Review

Application of AI Models for Preventing Surgical Complications: Scoping Review of Clinical Readiness and Barriers to Implementation

Kjersti Mevik[1,2,3], MD, PhD; Ashenafi Zebene Woldaregay[4], MSc, PhD; Eva Lindell Jonsson[5,6], MD, PhD; Miguel Tejedor[7,8], MSc, PhD; Claire Temple-Oberle[5], MSc, MMEd, MD

[1]Department of Breast and Endocrine Surgery, Nordland Hospital Trust, Bodø, Norway

[2]Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3]Department of Clinical Medicine, UiT The Arctic University of Norway, Tromsø, Norway

[4]The Norwegian Centre for Clinical Artificial Intelligence, University Hospital of North Norway, Tromsø, Norway

[5]Departments of Surgery and Oncology, Section of Plastic and Reconstructive Surgery, University of Calgary, Calgary, AB, Canada

[6]Department of Surgical Sciences, Plastic and Maxillofacial Surgery, Uppsala University, Uppsala, Sweden

[7]Technology and AI, Norwegian Centre for E-health Research, Tromsø, Norway

[8]Department of Mathematics and Statistics, UiT The Arctic University of Norway, Tromsø, Norway

Corresponding Author:

Kjersti Mevik, MD, PhD
Department of Breast and Endocrine Surgery
Nordland Hospital Trust
Bodø 8092
Norway
Phone: 47 97123977
Email: kjersti.mevik@nlsh.no

Abstract

Background: The impact of surgical complications is substantial and multifaceted, affecting patients and their families, surgeons, and health care systems. Despite the remarkable progress in artificial intelligence (AI), there remains a notable gap in the prospective implementation of AI models in surgery that use real-time data to support decision-making and enable proactive intervention to reduce the risk of surgical complications.

Objective: This scoping review aims to assess and analyze the adoption and use of AI models for preventing surgical complications. Furthermore, this review aims to identify barriers and facilitators for implementation at the bedside.

Methods: Following PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines, we conducted a literature search using IEEE Xplore, Scopus, Web of Science, MEDLINE, ProQuest, PubMed, ABI, Embase, Epistemonikos, CINAHL, and Cochrane registries. The inclusion criteria included empirical, peer-reviewed studies published in English between January 2013 and January 2025, involving AI models for preventing surgical complications (surgical site infections, and heart and lung complications or stroke) in real-world settings. Exclusions included retrospective algorithm-only validations, nonempirical research (eg, editorials or protocols), and non-English studies. Study characteristics and AI model development details were extracted, along with performance statistics (eg, sensitivity and area under the receiver operating characteristic curve). We then used thematic analysis to synthesize findings related to AI models, prediction outputs, and validation methods. Studies were grouped into three main themes: (1) duration of hypotension, (2) risk for complications, and (3) decision support tool.

Results: Of the 275 identified records, 19 were included. The included models frequently demonstrated strong technical accuracy with high sensitivity and area under the receiver operating characteristic curve, particularly among studies evaluating decision support tools. However, only a few models were adopted routinely in clinical practice. Two studies evaluated the clinicians' perceptions regarding the use of AI models, reporting predominantly positive assessments of their usefulness.

Conclusions: Overall, AI models hold potential to predict and prevent surgical complications as the validation studies demonstrated high accuracy. However, implementation in routine practice remains limited by usability barriers, workflow misalignment, trust concerns, and financial and ethical constraints. The evidence included in this scoping review was limited by the heterogeneity in study design and the predominance of small-scale feasibility studies, particularly for hypotension

prediction. Future research should prioritize prospectively validated models that use other physiologic features and address clinicians' concerns regarding generalizability and adoption.

Introduction

With more than 320 million surgical procedures performed worldwide annually, there is a global responsibility to enhance the quality of surgical care [1]. Complications such as surgical site infections and stroke or heart and lung complications, whether minor or severe, often lead to reoperations, morbidity, and prolonged hospital stay [2,3]. Death following surgery approaches 5% and every tenth patient experiences preventable surgical complications [1]. Adverse events also impact surgeons as a second victim [4], and health care resource use rises notably [5]. The added cost of surgical complications ranges from US $3.5 to $10 billion yearly and is associated with an average increase in hospital stay of 11 days [6]. Despite attempts to optimize adherence to clinical pathways to reduce the frequency of surgical complications, complications persist [7,8].

Artificial intelligence (AI) is transforming the field of surgery, offering unprecedented advancements in precision, efficiency, and patient outcomes that could possibly reduce surgical complications [9]. For example, AI-assisted frame reviews in neuroscience demonstrate that AI can help both junior and senior clinicians perform better [10]. In educational platforms, AI is outpacing traditional coaching programs as demonstrated by a gallbladder surgery program [11]. By leveraging the power of machine learning, natural language processing, and computer vision, AI can enhance various aspects of surgical practice from preoperative planning to intraoperative guidance to reduce surgical complications [12].

During preoperative discussions, the surgeon and patient must weigh the benefits of surgery against the risks. AI can rapidly process large amounts of health data, unburdening health personnel and allowing them to better inform their patients [13]. These AI models analyze patterns and their relationship to determine complex combinations that can indicate the patient's risk for surgical complications. There is a gap in studies focusing on the adoption and clinical validation of these AI models [14-21]. Existing literature is focused on the retrospective development of models aimed at preventing surgical complications [9]. There are several reasons why these models have not been widely adopted in actual clinical use, including the lack of validation, lack of supporting data, differences in culture and behavior, and organizational structure [22,23]. Most common problems are regulatory and ethical constraints given that AI in surgery is considered high risk, time-consuming, and expensive. Randomized controlled trials (RCTs) are generally needed, which do not exist [24]. Also, there is currently a gap in the literature regarding the evaluation of AI models in clinical practice. This review aims to provide an overview of the published studies. Previous reviews have focused exclusively on studies involving the development and validation of models conducted using retrospective data [25-27]. Our focus is to uncover AI models that have been prospectively tested with real-time data at the bedside and to pinpoint the barriers and facilitators to implementing these models for the prevention of surgical complications.

Methods

The review was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [28]. We used a scoping review protocol that addressed key concepts and types of evidence. The aim was to map the existing literature by systematically searching, selecting, and synthesizing current evidence on AI models designed to prevent surgical complications using real-time data [29]. In this review, our definition of AI models also encompasses traditional statistical models, such as the National Surgical Quality Improvement Program (NSQIP), as these were tested prospectively. We relied on the five-stage framework proposed by Pollock et al [30]: (1) developing the review objective, (2) applying the eligibility criteria, (3) selecting the articles, (4) extracting and analyzing the data, and (5) reporting the results. The inclusion criteria were implemented using the population-concept-context framework, where population represents the patients undergoing surgery, concept includes the use of AI models to prevent surgical complications, and context includes studies that are conducted pre-, peri-, and intraoperative with both real-time and retrospective data excluding retrospective validation studies without prospective evaluation. The research team comprised both surgeons and machine learning engineers. The primary research question was "Which AI models have been clinically tested for the prediction of surgical complications?" The secondary research question was "For AI models not yet implemented in routine clinical practice, what barriers hinder their implementation, and are there any models on the horizon that could readily be adopted? " We included prospective, observational, and interventional peer-reviewed studies in English that included model development, validation, and implementation. A comprehensive search strategy was developed in collaboration with a medical librarian to identify peer-reviewed original studies. We searched the following databases from October 2024 to January 2025: Scopus, CINAHL, the Cochrane Library, PubMed, MEDLINE, Web of Science, Embase, Epistemonikos, and IEEE Xplore [28]. The search strategy combined controlled vocabulary (eg, Medical Subject

Headings and Emtree) and natural language keywords using Boolean operators and truncation to capture variations in terminology. The specific search terms and keywords were defined following iterative literature searches and several rounds of discussion among the authors.

The search query was structured around 3 key concepts:
- AI Methodology: ("Artificial intelligence" OR "Machine learning" OR "AI tool*" OR "AI model*" OR "Validated algorithm*");
- Function: ("Predict*" OR "Prediction tool*" OR "Prediction index" OR "Clinical decision support tool");
- Outcome: ("Postoperative complication*" OR "Surgical adverse event*" OR "Adverse surgical outcome*").

The search was restricted to articles published in English. Studies were included if they described prospective model validation or clinical implementation.

