across intervention subthemes to prevent the overall effectiveness from being biased by high-frequency outcomes.

among the subcategories of NTS included teamwork, communication, leadership, and the aggregate NTS score (Table 3). For instance, a modified case-specific surgical safety checklist cognitive aid improved teamwork by allowing staff members to better distribute tasks, share the workload, and keep track of progress.[33]

Standardization interventions also improved NTS (100%; n = 2; Table 1). In particular, the NTS subcategories communication and the aggregate NTS score showed improvements, although sample sizes were low (Table 3). For example, the development of a standardized protocol for notifying staff about the presence of vaginal swabs in situ led to an improvement in communication during patient handover and patient transfer outside the OR and reduced the incidence of retained vaginal swabs and near misses.[34]

Patient Outcomes

Twenty-nine % (356/1221) of the outcomes were patient outcomes.

Standardization interventions showed the largest improvement on patient outcomes (85%; n = 22; Table 1). Specifically, the patient outcome subcategories including errors, situation severity, and recovery showed the largest improvement, followed by patient complications which also showed improvement (Table 4). For example, one study aiming to reduce infection rates in total joint surgeries showed that the use of standardized procedures such as using pulsed xenon ultraviolet light, 2 preoperative showers with chlorhexidine gluconate, skin cleansing with chlorhexidine gluconate immediately before surgery, and perioperative order sets for 12 months resulted in a statistical reduction in surgical site infections, compared with not using these standardized procedures.[35]

Cognitive aid interventions were also an intervention type that improved patient outcomes (69%; n = 9; Table 1). The patient outcome subcategories including situation severity and mortality showed improvements in particular (Table 4). For example, the use of a surgical debriefing checklist was associated with a decrease in surgical cases with reported adverse events and a significant reduction in the 30-day unadjusted surgical mortality rate.[36]

Culture

Of the 1221 outcomes extracted, 11% (132/1221) represented outcomes related to culture.

An Environment Redesign intervention showed an improvement in culture (100%; n = 1; Table 1), although this stemmed from only 1 study. Specifically, this study describes how the World Health Organization recommends OR noise levels to be maintained below 35 dB, yet average OR noise levels range from 61.6 to 69.9 dB, increasing the likelihood of adverse events in surgery. The aim of this study was to improve the safety culture of the OR by reducing sound levels during critical phases of surgery. The authors refer to the well-established connection between unnecessary noise, distraction, and errors in the OR, which are indicative of poorer safety culture. To address this, "No Interruption Zones" (ie, prohibiting nonessential conversations and activities) were implemented during these critical surgical phases. Significant reductions in sound levels (ie, a quiet OR environment) were observed within the No Interruption Zones, serving as an indicator of improved safety culture in the OR.[37]

Education interventions also showed evidence of improving outcomes related to culture (98%; n = 61; Table 1). For instance, an emotional intelligence training program not only improved the Emotional-Quotient Inventory of otolaryngology residents and faculty but also led to increased patient satisfaction, reflecting a positive shift in the safety culture.[38]

Professional Development

Nine % (106/1221) of the outcomes extracted were related to professional development.

Standardization interventions showed the largest improvement on professional development (100%; n = 1; Table 1), but this was represented by only 1 study. Specifically, in this study, a committee of total joint arthroplasty OR staff developed process changes and standardized protocols (eg, reducing and standardizing trays of instruments) to improve the efficiency of the OR. These changes led to improved OR efficiencies (eg, more on-time starts), OR staff productivity (eg, shorter average turnover times), and professional growth (eg, acquiring the skills to identify OR inefficiencies and proposing constructive solutions to identified problems).[39]

Education interventions also showed improvements in professional development (98%; n = 84; Table 1). For example, the development of a mentorship program, whereby newly appointed congenital heart surgeons were mentored by more experienced surgeons, led to independent practice of congenital heart surgery on pediatric patients by the end of the program.[40]

Resilience

Two % (28/1221) of the outcomes extracted were related to resilience.

TABLE 3.

Effectiveness of Intervention Types on Nontechnical Skill Subcategories

	Nontechnical Skill Subcategory Outcomes																							
	Teamwork				Decision-Making				Situation Awareness				Communication				Leadership				Aggregate NTS Score			
Intervention Types	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%
Person-based interventions																								
Education																								
Mentorship	3	0	0	100	0	0	0	-	0	0	0	-	5	1	0	83	4	1	0	80	2	1	0	67
Psychological well-being training	1	0	0	100	1	0	0	100	0	0	0	-	3	0	0	100	4	0	0	100	0	0	0	-
Simulation	11	0	0	100	3	2	0	60	5	1	0	83	9	0	0	100	6	0	1	100	15	1	0	94
Didactics	6	0	0	100	4	0	0	100	2	0	0	100	5	0	0	100	1	0	0	100	3	0	1	100
Warm-up	1	1	0	50	1	0	0	100	0	0	0	-	0	0	0	-	0	0	0	-	1	2	0	33
Weighted average	-	-	-	88	-	-	-	60	-	-	-	81	-	-	-	86	-	-	-	82	-	-	-	88
Policy																								
Process	0	1	0	0	0	1	0	0	0	1	0	0	2	0	0	100	0	1	0	0	0	0	0	-
Resourcing	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Weighted average	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	-	-	-	-	-	-	-	-
System-based interventions																								
Technology																								
Apps	0	0	0	-	0	0	0	-	0	0	0	-	1	0	0	100	1	0	0	100	0	0	0	-
Digital checklist	0	0	0	-	0	0	0	-	0	0	0	-	1	0	0	100	0	0	0	-	0	0	0	-
Vision	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Other (eg, VR)	0	0	0	-	0	0	0	-	0	0	0	-	0	1	0	0	1	0	0	100	0	0	0	-
Weighted average	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	50	-	-	-	100	-	-	-	-
Cognitive aids																								
Average	15	0	0	100	0	0	0	-	0	0	0	-	11	0	0	100	1	0	0	100	4	0	0	100
Equipment																								
Assist with practice	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Assist with surgery	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Weighted average	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Standardization																								
Average	0	0	0	-	0	0	0	-	0	0	0	-	1	0	0	100	0	0	0	-	1	0	0	100
Environment redesign																								
Average	0	0	0	-	0	0	0	-	0	1	0	0	1	1	0	50	0	0	0	-	0	0	0	-
Total	37	2	0	95	9	3	0	75	7	3	0	70	39	3	0	93	18	2	1	90	26	4	1	87

The numbers reflect the frequency each outcome was reported across studies for each outcome type.
% indicates intervention effectiveness (improved/improved + did not improve ×100); D, did not improve; I, improved; M, mixed; NTS, nontechnical skills; VR, virtual reality; weighted average, accounts for the varying outcome frequencies across intervention subthemes to prevent the overall effectiveness from being biased by high-frequency outcomes.

Technology interventions showed improvements in resilience (100%; n = 6; Table 1). For example, residents in the departments of general surgery, anesthesia and obstetrics and gynecology used a mindfulness and meditation smartphone app (ie, Headspace) on a self-guided basis and showed improvements in positive affect and mindfulness after 4 weeks.[41]

Education interventions also showed evidence of improving resilience (86%; n = 12; Table 1). For example, general surgery residents who participated in an Energy Leadership Well-Being and Resiliency Program for 1 year showed positive improvements in their perceived stress, emotional exhaustion, emotional intelligence, life satisfaction, and their perception of the residency program.[42]

Compliance and Protocol Quality

Of the 1221 outcomes extracted, 9% (112/1221) represented outcomes related to compliance and protocol quality.

Equipment interventions showed improvements in compliance and protocol quality (100%; n = 13; Table 1). For example, covering spinal implants with sterile surgical towels until the implants were required for the case compared with leaving the implants open and uncovered resulted in improvements in the quality of preparing implants for surgery and significantly lowered rates of implant contamination.[43]

Technology interventions also improved compliance and protocol quality-related outcomes (100%; n = 2; Table 1), although the sample size was low. For example, an aviation-style digital surgical safety checklist displayed on a large, centrally located screen in the OR significantly increased the proportion of cases in which all of the items on the checklist were completed and improved surgical team participation, compared with a wall poster surgical safety checklist.[44]

Ergonomics

Three % (34/1221) of the outcomes extracted were related to ergonomics.