A reference list of selected articles was used to extract additional articles to get a complete overview of the field. Detailed information on the search strategy can be found in Multimedia Appendix 1. The librarian vetted the initial search, using Mendeley (version 2.129.0; Elsevier). Eligibility assessment and screening were independently conducted by the primary investigator (KM) and co-investigator (ELJ) based on the established inclusion and exclusion criteria. After the initial screening, a full-text assessment was carried out. Disagreements were arbitrated by a third reviewer (CT-O). The reporting quality of the included studies was assessed by using the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) with the AI statement [31,32]. The checklist includes 22 items (27 for the AI statement) with the potential answer options: "yes," "no," and "not applicable." The 2 reviewers assessed the included studies for compliance with the items described in the TRIPOD+AI checklist. Furthermore, the following information was extracted: bibliographic details, study design and setting, surgical specialty and procedure type, AI model and technical details, predicted outcome or complications, stage of implementation, validation, and reported barriers and facilitators to clinical implementation. Quantitative characteristics derived from the included studies were summarized using tables and figures. Thematic analysis was performed on qualitative data related to barriers and facilitators. For this scoping review, we developed an analytical categorization framework to systematically classify the included studies according to the primary applications of the AI models they used (Table 1). This framework served to structure the evidence by grouping studies into conceptually coherent domains, thereby facilitating a clearer understanding of the thematic focus, methodological approaches, and applied contexts represented across the studies.

Table 1. Thematic classification of 19 included studies in the scoping review.

Theme	Studies (N=19), n (%)	AI model
Duration of hypotension	11 (58)	HPI[a]
Risk for complications	4 (21)	POTTER[b], Periop ORACLE[c], MuscleSound, and PPC -score
Decision support tool	4 (21)	My Surgery Risk, ACS[e] NSQIP[f], SURPAS[g], and MyRISK

[a]HPI: Hypotension Prediction Index.
[b]POTTER: Predictive Optimal Trees in Emergency Surgery Risk.
[c]ORACLE: Outcome Risk Assessment with Computer Learning Enhancement.
[d]PPC: postoperative pulmonary complications.
[e]ACS: American College of Surgeons.
[f]NSQIP: National Surgical Quality Improvement Program.
[g]SURPAS: Surgical Risk Preoperative Assessment System.

Results

A PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [33] flow diagram, as shown in Figure 1, illustrates the study selection process. The initial search yielded 199 articles, and 76 additional articles were gleaned from reference lists, for a total of 275 records screened. Of these 275 records, 19 studies met the inclusion criteria for this scoping review. The majority of studies were conducted in high-income countries, with the United States (n=7) being the most frequent contributor. The studies used a prospective study design, including RCTs, pilot interventional studies, and prospective cohorts. There were 8 RCTs and 11 prospective studies with a strong trend toward pilot-scale prospective studies. Large-scale validation or postdeployment studies were lacking. Few studies were evaluated outside of controlled research settings. External validation of the AI models was infrequent and adherence to TRIPOD+AI was poor. No study fully met the criteria for transparent reporting (Table 2).

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) flowchart (reproduced from Haddaway et al [34] and is covered by Creative Commons license). AI: artificial intelligence.
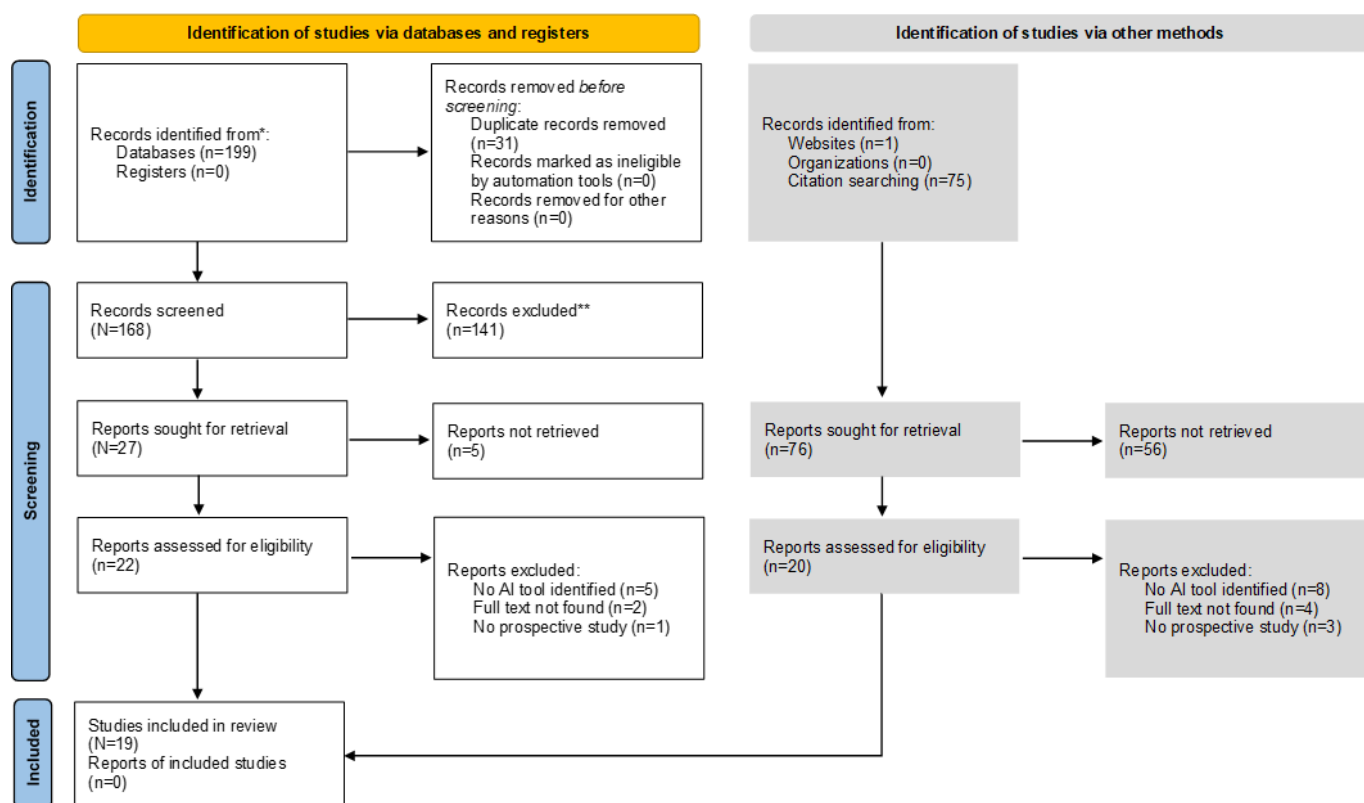

Table 2. Summary of adherence to TRIPOD+AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis + Artificial Intelligence) guidelines.

Study (year)	Model type	TRIPOD: outcome defined	TRIPOD: missing data	TRIPOD: internal validation	TRIPOD: external validation	TRIPOD-AI: algorithm description	TRIPOD-AI: explainability	TRIPOD-AI: bias assessment
Wijnberge et al (2020) [35]	ML[a]-derived HPI[b]	✓		✓				
Lorente et al (2023) [36]	HPI protocol	✓		✓				
Schneck et al (2020) [37]	HPI system	✓		✓				
Bao et al (2024) [38]	Acumen HPI	✓	✓	✓				
Tsoumpa et al (2021) [39]	HPI algorithm	✓		✓				
Cylwik et al (2024) [40]	HPI software	✓		✓				
Murabito et al (2022) [41]	ML proactive HPI	✓		✓				
Šribar et al (2023) [42]	HPI-guided therapy	✓		✓				
Maheshwari et al (2020) [43]	ML-HPI tool	✓		✓				
Andrzejewska et al (2023) [44]	HPI prediction	✓		✓				
Kouz et al (2023) [45]	HPI registry	✓		✓				
Ren et al (2022) [46]	ML postoperative	✓	✓	✓		✓	✓	

Study (year)	Model type	TRIPOD: outcome defined	TRIPOD: missing data	TRIPOD: internal validation	TRIPOD: external validation	TRIPOD-AI: algorithm description	TRIPOD-AI: explainability	TRIPOD-AI: bias assessment
Bilimoria et al (2013) [15]	ACS NSQIP Surgical Risk Calculator	✓	✓	✓	✓		✓	✓
Bronsert et al (2020) [47]	SURPAS[c] pilot	✓	✓	✓			✓	✓
El Moheb et al (2023) [48]	Surgeon's AI risk	✓		✓		✓		
Ferré et al (2023) [49]	MyRISK score	✓	✓	✓		✓		
Fritz et al (2024) [50]	ORACLE[d] ML model	✓	✓	✓		✓	✓	
Yik et al (2024) [51]	US sarcopenia AI	✓	✓	✓		✓	✓	✓
Li et al (2024) [52]	ML pulmonary outcome	✓	✓	✓		✓	✓	✓

[a]ML: Machine Learning.
[b]HPI: Hypotension Prediction Index.
[c]SURPAS: Surgical Risk Preoperative Assessment System.
[d]ORACLE: Outcome Risk Assessment with Computer Learning Enhancement.