Policy interventions demonstrated 100% improvement in ergonomics (n=3; Table 1). For example, one study showed that performing microbreak exercises during surgery led to reduced muscle strain and pain experienced by surgeons.[45]

Equipment interventions also improved ergonomics (100%; n=1; Table 1); however, this only stemmed from 1 study. This study showed that using an intra-abdominal laparoscopic cleaning device, which eliminates the multistep external cleaning process, improved surgical site visualization and reduced workflow interruptions, compared with the standard scope cleaning protocol.[46]

Sustained Intervention Effects

Among the 575 studies reviewed, 10% (56/575) reported whether the intervention effects were sustained postimplementation. Of these, 80% (45/56) reported effects that were sustained when tested postimplementation and 20% (11/56)

TABLE 4.

Effectiveness of Intervention Types on Patient Outcome Subcategories

	Patient Subcategory Outcomes																			
	Complications				Error				Mortality				Situation Severity				Recovery			
Intervention Types	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%	I	D	M	%
Person-based interventions																				
Education																				
Mentorship	1	12	0	8	2	1	0	67	0	4	0	0	1	0	0	100	3	2	0	60
Psychological well-being training	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Simulation	2	1	0	67	3	1	0	75	0	0	0	-	0	0	0	-	1	0	0	100
Didactics	2	5	1	29	2	2	0	50	2	1	0	67	0	0	0	-	0	4	0	0
Warm-up	0	2	0	0	1	0	0	100	0	1	0	0	1	0	0	100	0	1	0	0
Weighted average	-	-	-	17	-	-	-	60	-	-	-	40	-	-	-	100	-	-	-	45
Policy																				
Process	59	24	7	71	5	1	0	83	3	6	0	33	6	1	1	86	11	13	2	46
Resourcing	3	0	3	100	0	0	0	-	0	0	0	-	4	0	1	100	5	1	0	83
Weighted average	-	-	-	70	-	-	-	83	-	-	-	33	-	-	-	84	-	-	-	43
System-based interventions																				
Technology																				
Apps	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Digital checklist	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-	0	0	0	-
Vision	1	7	1	13	2	1	0	67	0	2	0	-	0	0	0	-	2	5	0	29
Other (eg, VR)	0	1	0	0	0	0	0	-	0	0	0	-	0	0	0	-	0	1	0	0
Weighted average	-	-	-	14	-	-	-	67	-	-	-	-	-	-	-	-	-	-	-	32
Cognitive aids																				
Average	3	2	0	60	2	1	0	67	3	1	0	75	1	0	0	100	0	0	0	-
Equipment																				
Assist with practice	8	1	0	89	0	0	0	-	0	1	0	0	0	0	0	-	0	0	0	-
Assist with surgery	21	9	5	70	1	0	0	100	0	7	0	0	0	0	0	-	9	14	0	39
Weighted average	-	-	-	68	-	-	-	100	-	-	-	-	-	-	-	-	-	-	-	39
Standardization																				
Average	12	3	2	80	3	0	0	100	2	1	0	67	1	0	0	100	4	0	0	100
Environment redesign																				
Average	0	0	0	0	0	0	0	-	0	0	0	-	0	0	0	-	0	1	0	0
Total	112	68	19	62	21	7	0	75	10	24	0	29	14	1	2	93	35	42	2	45

The numbers reflect the frequency each outcome was reported across studies for each outcome type.

% indicates intervention effectiveness (improved/improved + did not improve ×100); D, did not improve; I, improved; M, mixed; VR, virtual reality; weighted average, accounts for the varying outcome frequencies across intervention subthemes to prevent the overall effectiveness from being biased by high-frequency outcomes.

reported effects that were not sustained when tested postimplementation. Seventy-one percent (32/45) of the studies with sustained effects postimplementation were person-based interventions: 25 education interventions had sustained effects when followed 8.2 months on average with a range of 1 to 48 months postimplementation (6 education interventions did not have sustained effects when followed 8.6 months on average postimplementation); 7 policy interventions had sustained effects when followed 28.3 months on average with a range of 12 to 36 months postimplementation (3 policy interventions did not have sustained effects when followed 9.7 months on average postimplementation). Twenty-nine percent (13/45) of the studies with sustained effects postimplementation were system-based interventions: 6 standardization interventions had sustained effects when followed 17.8 months on average with a range of 8 to 36 months postimplementation; 6 cognitive aid interventions had sustained effects when followed 7.3 months on average with a range of 3 to 15 months postimplementation (6 cognitive aid interventions did not have sustained effects when followed 7 months on average postimplementation), and 1 technology intervention showed sustained effects when tested 6 months postimplementation. See Table S2 for which studies reported on whether intervention effects were sustained postimplementation and for how long.

DISCUSSION

The objectives of this systematic review were to (1) identify the most common types of intraoperative interventions and determine their prevalence and (2) examine how effective these intervention types were at improving surgical outcomes.

Frequency of Intraoperative Intervention Types

Among the 575 studies reviewed, 58% implemented person-based interventions (ie, education and policy), making them more common than the broader range of system-based interventions (ie, technology, equipment, cognitive aids, standardization, and environment redesign), which were implemented in 42% of the studies. These findings align with prior research showing that person- as opposed to system-based interventions are the most common interventions in surgery,[3,10] likely due to their relative ease of implementation. Yet, as we will discuss below, system-based interventions are far more likely to improve outcomes in surgery.

Frequency of Outcome Types

Our systematic review found that the majority of studies assessing the effectiveness of intraoperative interventions have predominantly focused on improving patient outcomes (40%), specifically postoperative complications, and technical skills (40%), specifically proficiency and performance time. However, there was a noticeable gap in the literature concerning the impact of interventions on culture, professional development, compliance and protocol quality, resilience, and ergonomics. The paucity of research in these areas highlights the need to broaden the scope of research to include these critical yet underrepresented areas, so that we can gain a more comprehensive understanding of how to optimize intraoperative performance.

Intervention Effectiveness on Surgical Outcomes

The effectiveness of person- and system-based interventions varies depending on the outcomes being measured. Our analysis compares intervention effectiveness within each outcome type, ensuring that each intervention is evaluated against similar metrics. System-based interventions consistently ranked among the top 2 highest effectiveness scores across all outcome types, indicating their broad and significant impact on improving outcomes compared to other types of interventions. The most impactful system-based interventions included standardization, which improved technical and NTS, patient outcomes, and professional development; environment redesign interventions that improved culture; cognitive aid interventions that improved NTS, and compliance and protocol quality; equipment and technology interventions that boosted compliance and protocol quality, and ergonomics, with technology also enhancing technical skills and resilience.

In contrast, person-based interventions, specifically education, ranked among the top 2 interventions for improving culture, professional development, and resilience outcomes. However, it is important to note that interventions not ranking in the top 2 may still be valuable. For example, although education did not make the top 2 for improving NTS (despite showing 91% effectiveness across 96 studies), it remains a strong option for this outcome and was critical for improving safety culture and interpersonal outcomes. Our systematic review specifically excluded studies focusing solely on technical skill improvements through simulations, as the link between simulations and technical skill enhancements is well-established. Nevertheless, person-based interventions involving simulations often lead to improvements in technical skills, and our findings also indicate that such interventions can enhance NTS, resilience, professional development, and safety culture. This highlights the important, complementary role that person-based interventions, particularly education, can play.

While system-based interventions were ranked among the top 2 most effective for all outcome types, it is important to recognize that person-based interventions, especially educational strategies, ranked highly for improving safety culture and interpersonal dynamics within surgical teams. We also found that policy interventions, while less frequently cited, showed effectiveness in improving ergonomics—though this finding was based on only 3 studies. The results from this review challenge the notion that education and policy are inherently weak interventions. Instead, our findings emphasize that these interventions play a vital role in improving safety culture, resilience, and professional development across healthcare systems.

The human factors engineering literature indicates that healthcare problems are often deeply rooted in systemic issues rather than individual shortcomings.[3,10] System-based changes have shown to improve various healthcare outcomes.[47] Our findings support this perspective, by demonstrating that system-based interventions consistently improved a wide-range of surgical outcomes. However, our review also underscores that person-based interventions—particularly education—remain a powerful tool for addressing interpersonal and cultural outcomes in surgery, highlighting the importance of a dual approach that combines both system-level changes and individual skill development.