To organize the heterogeneous literature, we developed a thematic categorization framework based on the primary intended function of each model as described in the original publications. The categorization was not based on the underlying statistical methodology, but on how the model output was framed and intended to be used in the clinical setting. Models categorized as "risk for complications" primarily focused on estimating the probability of specific postoperative outcomes, with performance evaluation as the main objective and limited emphasis on downstream clinical use. In contrast, models categorized as "decision support tools" were explicitly presented as supporting clinical decision-making processes, such as preoperative planning, shared decision-making, or patient counseling. We acknowledge that these categories are not mutually exclusive and that several models could reasonably fit more than one theme. In such cases, classification was based on the dominant emphasis in the study objectives and presentation (Table 1).

Study demographics and their respective funding are presented in Table 3. A total of 11 studies tested intraoperative hemodynamic monitoring and complication prediction using the Hypotension Prediction Index (HPI), while the remaining studies included AI models that addressed predicting general complications (n=7) and image analysis (n=1). While this review identified diversity of AI model applications, the majority of studies evaluated HPI, potentially skewing the findings through over-representation of intraoperative hemodynamic monitoring as the primary area of AI use in surgery. As such, the generalizability of results across surgical domains is narrowed. The predominance of HPI-related research may reflect a commercially available and well-integrated AI model, with greater funding and dissemination of pathways than other early-phase AI models. This may introduce publication and funding biases, where more rigorously tested, industry-supported models are over-represented compared to academic, noncommercial exploratory models. However, most of the studies testing the HPI had funding from the manufacturer Edwards Lifesciences. The authors stated that manufacturers were not involved in the conduct of the studies and did not approve or disapprove of the manuscript. Moreover, the clinical end points targeted by HPI are relatively narrow compared to the diverse risks associated with surgery. This limits the scope of AI use in surgery in terms of complication prediction, workflow optimization, and personalized surgical planning. Few studies evaluated AI models for long-term surgical complications, and the under-representation of other AI models narrows applicability. The main population studied was high-risk patients in 4 studies, with a subspecialty lens of noncardiac surgical patients. Table 4 delineated performance metrics of the AI models.

Table 3. Characteristics of the included articles.

Author (year) of publication	Country of origin	Clinical domain	Enrolled or planned participants	Type of study	Funding
Wijnberge et al (2020) [35]	The Netherlands	Noncardiac surgery	60	Unblinded randomized clinical trial: (1) early warning system with HPI[a] and (2) standard care	Edwards Lifesciences
Lorente et al (2023) [36]	Spain	High-risk surgical patients for elective major abdominal surgery	80	Parallel-arm double-blinded multicenter randomized trial: (1) HPI protocol and (2) standard care	Edwards Lifesciences
Schneck et al (2020) [37]	Germany	Patients undergoing primary hip arthroplasty.	99	Single center randomized blinded prospective trial: (1) therapy algorithm HPI, (2) standard care, and (3) historic control group	Edwards Lifesciences
Bao et al (2024) [38]	The United States	Patients with ASA[b] 3 or 4 moderate or high-risk noncardiac surgery >3 hours.	Prospective: 425 and post hoc analysis: 457 verss 15,796	Prospective single-arm multicenter (n=11) trial study: (1) continuous blood pressure measurements from study monitors compared to historical cohort with standard care study and (2) subset of trial participants versus a propensity score-weighted contemporaneous comparison group	Edwards Lifesciences
Tsoumpa et al (2021) [39]	Greece	Moderate or high-risk noncardiac surgery	99	Single-center prospective randomized trial: (1) HPI with hemodynamic treatment protocol and (2) standard care	None
Cylwik et al (2024) [40]	Poland	Patients undergoing oncological gastrointestinal surgery with ASA 3 or 4	46	Prospective single-center where HPI was used for 50 patients	None
Murabito et al (2022) [41]	Italy	Patients for elective major general surgery	40	Single-center pilot randomized clinical trial: (1) early warning system and (2) standard care	Edwards Lifesciences and the University of Catania
Šribar et al (2023) [42]	Croatia	Patients for elective major thoracic surgery with single lung ventilation	34	Prospective randomized single-center blinded trial: (1) "machine learning algorithm" (AcumenIQ) and (2) "conventional pulse contour analysis" (Flotrac)	None
Maheshwari et al (2020) [43]	The United States	Patients with ASA 3 or 4 for moderate or high-risk noncardiac surgery	214	Randomized multicenter controlled trial (n=2): (1) HPI guided group and (2) standard care	Edwards Lifesciences
Andrzejewska et al (2023) [44]	Poland	Patients undergoing posterior fusion for adolescent idiopathic scoliosis	59 adolescents	Prospective single-center, non-randomized, case-control study: (1) goal-directed therapy with HPI and (2) standard care	None

Author (year) of publication	Country of origin	Clinical domain	Enrolled or planned participants	Type of study	Funding
Ren et al (2022) [46]	The United States	Preoperative	67 surgeons testing the tool on 100 cases	Prospective	University of Florida, NIBIB[c], NIDDK[d], NIGMS[e], and the National Science Foundation
Bilimoria et al (2013) [15]	The United States	Preoperative	80 surgeons testing the tool on 10 cases	Prospective	Agency for Healthcare Research and Quality
Bronsert et al (2020) [47]	The United States	Preoperative	197 patients assessed by 9 surgeons, but 166 were assessed by the tool	Convergent prospective mixed methods with both quantitative and qualitative data.	Agency for Healthcare Research and Quality
El Moheb et al (2023) [48]	The United States	Emergency surgery	150 patients, 15 surgeons in each group	Prospective, nonblinded, single-center: (1) prediction with use of POTTER[f] and (2) standard prediction.	CRICO[g] or RMF[h] grant
Ferre et al (2023) [49]	France	Preoperative	389	Single-center prospective observational study	None
Kouz et al (2023) [45]	France, Germany, Italy, Spain, and the United Kingdom	Elective major noncardiac surgery	702	European multicenter (n=12) prospective observational trial	Edwards Lifescience
Fritz et al (2024) [50]	The United States	Patients for elective surgery during daytime weekdays	5071	Single-center prospective randomized clinical trial: (1) AlertWatch + ML[i] display and (2) standard care	National Institute of Nursing Research, the Foundation for Anesthesia Education and Research, and the Washington University School of Medicine
Yik et al (2024) [51]	Singapore	Elective major gastrointestinal surgery	36	Prospective cohort study	SingHealth Medical Student Talent Development Award
Li et al (2024) [52]	China	Patient underwent surgical procedure with general anesthesia and mechanical ventilation	307	Prospective cohort in a single-center	The National Natural Science Foundation of China, Technology Project of Sichuan, Postdoctoral Science Foundation, Postdoctoral Program of Sichuan University, the Postdoctoral Program of West China Hospital, Sichuan, the 1$3$5 Project for Disciplines of excellence, West China Hospital, the Sichuan Province Natural Science Foundation of China, and the CAMS[j] Innovation Fund for Medical Sciences

[a]HPI: Hypotension Prediction Index.

[b]ASA: American Society of Anesthesiologist Physical Status Classification System

[c]NIBIB: National Institutes of Health

[d]NIDDK: National Institute of Diabetes and Digestive and Kidney Diseases

[e]NIGMS: National Institute of General Medical Sciences

[f]POTTER: Predictive Optimal Trees in Emergency Surgery Risk.

[g]CRICO: Controlled Risk Insurance Company.

[h]RMF: Risk Management Foundation.

[i]ML: machine learning.

[j]CAMS: Chinese Academy of Medical Sciences

Table 4. Features and performance of the included artificial intelligence (AI) models.