Thus, while system-based interventions tend to show broad, consistent improvements across diverse surgical outcomes, person-based interventions, especially those focused on education and policy, continue to play a crucial role in shaping safety culture, resilience, and professional development. These interventions are not inherently weak, but rather serve specific, critical functions in improving healthcare delivery.

There is a plethora of heterogenous problems to be worked on in surgery—reducing provider burn-out, patient complications and mortality, improving surgical skills, task efficiency, ergonomics, and safety culture—and there is no "one size fits all" intervention type to tackle these. Healthcare challenges often stem from complex interactions among multiple variables that are difficult to identify and may require a combination of intervention types to yield improvement. For example, McCulloch et al[4] investigated the effectiveness of person-based interventions (ie, team training) and system-based interventions (ie, standard operating procedures to analyze, redesign and standardize work systems, processes, and lean thinking techniques to eliminate waste, make error visible and engage staff) alone and in combination.[48] They found that combining person- and system-based interventions enhanced a greater variety of surgical outcomes, specifically technical and NTS, and surgical safety checklist compliance, highlighting the synergistic benefits of combining interventions. By distinguishing the areas of impact for each type of intervention, this review highlights the importance of thoughtfully selecting interventions that align with specific outcomes, helping to clarify how targeted interventions can be strategically deployed to achieve desired improvements in individual or combined surgical outcomes.

Future Research

We call for more attention to the work system factors that may impact the surgical team, their tasks, and downstream clinical outcomes. The Systems Engineering Initiative for Patient Safety model,[49] which provides a framework for understanding the structures, processes and outcomes and their relationships in healthcare settings, may be particularly helpful in determining which factors are impacting certain surgical outcomes and which interventions may yield the greatest impact. In particular, there is little research on the interventions that optimize the physical layout and equipment set-up of the OR. Further, researchers may want to consider a wider lens of intervention options when determining which interventions to implement. For instance, we observed an unexpected association between an environment redesign intervention and improvements in safety culture, in which designated "No Interruption Zones" were implemented during critical surgical phases to enforce a quiet OR and ensure that each surgical team member could concentrate fully on their respective tasks, improving the safety culture of the OR. When system-based interventions are implemented, providers may experience increased support and investment from their organization, as it reflects a shared commitment to improvement and surgical safety. This alignment may, in turn, enhance the safety culture as a secondary benefit.

In addition, we argue that a standardized approach of research impact evaluation may address methodological discrepancies and better inform researchers on which interventions are most impactful.[50] For example, there was limited research on whether the effects of interventions were sustained over time, whether they received institutional support, the detailed implementation processes, or the facilitators and barriers to their implementation. Other gaps in the literature that warrant investigation include determining whether provider gender, expertise level, country income level,[51] or surgical risk level impact intervention effectiveness. More efforts should be directed toward developing surgical interventions that support anesthesia and nursing teams (eg, improving task efficiency for their roles), as the majority of interventions have focused primarily on surgeons or the surgical team as a whole.

LIMITATIONS

The limitations of the evidence in this review include (1) a lack of standardization concerning how intervention effectiveness is reported, making it difficult to synthesize and (2) a paucity of research investigating the impact of interventions on critical surgical outcomes such as culture, professional development, compliance and protocol quality, resilience, and ergonomics.

The limitations of our review processes include the following. (1) Our search strategy aimed to identify the most common types of intraoperative interventions implemented in the surgical literature by searching through published reviews. However, it is likely that our approach did not capture all common intervention types in surgery, particularly interventions that have not been investigated by way of review-level analysis. (2) A meta-analysis to statistically test intervention effectiveness was not conducted because there was substantial heterogeneity across studies in terms of research aims and many important surgical outcome variables were qualitative rather than quantitative. (3) Our data aggregation approach may have inadvertently eliminated important between- and within-study variation. A high-level, coarse overview of intervention and outcome types was developed to manage the large and highly variable dataset and answer our *a priori* research questions. The dichotomy of person-based and system-based interventions may oversimplify their complexity, as many interventions include both components. However, most had a dominant component, allowing for meaningful categorization and insights into effectiveness. Mixed outcomes (both positive and negative results within a study) were excluded from effectiveness calculations to reduce data variability. While this may have underestimated effectiveness when minor negatives offset positives, it was justified given that only 5% of studies showed mixed results (63/1221; Table 1). (4) A limitation of focusing on the top 2 interventions is that effective options, such as the third-ranked intervention, may be overlooked. Rather than applying an arbitrary threshold (eg, effectiveness >75%), we prioritized the top 2 interventions within each outcome type for more context-specific comparisons. We also calculated effectiveness without overemphasizing sample size, ensuring that novel, less-implemented interventions were highlighted. All effectiveness scores are available in Table 1 for reader review. (5) Our original protocol was amended in PROSPERO for clarity. Initially labeled as an umbrella review, it was recategorized as a systematic review based on the Cochrane Handbook for Systematic Reviews of Intervention's[17] definition, as we focused on analyzing original studies within systematic reviews and meta-analyses. The methodology remains unchanged from our original PROSPERO protocol. We also expanded our criteria from only extracting statistically tested technical and NTS outcomes to include all surgical outcomes reported in the intraoperative intervention literature, regardless of statistical testing, for a more comprehensive review. Our PROSPERO registration has been updated accordingly.

CONCLUSION

Effective interventions in healthcare depend not only on the intervention type but also on its alignment with the specific outcome being targeted. This review highlighted that person-based interventions, particularly educational strategies, ranked among the top 2 most effective interventions for improving safety culture, interpersonal outcomes, and resilience. These interventions were also the most commonly implemented across healthcare settings. However, person-based interventions were often less effective in improving patient outcomes and technical skills, highlighting the need for careful alignment between the intervention type and desired outcome. In contrast, system-based interventions consistently ranked among the top 2 most effective interventions across all outcome types, but were less frequently implemented, possibly due to factors such as higher implementation costs or complexity. In light of this, healthcare providers and administrators should carefully align interventions to the outcomes they aim to improve by considering a broader range of options—including system-based interventions—rather than defaulting to person-based approaches only. Thoughtfully aligning interventions to their respective outcomes can help healthcare systems achieve meaningful improvements, optimize resource allocation, and avoid ineffective or mismatched interventions.

ACKNOWLEDGMENTS

The authors thank Eleanor Cheatham, Anbreen Zulfiqar, Taylor Wu, and Jeneen Massad for helping to screen studies, extract data, rate the quality of the included studies, and review table data.

All authors listed below meet the following criteria: (1) participated sufficiently in the intellectual content, (2) the analysis of data, (3) the writing of the manuscript, and (4) have reviewed the manuscript, believe it represents valid work, and approve it for submission. B.A.A.: idea generation, study screening, data extraction, data cleaning, analysis, risk of bias assessments, manuscript writing and revisions, Supplemental Digital Content (SDC) http://links.lww.com/AOSO/A480 writing and revisions, and final approval was given. A.T.: idea generation, study screening, data extraction, data cleaning, analysis, risk of bias assessments, manuscript writing and revisions, SDC http://links.lww.com/AOSO/A480 revisions, and final approval was given. L.L., J.O., J.W., and J.J.: study screening, data extraction, data cleaning, analysis, risk of bias assessments, manuscript revisions, and final approval was given. S.S.A. and M.Z.: study screening, data extraction, data cleaning, analysis, risk of bias assessments, manuscript revisions, SDC http://links.lww.com/AOSO/A480 revisions, and final approval was given. J.M.: development of search strategy, ran searches and obtained full list of studies for screening, and final approval was given. P.T.: idea generation, data cleaning, analysis, manuscript writing and revisions, and final approval was given.