Predicted outcome	Author (year) of publication	Features	Name of AI models	Performance metrics or clinical end points	Web-based calculators available
Duration of hypotension	Wijnberge et (2020) [35]	Arterial pressure waveform (28 variables)	HPI[a]	• Median average hypotension: 0.10 mmHg versus 0.44 mmHg • Median time of hypotension: 8 minutes and 33 minutes but no differences in adverse events • The algorithm (HPI) tested with AUC=0.89 of 5-minute prediction time	Algorithm derivation only [53]
Duration of hypotension	Lorente et al (2023) [36]	Intraoperative time-weighted average of MAP[c] <65 mm Hg, number of hypotension episodes, total time of hypotension, biomarkers of acute kidney distress, and tissue oxygenation	HPI	• Median average hypotension: 0.06 mmHg versus 0 mmHg • Median time of hypotension: 5 minutes and 0 minutes • No differences in oxygen saturation and acute kidney injury	Algorithm derivation only [53]
Duration of hypotension	Schneck et al (2020) [37]	—	HPI	• Duration of hypotension episodes: 0, 640, and 660 seconds	Algorithm derivation only [53]
Duration of hypotension	Bao et al (2024) [38]	Arterial pressure waveform and demographics, comorbidity, procedures, and acute kidney injury for post hoc analysis	HPI	• 58% reduction of MAP < 65 mmHg. In post hoc analysis, 35% reduction in minutes of hypotension • Median time of hypotension: 9 minutes versus 15 minutes • No difference in AKI[d] 13.8% versus 15.8%	Algorithm derivation only [53]
Duration of hypotension	Tsoumpa et al (2021) [39]	Intraoperative time-weighted average of MAP <65 mm Hg, number and time of hypotension, amount of medicines, IV fluid, transfusion, morbidity, and complications	HPI	• Median average hypotension: 0.16 mmHg and 0.50 mmHg • Median time of hypotension: 9 minutes and 24 minutes • No differences in complications or LOS[e] or use of medicine or IV	Algorithm derivation only [53]
Duration of hypotension	Cylwik et al (2024) [40]	Pre- and postoperatively proBNP[f] and troponin and acute kidney injury	HPI	• Median average hypotension: 0.085 mmHg • Median time of hypotension: 2 minutes • Hypotension associated with acute kidney injury but not with myocardial injury	Algorithm derivation only [53]
Duration of hypotension	Murabito et al (2022) [41]	Time-weighted average of hypotension and biomarkers	HPI	• Median average of hypotension: 0.12 mmHg and 0.37 mmHg • Median time of hypotension: 4.3 minutes and 21.3 minutes • Use of HPI reduced the intraoperative hypotension and biomarker for brain and oxidative stress	Algorithm derivation only [53]

Predicted outcome	Author (year) of publication	Features	Name of AI models	Performance metrics or clinical end points	Web-based calculators available
Duration of hypotension	Šribar et al (2023) [42]	Time-weighted average of hypotension, intravenous fluids, medicines, ICU stay, length of stay, acute kidney injury, coronary syndrome, or cerebrovascular infarction	HPI	• Median average of hypotension: 0.01 mmHg and 0.08 mmHg • Median time of hypotension: 0 and 3.7 minutes	Algorithm derivation only [53]
Duration of hypotension	Maheshwari et al (2020) [43]	Arterial pressure waveform	HPI	• Median average hypotension: 0.14 mmHg and 0.14 mmHg • Median time of hypotension: 2 minutes and 2 minutes	Algorithm derivation only [53]
Duration of hypotension	Andrzejewska et al (2023) [44]	Surgical time, intravenous fluids, blood values, length of stay, and cardiac and neurological complications	HPI	• Median hypotension time: 8 minutes and 40 minutes • Less time to extubation time for the HPI group (median 5 vs 27.5 min) • Noncardiac and neurological complications in the HPI group, while they were 4 in the control group	None
Duration of hypotension	Kouz et al (2023) [45]	Acute myocardial injury, acute kidney injury, death within 30 days after surgery, and hospital readmission within 30 days after surgery	HPI	• Median time of hypotension: 2 minutes. Median average hypotension: 0.03 mm Hg • 3% had acute myocardial injury, 9% had acute kidney injury • Postoperative mortality within 30 days after surgery was observed in 2%	Algorithm derivation only [53]
Decision support tool	Ren et al (2022) [46]	285 inputs and 8 outcomes: complications and death	My Surgery Risk	• Models with 135 features had AUC 0.80 to 0.92 for the different outcomes compared to lower AUC for models with 55 and 101 features • Surgeon predictive performance did not change significantly after viewing predictions generated by the algorithm	Algorithm development [54, 55]
Decision support tool	Bilimoria et al (2013) [15]	21 preoperative factors, 8 outcomes: mortality, morbidity, and 6 others	ACS[g] NSQIP[h] Surgical Risk Calculator	• Brier score for mortality: 0.011 and morbidity: 0.069 • Surgeons' agreement ranging from 80% to 100%	Algorithm development [56, 57]
Decision support tool	Bronsert et al (2020) [47]	Mortality, overall morbidity, unplanned readmission, and 19 preoperative variables	SURPAS[i]	• 98.8% reported they understood their surgical risks very or quite well after exposure to SURPAS; 92.7% reported SURPAS was very helpful or helpful. Providers shared that 83.4% of the time they reported SURPAS was very or somewhat helpful; 44.7% of the time the providers reported it changed their interaction with the	None, but the algorithm can be found [58]

Predicted outcome	Author (year) of publication	Features	Name of AI models	Performance metrics or clinical end points	Web-based calculators available
				patient and this change was beneficial 94.3% of the time	
Risk for complication	El Moheb et al (2023) [48]	8 variables	POTTER[j]	• POTTER outperformed surgeons in predicting mortality—AUC: 0.880 versus 0.841; ventilator dependence—AUC: 0.928 versus 0.833; bleeding—AUC: 0.832 versus 0.735; pneumonia—AUC: 0.837 versus 0.753	None, but the algorithm can be found [17]
Decision support tool	Ferre et al (2023) [49]	25 variables	MyRISK	• AUC 0.71, sensitivity 94%, NPV[k] 99%, specificity 49%, and PPV[l] 7% • Patient satisfaction 8/10 and usability 90/100	None, but validated in the same study.
Risk for complications	Fritz et al (2024) [50]	Variables within comorbidity, preoperative vital sign, preoperative laboratories, intraoperative time series, and medication and fluids	Periop ORACLE[m]	• AUC for AKI: 0.73 and 0.69 • Death: 0.79 and 0.78 • No significant difference in prediction with the use of the model	None, but the algorithm can be found [59]
Risk for complications	Yik et al (2024) [51]	Intramuscular adipose tissue as a proxy for muscle quality obtained by ultrasound	MuscleSound	• AUC 0.73 • Clinicians using the tool can have a robust diagnostic tool to help predict surgical risk and outcomes	None
Risk for complications	Li et al (2024) [52]	20 variables	PPC[n] score	• AUC 0.88, simplified model AUC 0.86 • Real-time identification of surgical patients' risk of postoperative pulmonary complications could help personalize intraoperative ventilatory strategies and reduce postoperative pulmonary complications	[60]

[a]HPI: Hypotension Prediction Index.
[b]AUC: area under the curve.
[c]MAP: mean arterial pressure
[d]AKI: acute kidney injury.
[e]LOS: length of stay.
[f]proBNP: pro-B-type natriuretic peptide.
[g]ACS: American College of Surgeons.
[h]NSQIP: National Surgical Quality Improvement Program.
[i]SURPAS: Surgical Risk Preoperative Assessment System.
[j]POTTER: Predictive Optimal Trees in Emergency Surgery Risk.
[k]NPV: negative predictive value.
[l]PPV: positive predictive value.
[m]ORACLE: Outcome Risk Assessment with Computer Learning Enhancement.
[n]PPC: postoperative pulmonary complications.

The American College of Surgeons (ACS) NSQIP Surgical Risk calculator is a widely adopted quality improvement tool used globally, and the HPI is a commercially available, regulatory-approved medical device (Acumen IQ) deployed in operating rooms. However, few of the other tools are deployed in surgical practice. In this regard, some of the common barriers included lack of external validation, limited generalizability, and black box model opacity (Table 5). Clinicians reported low trust in AI models (AI illiteracy and workflow issues as barriers) but noted real-time performance

benefits (integration with existing platforms and clinical support as facilitators). Most of the AI models required high implementation costs, and together with the lack of financial incentives and reimbursement structures, these represented the greatest challenges to implementation.

Table 5. The most frequently cited barriers and facilitators in each of the studies.

Author (year of publication)	Barriers	Facilitators
Wijnberge et al (2020) [35]	Early warning system software needed (Flotrac IQ pressure transducer connected to the HemoSphere monitor)	No facilitator identified
Lorente et al (2023) [36]	Defining the correct range for normal blood pressure. FloTrac sensor (GDHT[a] protocol) and AcumenIQ sensor (HPI[b] protocol) needed.	No facilitator identified
Schneck et al (2020) [37]	Not mentioned	Implementation of HPI was considered uncomplicated providing a high user compliance
Bao et al (2024) [38]	FloTrac IQ sensor and EV1000 platform needed	No facilitator identified
Tsoumpa et al (2021) [39]	Acumen Flo-Traq transducer and EV1000 platform needed	No facilitator identified
Cylwik et al (2024) [40]	Need the HemoSphere monitoring platform, equipped with the AcumenTM IQ sensor	Anesthesiologists need training in the use of HPI software
Kouz et al (2023) [45]	Acumen IQ sensor (Edwards Lifesciences) and the HemoSphere monitoring platform are needed	Each center needs a clinical routine for hypotension procedures
Murabito et al (2022) [41]	FloTrac IQ sensor with EWS[c] software needed	No facilitator identified
Šribar et al (2023) [42]	Hemosphere monitoring platform using either AcumenIQ or Flotrac sensors is needed	No facilitator identified
Maheshwari et al (2020) [43]	EV1000 is needed	The waveform needs to be acceptable using a fast flush test
Andrzejewska et al (2023) [44]	Acumen sensor and Hemosphere monitor are needed	Fast flush test was needed
Ren et al (2022) [46]	Fully automated data entry and mobile device outputs require a system architecture as a scalable real-time platform	Model outputs were provided to mobile device apps
Bilimoria et al (2013) [15]	Variables are manually added in the calculator	Allows clinicians to decrease the risk of surgery within the confidence interval for the predicted risk
Bronsert et al (2020) [47]	Variables are manually added in the calculator	Increase the interaction between the patient and the surgeon and make the patients able to understand the procedure and risk of the surgery.
El Moheb et al (2023) [48]	Variables are manually added in the calculator	Improved the surgeons' prediction
Ferre et al (2023) [49]	A digital questionnaire had to be filled out by the patients	The risks were visually illustrated with green (low), orange (intermediate), and red (high)
Fritz et al (2024) [50]	Variables are manually added in the calculator	No facilitator identified
Yik et al (2024) [51]	Need an ultrasound and the software Musclesound	Bedside and easy to use
Li et al (2024) [52]	Data developed on an older adult demographic	Easy to use and available online

[a]GDHT: goal-directed hemodynamic therapy
[b]HPI: Hypotension Prediction Index.
[c]EWS: early warning system

Discussion

Principal Findings

This scoping review highlights that only a small number of AI-based models have progressed to clinical use. Notably, ACS NSQIP is widely implemented as a quality improvement and risk stratification tool, whereas the HPI represents one of the few regulatory-approved AI-based medical devices that is ready for integration into routine clinical practice. The review identified several promising AI models that could help clinicians improve outcomes for surgical patients [61]. Although the models have demonstrated usefulness, important limitations remain regarding clinical use. Most of the studies reported a lack of widespread adoption. Although this review uses the umbrella term "AI-based models," it is important to acknowledge the methodological heterogeneity of the included tools. Several widely used systems, such as ACS NSQIP, are based on traditional statistical approaches, primarily logistic regression, rather than modern machine-learning techniques. These models were included in accordance with our predefined search strategy, which intentionally

captured both established statistical risk calculators and newer machine learning–based models used for surgical risk prediction and decision support. Importantly, traditional statistical models and machine learning algorithms differ in terms of model development, interpretability, data requirements, and generalizability. While logistic regression–based tools such as ACS NSQIP remain highly influential due to their transparency, validation history, and clinical acceptance, newer machine learning approaches offer potential advantages in handling complex, high-dimensional data but often face greater challenges related to interpretability, external validation, and clinical implementation. Distinguishing between these methodological paradigms is essential when interpreting the maturity and clinical readiness of AI-based tools in surgery.

Among the included models, HPI stood out as the most clinically mature and widely tested model, supported by multiple RCTs and integrated into well-known platforms. Its performance consistency and real-time application make it the most implementation-ready model in the surgical field. In contrast, the ACS NSQIP and ORACLE (Outcome Risk Assessment with Computer Learning Enhancement) demonstrated strong interpretability and user engagement but are designed exclusively for a preoperative decision support context, not intraoperative intervention. We observed that regulatory and ethical uncertainty is the most common reason why these models are not adopted into clinical practice. The lack of financial incentives to deploy AI is another barrier, elucidating why AI model development is progressing rapidly while translational science and implementation research lag behind [62]. A mixed method study suggested that barriers to implementing AI in clinical practice could be overcome by identifying and preparing champions, conducting educational meetings, promoting adaptability, and developing and disseminating educational materials on the AI model [63].

MySurgeryRisk, a tool developed by researchers at the University of Florida, uses machine learning to process vast amounts of patient data and clinical metrics and represents a promising predictive AI model with a high degree of accuracy. It is designed to provide real-time, actionable insights to surgeons, leading to better patient outcomes and optimized resource allocation [46,54]. A limitation of this model is that the predictions are mostly linear and do not account for combinations of variables that should be given greater weight when calculating risk. The augmentOR Portal developed by Asensus Surgical specifically evaluates the surgeon's performance and identifies areas for improvement. This could reduce surgical complications by enhancing technical surgical skills but has yet to be trialed clinically [64]. The implementation of AI in predicting surgical complications is marked by these innovative approaches and promising results, yet its integration into routine clinical practice faces barriers [12,65].

Barriers that impede the widespread adoption of potentially transformative AI models in health care are several [66,67]. First, the quality and comprehensiveness of the data used to train these models are critical. AI models require extensive, well-annotated clinical data to learn effectively, and this data must be continually updated to reflect contemporary medical knowledge and practice [66]. Integrating AI models into existing health care IT infrastructures can be technically challenging and costly, necessitating significant upfront investment and ongoing maintenance, as well as extensive training of health care providers [68]. There are also substantial regulatory hurdles. AI models require rigorous testing and approval processes to ensure they meet clinical safety and efficacy standards. Ethical considerations, such as protecting patient privacy and avoiding biases in AI models, must be carefully managed to prevent disparities in health care outcomes [69]. For instance, the review by de Keijzer et al [70] highlighted that despite the potential of AI to transform clinical decision-making, there is a notable translational gap from proof-of-concept to clinical use. This gap is often due to regulatory uncertainties, organizational challenges, and attitudinal barriers among health care professionals. These barriers slow the uptake and adoption of AI models, even in cases where they have proved to significantly benefit patient care, such as in managing stroke complications [71]. Another reason for resistance to the implementation of AI is the skepticism of health care professionals towards AI models. Clinicians are cautious about relying on AI for decision-making, concerned that it may overlook individual patient nuances or erode their clinical autonomy [72,73]. Additionally, economic implications cannot be overlooked. The development, testing, and deployment of AI models require substantial financial resources, which can be a barrier for less well-funded health care institutions [74]. Finally, there is a worry that AI could become a substitute rather than a support for clinical decision-making, potentially leading to an erosion of clinicians' professional skills [73]. Moreover, research exploring clinicians' perceptions of AI underscores concerns regarding workload, risk, trust, and the integration of AI into clinical settings. Many clinicians fear that AI may increase their workload or change their workflow in ways that could compromise patient care. They also voice concerns about relying too heavily on technology that may not always account for the complex realities of medical practice [75]. Manual data entry is not feasible when the number of features is even moderately high. There would ideally be a bridge between AI models and the electronic health records, minimizing the effort for the clinicians to use the models. The integration should be a seamless solution, preferably as automated data pipelines that would facilitate implementations.

This review also reveals less explored but potentially transformative opportunities for advancing AI in surgery. For instance, embedding AI models into surgical training programs may foster early adoption and familiarity among new clinicians. Training curricula that include model interpretation and ethical consideration could empower the next generation of clinicians to embrace AI. Another often overlooked area is adaptive interface design. Many AI models fail because of poor integration into surgical workflows. Designing interfaces that adapt in real-time to the clinicians' needs could make adoption more intuitive. Collaboration

between data scientists and clinicians could advance in that direction. Moreover, the implementation of AI models in surgery still faces challenges in regulatory requirements, with a lack of alignment with existing clinical guidelines. A potential solution is the creation of sandbox-controlled environments where AI models can be evaluated and tested under close clinical and ethical oversight. Such frameworks, already explored in Fintech and digital health, could allow iterative deployment without compromising patient safety [76].