REFERENCES

1. Wacker J, Kolbe M. The challenge of learning from perioperative patient harm. *Trends Anaesth Crit Care*. 2016;7-8:5–10.
2. Dell-Kuster S, Gomes NV, Gawria L, et al. Prospective validation of classification of intraoperative adverse events (ClassIntra): international, multicentre cohort study. *BMJ*. 2020;2020:m2917.
3. Trbovich P, Shojania KG. Root-cause analysis: swatting at mosquitoes versus draining the swamp. *BMJ Qual Saf*. 2017;26:350–353.
4. McCulloch P, Morgan L, New S, et al. Combining systems and teamwork approaches to enhance the effectiveness of safety improvement interventions in surgery. *Ann Surg*. 2017;265:90–96.
5. Wu AW, Lipshutz AKM, Pronovost PJ. Effectiveness and efficiency of root cause analysis in medicine. *JAMA*. 2008;299:685–687.
6. National Patient Safety Foundation. RCA2 improving root cause analyses and actions to prevent harm. 2015.
7. Kanji F, Catchpole K, Choi E, et al. Work-system interventions in robotic-assisted surgery: a systematic review exploring the gap between challenges and solutions. *Surg Endosc*. 2021;35:1976–1989.
8. Percarpio KB, Watts BV, Weeks WB. The effectiveness of root cause analysis: what does the literature tell us? *Jt Comm J Qual Patient Saf*. 2008;34:391–398.
9. Mills PD, Neily J, Luan D, et al. Actions and implementation strategies to reduce suicidal events in the Veterans Health Administration. *Jt Comm J Qual Patient Saf*. 2006;32:130–141.
10. Kellogg KM, Hettinger Z, Shah M, et al. Our current approach to root cause analysis: is it contributing to our failure to improve patient safety? *BMJ Qual Saf*. 2017;26:381–387.
11. Peerally MF, Carr S, Waring J, et al. The problem with root cause analysis. *BMJ Qual Saf*. 2017;26:417–422.
12. Okrainec A, Henao O, Azzie G. Telesimulation: an effective method for teaching the fundamentals of laparoscopic surgery in resource-restricted countries. *Surg Endosc*. 2010;24:417–422.
13. Youssef SC, Aydin A, Canning A, et al. Learning surgical skills through video-based education: a systematic review. *Surg Innov*. 2023;30:220–238.
14. Engelmann C, Schneider M, Kirschbaum C, et al. Effects of intraoperative breaks on mental and somatic operator fatigue: a randomized clinical trial. *Surg Endosc*. 2011;25:1245–1250.