Regarding hypotension specifically, there have been several evaluations of HPI [77,78], including an RCT showing the efficacy of the HPI [70]. Retrospective studies have demonstrated that HPI reduces hypotension [79] which is associated with acute kidney injury and myocardial injury [80,81], and decreases mechanical ventilation time and length of intensive care unit stay [82]. To monitor the effect of HPI, a European registry has been established [83]. HPI is also used in a protocol to measure oxygen saturation and predict free flap survival [84]. HPI is ready for broader adoption, which may pave the way for more AI models in surgery. However, when it comes to decision support tools, once surgery is planned, the tool may not alter clinicians' decisions due to the complexity of the inputs into existing prediction models. One concern with using AI for surgical decisions is the difficulty of integrating complex AI predictions into the nuanced and highly individualized process of surgical planning. AI models must accurately interpret and analyze medical images, the patient's history, and other data to suggest surgical interventions. Currently, clinical judgment that is required for surgical decisions involves factors beyond what AI can predict, such as patient preferences, surgeon experience, and intraoperative findings. Thus, while AI can support decision-making and enhance specific tasks, it cannot replace the expert judgment of experienced surgeons [85].

Conclusions

In conclusion, this scoping review demonstrates that despite substantial research activity, only a limited number of predictive models have been adopted into routine surgical practice. Most clinically implemented systems are based on traditional statistical models, such as ACS NSQIP, whereas only a few machine learning–based models, including the regulatory-approved HPI, have progressed toward clinical deployment. While these technologies show promise in improving perioperative risk prediction and physiological monitoring, current evidence does not consistently demonstrate downstream improvements in surgical outcomes. Continued technological advancements that can be deployed prospectively in controlled environments are important next steps. Such efforts are essential to safeguard patient safety, support the development of AI-specific reimbursement pathways within hospital budgets, and facilitate the integration of AI concepts into medical education to prepare future clinicians for AI-assisted clinical practice.

Authors' Contributions

Conceptualization: KM, MT, CT-O
Data curation: KM, ELJ
Formal analysis: KM, ELJ, CT-O
Funding acquisition: CT-O
Investigation: KM, ELJ, CT-O
Methodology: KM, AZW, ELJ, MT
Project administration: KM
Resources: KM, AZW, ELJ
Supervision: KM
Visualization: KM, AZW
Writing – original draft: KM, AZW
Writing – review & editing: ELJ, MT, CT-O

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.
[DOCX File (Microsoft Word File), 28 KB-Multimedia Appendix 1]

Checklist 1

PRISMA-ScR checklist.
[DOCX File (Microsoft Word File), 79 KB-Checklist 1]

References

1. Dobson GP. Trauma of major surgery: a global problem that is not going away. Int J Surg. Sep 2020;81:47-54. [doi: 10.1016/j.ijsu.2020.07.017] [Medline: 32738546]

2. Ludbrook GL. The hidden pandemic: the cost of postoperative complications. Curr Anesthesiol Rep. 2022;12(1):1-9. [doi: 10.1007/s40140-021-00493-y] [Medline: 34744518]

3. Siddaiah-Subramanya M, To H, Haigh C. The psychosocial impact of surgical complications on the operating surgeon: a scoping review. Ann Med Surg (Lond). Jul 2021;67:102530. [doi: 10.1016/j.amsu.2021.102530] [Medline: 34276982]

4. Luu S, Patel P, St-Martin L, et al. Waking up the next morning: surgeons' emotional reactions to adverse events. Med Educ. Dec 2012;46(12):1179-1188. [doi: 10.1111/medu.12058] [Medline: 23171260]

5. Chong RIH, Yaow CYL, Chong NZY, et al. Scoping review of the second victim syndrome among surgeons: understanding the impact, responses, and support systems. Am J Surg. Mar 2024;229:5-14. [doi: 10.1016/j.amjsurg.2023.09.045] [Medline: 37838505]

6. Seidelman JL, Mantyh CR, Anderson DJ. Surgical site infection prevention: a review. JAMA. Jan 17, 2023;329(3):244-252. [doi: 10.1001/jama.2022.24075] [Medline: 36648463]

7. Javed H, Olanrewaju OA, Ansah Owusu F, et al. Challenges and solutions in postoperative complications: a narrative review in general surgery. Cureus. Dec 2023;15(12):e50942. [doi: 10.7759/cureus.50942] [Medline: 38264378]

8. Bolcato M, Rodriguez D, Aprile A. Guiding principles for surgical pathways: a tool for improving outcomes and patient safety. Front Public Health. 2022;10:869607. [doi: 10.3389/fpubh.2022.869607] [Medline: 35462846]

9. Hassan AM, Rajesh A, Asaad M, et al. Artificial intelligence and machine learning in prediction of surgical complications: current state, applications, and implications. Am Surg. Jan 2023;89(1):25-30. [doi: 10.1177/00031348221101488] [Medline: 35562124]

10. Williams SC, Zhou J, Muirhead WR, et al. Artificial intelligence assisted surgical scene recognition: a comparative study amongst healthcare professionals. Ann Surg. Oct 30, 2024. [doi: 10.1097/SLA.0000000000006577] [Medline: 39474680]

11. Wu S, Tang M, Liu J, et al. Impact of an AI-based laparoscopic cholecystectomy coaching program on the surgical performance: a randomized controlled trial. Int J Surg. Dec 1, 2024;110(12):7816-7823. [doi: 10.1097/JS9.0000000000001798] [Medline: 38896869]

12. Guni A, Varma P, Zhang J, Fehervari M, Ashrafian H. Artificial intelligence in surgery: the future is now. Eur Surg Res. Jan 22, 2024. [doi: 10.1159/000536393] [Medline: 38253041]

13. Loftus TJ, Altieri MS, Balch JA, et al. Artificial intelligence–enabled decision support in surgery: state-of-the-art and future directions. Ann Surg. Jul 1, 2023;278(1):51-58. [doi: 10.1097/SLA.000000000005853] [Medline: 36942574]

14. Raymond BL, Wanderer JP, Hawkins AT, et al. Use of the American College of Surgeons National Surgical Quality Improvement Program Surgical Risk Calculator during preoperative risk discussion: the patient perspective. Anesth Analg. Apr 2019;128(4):643-650. [doi: 10.1213/ANE.0000000000003718] [Medline: 30169413]

15. Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. J Am Coll Surg. Nov 2013;217(5):833-42.e1-3. [doi: 10.1016/j.jamcollsurg.2013.07.385] [Medline: 24055383]

16. Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. Circulation. Jul 26, 2011;124(4):381-387. [doi: 10.1161/CIRCULATIONAHA.110.015701] [Medline: 21730309]

17. Bertsimas D, Dunn J, Velmahos GC, Kaafarani HMA. Surgical risk is not linear: derivation and validation of a novel, user-friendly, and machine-learning—based Predictive OpTimal Trees in Emergency Surgery Risk (POTTER) calculator. Ann Surg. Oct 2018;268(4):574-583. [doi: 10.1097/SLA.0000000000002956] [Medline: 30124479]

18. Podrat JL, Del Val FR, Pei KY. Evolution of risk calculators and the dawn of artificial intelligence in predicting patient complications. Surg Clin North Am. Feb 2021;101(1):97-107. [doi: 10.1016/j.suc.2020.08.012] [Medline: 33212083]

19. Zhang H, Zhao J, Farzan R, Alizadeh Otaghvar H. Risk predictions of surgical wound complications based on a machine learning algorithm: a systematic review. Int Wound J. Jan 2024;21(1):e14665. Retracted in: Int Wound J. 2025;22(3):e70280. [doi: 10.1111/iwj.70280] [Medline: 40031913]

20. Mevik K, Zebene Woldaregay A, Ringdal A, Øyvind Mikalsen K, Xu Y. Exploring surgical infection prediction: a comparative study of established risk indexes and a novel model. Int J Med Inform. Apr 2024;184:105370. [doi: 10.1016/j.ijmedinf.2024.105370] [Medline: 38341999]

21.   Chen W, Lu Z, You L, Zhou L, Xu J, Chen K. Artificial intelligence–based multimodal risk assessment model for surgical site infection (AMRAMS): development and validation study. JMIR Med Inform. Jun 15, 2020;8(6):e18186. [doi: 10.2196/18186] [Medline: 32538798]

22.   Scardoni A, Balzarini F, Signorelli C, Cabitza F, Odone A. Artificial intelligence–based tools to control healthcare associated infections: a systematic review of the literature. J Infect Public Health. Aug 2020;13(8):1061-1077. [doi: 10.1016/j.jiph.2020.06.006] [Medline: 32561275]

23.   Fischer UM, Shireman PK, Lin JC. Current applications of artificial intelligence in vascular surgery. Semin Vasc Surg. Dec 2021;34(4):268-271. [doi: 10.1053/j.semvascsurg.2021.10.008] [Medline: 34911633]

24.   Cobianchi L, Verde JM, Loftus TJ, et al. Artificial intelligence and surgery: ethical dilemmas and open issues. J Am Coll Surg. Aug 1, 2022;235(2):268-275. [doi: 10.1097/XCS.0000000000000242] [Medline: 35839401]

25.   Amin A, Cardoso SA, Suyambu J, et al. Future of artificial intelligence in surgery: a narrative review. Cureus. Jan 2024;16(1):e51631. [doi: 10.7759/cureus.51631] [Medline: 38318552]

26.   Varghese C, Harrison EM, O'Grady G, Topol EJ. Artificial intelligence in surgery. Nat Med. May 2024;30(5):1257-1268. [doi: 10.1038/s41591-024-02970-3] [Medline: 38740998]

27.   Morris MX, Fiocco D, Caneva T, Yiapanis P, Orgill DP. Current and future applications of artificial intelligence in surgery: implications for clinical practice and research. Front Surg. 2024;11:1393898. [doi: 10.3389/fsurg.2024.1393898] [Medline: 38783862]

28.   McGowan J, Straus S, Moher D, et al. Reporting scoping reviews—PRISMA ScR extension. J Clin Epidemiol. Jul 2020;123:177-179. [doi: 10.1016/j.jclinepi.2020.03.016] [Medline: 32229248]

29.   Colquhoun HL, Levac D, O'Brien KK, et al. Scoping reviews: time for clarity in definition, methods, and reporting. J Clin Epidemiol. Dec 2014;67(12):1291-1294. [doi: 10.1016/j.jclinepi.2014.03.013] [Medline: 25034198]

30.   Pollock D, Davies EL, Peters MDJ, et al. Undertaking a scoping review: a practical guide for nursing and midwifery students, clinicians, researchers, and academics. J Adv Nurs. Apr 2021;77(4):2102-2113. [doi: 10.1111/jan.14743] [Medline: 33543511]

31.   Collins GS, Dhiman P, Andaur Navarro CL, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open. Jul 9, 2021;11(7):e048008. [doi: 10.1136/bmjopen-2020-048008] [Medline: 34244270]

32.   Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. Apr 16, 2024;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]

33.   Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. Mar 29, 2021;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]

34.   Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. *PRISMA2020*: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. Campbell Syst Rev. Jun 2022;18(2):e1230. [doi: 10.1002/cl2.1230] [Medline: 36911350]

35.   Wijnberge M, Geerts BF, Hol L, et al. Effect of a machine learning–derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the HYPE randomized clinical trial. JAMA. Mar 17, 2020;323(11):1052-1060. [doi: 10.1001/jama.2020.0592] [Medline: 32065827]

36.   Lorente JV, Ripollés-Melchor J, Jiménez I, et al. Intraoperative hemodynamic optimization using the hypotension prediction index vs. goal-directed hemodynamic therapy during elective major abdominal surgery: the Predict-H multicenter randomized controlled trial. Front Anesthesiol. 2023;2. [doi: 10.3389/fanes.2023.1193886]

37.   Schneck E, Schulte D, Habig L, et al. Hypotension Prediction Index based protocolized haemodynamic management reduces the incidence and duration of intraoperative hypotension in primary total hip arthroplasty: a single centre feasibility randomised blinded prospective interventional trial. J Clin Monit Comput. Dec 2020;34(6):1149-1158. [doi: 10.1007/s10877-019-00433-6] [Medline: 31784852]

38.   Bao X, Kumar SS, Shah NJ, et al. AcumenTM hypotension prediction index guidance for prevention and treatment of hypotension in noncardiac surgery: a prospective, single-arm, multicenter trial. Perioper Med (Lond). Mar 4, 2024;13(1):13. [doi: 10.1186/s13741-024-00369-9] [Medline: 38439069]

39.   Tsoumpa M, Kyttari A, Matiatou S, et al. The use of the Hypotension Prediction Index integrated in an algorithm of goal directed hemodynamic treatment during moderate and high-risk surgery. J Clin Med. Dec 15, 2021;10(24):5884. [doi: 10.3390/jcm10245884] [Medline: 34945177]

40.   Cylwik J, Celińska-Spodar M, Dudzic M. Individualized perioperative hemodynamic management using Hypotension Prediction Index software and the dynamics of troponin and NTproBNP concentration changes in patients undergoing oncological abdominal surgery. J Pers Med. Feb 16, 2024;14(2):211. [doi: 10.3390/jpm14020211] [Medline: 38392644]

41. Murabito P, Astuto M, Sanfilippo F, et al. Proactive management of intraoperative hypotension reduces biomarkers of organ injury and oxidative stress during elective non-cardiac surgery: a pilot randomized controlled trial. J Clin Med. Jan 13, 2022;11(2):392. [doi: 10.3390/jcm11020392] [Medline: 35054083]

42. Šribar A, Jurinjak IS, Almahariq H, et al. Hypotension Prediction Index guided versus conventional goal directed therapy to reduce intraoperative hypotension during thoracic surgery: a randomized trial. BMC Anesthesiol. Mar 30, 2023;23(1):101. [doi: 10.1186/s12871-023-02069-1] [Medline: 36997847]

43. Maheshwari K, Shimada T, Yang D, et al. Hypotension Prediction Index for prevention of hypotension during moderate-to high-risk noncardiac surgery. Anesthesiology. Dec 1, 2020;133(6):1214-1222. [doi: 10.1097/ALN.0000000000003557] [Medline: 32960954]

44. Andrzejewska A, Miegoń J, Zacha S, et al. The impact of intraoperative haemodynamic monitoring, prediction of hypotension and goal-directed therapy on the outcomes of patients treated with posterior fusion due to adolescent idiopathic scoliosis. J Clin Med. Jul 9, 2023;12(14):4571. [doi: 10.3390/jcm12144571] [Medline: 37510686]

45. Kouz K, Monge García MI, Cerutti E, et al. Intraoperative hypotension when using Hypotension Prediction Index software during major noncardiac surgery: a European multicentre prospective observational registry (EU HYPROTECT). BJA Open. Jun 2023;6:100140. [doi: 10.1016/j.bjao.2023.100140] [Medline: 37588176]

46. Ren Y, Loftus TJ, Datta S, et al. Performance of a machine learning algorithm using electronic health record data to predict postoperative complications and report on a mobile platform. JAMA Netw Open. May 2, 2022;5(5):e2211973. [doi: 10.1001/jamanetworkopen.2022.11973] [Medline: 35576007]

47. Bronsert MR, Lambert-Kerzner A, Henderson WG, et al. The value of the "Surgical Risk Preoperative Assessment System" (SURPAS) in preoperative consultation for elective surgery: a pilot study. Patient Saf Surg. 2020;14(1):31. [doi: 10.1186/s13037-020-00256-4] [Medline: 32724336]

48. El Moheb M, Gebran A, Maurer LR, et al. Artificial intelligence versus surgeon gestalt in predicting risk of emergency general surgery. J Trauma Acute Care Surg. Oct 1, 2023;95(4):565-572. [doi: 10.1097/TA.0000000000004030] [Medline: 37314698]

49. Ferré F, Laurent R, Furelau P, et al. Perioperative risk assessment of patients using the MyRISK digital score completed before the preanesthetic consultation: prospective observational study. JMIR Perioper Med. Jan 16, 2023;6:e39044. [doi: 10.2196/39044] [Medline: 36645704]

50. Fritz BA, King CR, Abdelhack M, et al. Effect of machine learning models on clinician prediction of postoperative complications: the perioperative ORACLE randomised clinical trial. Br J Anaesth. Nov 2024;133(5):1042-1050. [doi: 10.1016/j.bja.2024.08.004] [Medline: 39261226]

51. Yik V, Kok SSX, Chean E, et al. Diagnosing sarcopenia with AI-aided ultrasound (DINOSAUR)—a pilot study. Nutrients. Aug 20, 2024;16(16):2768. [doi: 10.3390/nu16162768] [Medline: 39203903]

52. Li P, Gao S, Wang Y, et al. Utilising intraoperative respiratory dynamic features for developing and validating an explainable machine learning model for postoperative pulmonary complications. Br J Anaesth. Jun 2024;132(6):1315-1326. [doi: 10.1016/j.bja.2024.02.025] [Medline: 38637267]

53. Hatib F, Jian Z, Buddi S, et al. Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. Anesthesiology. Oct 2018;129(4):663-674. [doi: 10.1097/ALN.0000000000002300] [Medline: 29894315]

54. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, et al. MySurgeryRisk: development and validation of a machine-learning risk algorithm for major complications and death after surgery. Ann Surg. Apr 2019;269(4):652-662. [doi: 10.1097/SLA.0000000000002706] [Medline: 29489489]