15. Koshy K, Syed H, Luckiewicz A, et al. Interventions to improve ergonomics in the operating theatre: a systematic review of ergonomics training and intra-operative microbreaks. *Ann Med Surg (Lond)*. 2020;55:135–142.
16. Liu Z, Dumville JC, Norman G, et al. Intraoperative interventions for preventing surgical site infection: an overview of Cochrane reviews. *Cochrane Database Syst Rev*. 2018;2. doi:10.1002/14651858.CD012653.pub2
17. Higgins J, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.3*. Cochrane; 2022.
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
19. Tong A, Flemming K, McInnes E, et al. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Med Res Methodol*. 2012;12:181.
20. Scottish Intercollegiate Guidelines Network. Search Filters. 2021. Available at: https://www.sign.ac.uk/what-we-do/methodology/search-filters/. Accessed August 24, 2024.
21. Palter VN, Grantcharov TP. Simulation in surgical education. *CMAJ*. 2010;182:1191–1196.
22. Cardoso SA, Suyambu J, Iqbal J, et al. Exploring the role of simulation training in improving surgical skills among residents: a narrative review. *Cureus*. 2023;15:e44654.
23. Mazzone E, Puliatti S, Amato M, et al. A systematic review and meta-analysis on the impact of proficiency-based progression simulation training on performance outcomes. *Ann Surg*. 2021;274:281–289.
24. Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15:1277–1288.
25. Yule S, Flin R, Maran N, et al. Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World J Surg*. 2008;32:548–556.
26. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010;25:603–605.
27. National Heart Lung and Blood Institute. Study quality assessment tools. 2017. Available at: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools. Accessed August 24, 2024.
28. Nagpal K, Vats A, Lamb B, et al. Information transfer and communication in surgery. *Ann Surg*. 2010;252:225–239.
29. Ding H, Hu GL, Zheng XY, et al. The method quality of cross-over studies involved in Cochrane systematic reviews. *PLoS One*. 2015;10:e0120519.
30. Armstrong BA, Dutescu IA, Nemoy L, et al. Effect of the surgical safety checklist on provider and patient outcomes: a systematic review. *BMJ Qual Saf*. 2022;31:463–478.
31. Flynn JM, Striano BM, Muhly WT, et al. A dedicated pediatric spine deformity team significantly reduces surgical time and cost. *J Bone Joint Surg Am*. 2018;100:1574–1580.
32. Badani KK, Bhandari A, Tewari A, et al. Comparison of two-dimensional and three-dimensional suturing: is there a difference in a robotic surgery setting? *J Endourol*. 2005;19:1212–1215.
33. Jing J, Honey MLL. Using a checklist in robotic-assisted laparoscopic radical prostatectomy procedures. *AORN J*. 2016;104:145–152.
34. Lean K, Page BF, Vincent C. Improving communication at handover and transfer reduces retained swabs in maternity services. *Eur J Obstet Gynecol Reprod Biol*. 2018;220:50–56.
35. Fornwalt L, Ennis D, Stibich M. Influence of a total joint infection control bundle on surgical site infection rates. *Am J Infect Control*. 2016;44:239–241.
36. Rose MR, Rose KM. Use of a surgical debriefing checklist to achieve higher value health care. *Am J Med Qual*. 2018;33:514–522.
37. Wright MI. Implementing no interruption zones in the perioperative environment. *AORN J*. 2016;104:536–540.
38. Dugan JW, Weatherly RA, Girod DA, et al. A longitudinal study of emotional intelligence training for otolaryngology residents and faculty. *JAMA Otolaryngol Head Neck Surg*. 2014;140:720–726.
39. Attarian DE, Wahl JE, Wellman SS, et al. Developing a high-efficiency operating room for total joint arthroplasty in an academic setting. *Clin Orthop Relat Res*. 2013;471:1832–1836.
40. Mussa S, Drury NE, Stickley J, et al. Mentoring new surgeons: can we avoid the learning curve?†. *Eur J Cardiothorac Surg*. 2017;51: 291–299.
41. Wen L, Sweeney TE, Welton L, et al. Encouraging mindfulness in medical house staff via smartphone app: a pilot study. *Acad Psychiatry*. 2017;41:646–650.
42. Riall TS, Teiman J, Chang M, et al. Maintaining the fire but avoiding burnout: implementation and evaluation of a resident well-being program. *J Am Coll Surg*. 2018;226:369–379.
43. Bible JE, O'Neill KR, Crosby CG, et al. Implant contamination during spine surgery. *Spine J*. 2013;13:637–640.
44. Jelacic S, Bowdle A, Nair BG, et al. Aviation-style computerized surgical safety checklist displayed on a large screen and operated by the anesthesia provider improves checklist performance. *Anesth Analg*. 2020;130:382–390.
45. Hallbeck MS, Lowndes BR, Bingener J, et al. The impact of intraoperative microbreaks with exercises on surgeons: a multi-center cohort study. *Appl Ergon*. 2017;60:334–341.
46. Cassera MA, Goers TA, Spaun GO, et al. Efficacy of using a novel endoscopic lens cleaning device: a prospective randomized controlled trial. *Surg Innov*. 2011;18:150–155.
47. Kaplan G, Bo-Linn G, Carayon P, et al. Bringing a systems approach to health. *NAM Perspect*. 2013;3.
48. McCulloch P, Catchpole K. A three-dimensional model of error and safety in surgical health care microsystems. Rationale, development and initial testing. *BMC Surg*. 2011;11:23.
49. Carayon P, Schoofs Hundt A, Karsh B-T, et al. Work system design for patient safety: the SEIPS model. *Qual Saf Health Care*. 2006;15(Suppl 1):i50–i58.
50. van Munster JJCM, Zamanipoor Najafabadi AH, de Boer NP, et al. Impact of surgical intervention trials on healthcare: a systematic review of assessment methods, healthcare outcomes, and determinants. *PLoS One*. 2020;15:e0233318.
51. Bickler SW, Spiegel D. Improving surgical care in low- and middle-income countries: a pivotal role for the World Health Organization. *World J Surg*. 2010;34:386–390.