55. Welcome to mysurgeryrisk calculator. URL: https://prismap.medicine.ufl.edu/x/com-dom-prisma-p.sites.medinfo.ufl.edu/MySurgeryRisk.html [Accessed 2026-02-12]

56. Cohen ME, Bilimoria KY, Ko CY, Hall BL. Development of an American College of Surgeons National Surgery Quality Improvement Program: morbidity and mortality risk calculator for colorectal surgery. J Am Coll Surg. Jun 2009;208(6):1009-1016. [doi: 10.1016/j.jamcollsurg.2009.01.043] [Medline: 19476884]

57. ACS NSQIP surgical risk calculator; version: 4.0.2. URL: https://riskcalculator.facs.org/RiskCalculator/ [Accessed 2026-02-12]

58. Meguid RA, Bronsert MR, Juarez-Colunga E, Hammermeister KE, Henderson WG. Surgical Risk Preoperative Assessment System (SURPAS): III. Accurate preoperative prediction of 8 adverse outcomes using 8 predictor variables. Ann Surg. Jul 2016;264(1):23-31. [doi: 10.1097/SLA.0000000000001678] [Medline: 26928465]

59. Cui Z, Fritz BA, King CR, Avidan MS, Chen Y. A factored generalized additive model for clinical decision support in the operating room. AMIA Annu Symp Proc. 2020;2019:343-352. [Medline: 32308827]

60. Risk calculator. URL: https://aorm.wchscu.cn/ [Accessed 2026-02-12]

61. Xu H, Fu C, Zhao W, et al. Anesthesia transformed: AI pioneering a new era in perioperative medicine. Anesthesiol Perioper Sci. 2025;3(1):1-4. [doi: 10.1007/s44254-025-00091-9]

62. Nair M, Svedberg P, Larsson I, Nygren JM. A comprehensive overview of barriers and strategies for AI implementation in healthcare: mixed-method design. PLoS One. 2024;19(8):e0305949. [doi: 10.1371/journal.pone.0305949] [Medline: 39121051]

63. Schouten B, Schinkel M, Boerman AW, et al. Implementing artificial intelligence in clinical practice: a mixed-method study of barriers and facilitators. J Med Artif Intell. 2022;5:12. [doi: 10.21037/jmai-22-71]

64. The augmentOR™ Portal: Empowering Surgeons with Surgical Insights. Asensus Surgical. 2024. URL: https://www.asensus.com/documents/augmentortm-portal-empowering-surgeons-surgical-insights [Accessed 2026-01-30]

65. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg. Jul 2018;268(1):70-76. [doi: 10.1097/SLA.0000000000002693] [Medline: 29389679]

66. Leeds IL, Rosenblum AJ, Wise PE, et al. Eye of the beholder: risk calculators and barriers to adoption in surgical trainees. Surgery. Nov 2018;164(5):1117-1123. [doi: 10.1016/j.surg.2018.07.002] [Medline: 30149939]

67. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. Cureus. Oct 2023;15(10):e46454. [doi: 10.7759/cureus.46454] [Medline: 37927664]

68. Olaye IM, Seixas AA. The gap between AI and bedside: participatory workshop on the barriers to the integration, translation, and adoption of digital health care and AI startup technology into clinical practice. J Med Internet Res. May 2, 2023;25:e32962. [doi: 10.2196/32962] [Medline: 37129947]

69. EU AI act: first regulation on artificial intelligence. European Parliament. URL: https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence [Accessed 2026-01-30]

70. de Keijzer IN, Vos JJ, Scheeren TWL. Hypotension Prediction Index: from proof-of-concept to proof-of-feasibility. J Clin Monit Comput. Dec 2020;34(6):1135-1138. [doi: 10.1007/s10877-020-00465-3] [Medline: 31974829]

71. Amann J, Vayena E, Ormond KE, Frey D, Madai VI, Blasimme A. Expectations and attitudes towards medical artificial intelligence: a qualitative study in the field of stroke. PLoS One. 2023;18(1):e0279088. [doi: 10.1371/journal.pone.0279088] [Medline: 36630325]

72. Gumbs AA, Frigerio I, Spolverato G, et al. Artificial intelligence surgery: How do we get to autonomous actions in surgery? Sensors (Basel). Aug 17, 2021;21(16):5526. [doi: 10.3390/s21165526] [Medline: 34450976]

73. Lambert SI, Madi M, Sopka S, et al. An integrative review on the acceptance of artificial intelligence among healthcare professionals in hospitals. NPJ Digit Med. Jun 10, 2023;6(1):111. [doi: 10.1038/s41746-023-00852-5] [Medline: 37301946]

74. Khanna NN, Maindarkar MA, Viswanathan V, et al. Economics of artificial intelligence in healthcare: diagnosis vs. treatment. Healthcare (Basel). Dec 9, 2022;10(12):2493. [doi: 10.3390/healthcare10122493] [Medline: 36554017]

75. Shamszare H, Choudhury A. Clinicians' perceptions of artificial intelligence: focus on workload, risk, trust, clinical decision making, and clinical integration. Healthcare (Basel). Aug 16, 2023;11(16):2308. [doi: 10.3390/healthcare11162308] [Medline: 37628506]

76. Zhou K, Gattinger G. The evolving regulatory paradigm of AI in MedTech: a review of perspectives and where we are today. Ther Innov Regul Sci. May 2024;58(3):456-464. [doi: 10.1007/s43441-024-00628-3] [Medline: 38528278]

77. Ranucci M, Barile L, Ambrogi F, Pistuddi V, Surgical and Clinical Outcome Research (SCORE) Group. Discrimination and calibration properties of the hypotension probability indicator during cardiac and vascular surgery. Minerva Anestesiol. Jul 2019;85(7):724-730. [doi: 10.23736/S0375-9393.18.12620-4] [Medline: 30481996]

78. Davies SJ, Vistisen ST, Jian Z, Hatib F, Scheeren TWL. Ability of an arterial waveform analysis–derived Hypotension Prediction Index to predict future hypotensive events in surgical patients. Anesth Analg. Feb 2020;130(2):352-359. [doi: 10.1213/ANE.0000000000004121] [Medline: 30896602]

79. Grundmann CD, Wischermann JM, Fassbender P, Bischoff P, Frey UH. Hemodynamic monitoring with Hypotension Prediction Index versus arterial waveform analysis alone and incidence of perioperative hypotension. Acta Anaesthesiol Scand. Nov 2021;65(10):1404-1412. [doi: 10.1111/aas.13964] [Medline: 34322869]

80. Salmasi V, Maheshwari K, Yang D, et al. Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: a retrospective cohort analysis. Anesthesiology. Jan 2017;126(1):47-65. [doi: 10.1097/ALN.0000000000001432] [Medline: 27792044]

81. Solares GJ, Garcia D, Monge Garcia MI, et al. Real-world outcomes of the hypotension prediction index in the management of intraoperative hypotension during non-cardiac surgery: a retrospective clinical study. J Clin Monit Comput. Feb 2023;37(1):211-220. [doi: 10.1007/s10877-022-00881-7] [Medline: 35653007]

82. Reddy VS, Stout DM, Fletcher R, et al. Advanced artificial intelligence–guided hemodynamic management within cardiac enhanced recovery after surgery pathways: a multi-institution review. JTCVS Open. Dec 2023;16:480-489. [doi: 10.1016/j.xjon.2023.06.023] [Medline: 38204636]

83.   Monge García MI, García-López D, Gayat É, et al. Hypotension Prediction Index software to prevent intraoperative hypotension during major non-cardiac surgery: protocol for a European multicenter prospective observational registry (EU-HYPROTECT). J Clin Med. Sep 23, 2022;11(19):5585. [doi: 10.3390/jcm11195585] [Medline: 36233455]

84.   Pražetina M, Šribar A, Sokolović Jurinjak I, Matošević J, Peršec J. Effect of machine learning–guided haemodynamic optimization on postoperative free flap perfusion in reconstructive maxillofacial surgery: a study protocol. Br J Clin Pharmacol. Mar 2024;90(3):684-690. [doi: 10.1111/bcp.15942] [Medline: 37876305]

85.   Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. Oct 29, 2019;17(1):195. [doi: 10.1186/s12916-019-1426-2] [Medline: 31665002]

Abbreviations

ACS: American College of Surgeons
AI: artificial intelligence
HPI: Hypotension Prediction Index
NSQIP: National Surgical Quality Improvement Program
ORACLE: Outcome Risk Assessment with Computer Learning Enhancement
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RCT: randomized controlled trial
TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis + Artificial Intelligence