

Review

AI Trustworthiness in the Perioperative Period for Patients with Serious Illness: Scoping Narrative Review

Bhagvat J Maheta^{1,2}, MD; Rachel L Ross³, PhD, MPH; Isabella Raspi², BA; Christina Keny^{4,5}, MHA, PhD, RN; Marti N Okech³; Selen Bozkurt⁶, PhD; Karleen F Giannitrapani^{2,3}, MPH, PhD

¹Department of Surgery, Northwestern University, Chicago, IL, United States

²VA Palo Alto Health Care System, Palo Alto, CA, United States

³Department of Primary Care and Population Health, Stanford University, Palo Alto, CA, United States

⁴Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, United States

⁵School of Nursing, University of California, , San Francisco, CA

⁶Department of Biomedical Informatics, Emory University School of Medicine, GA, United States

Corresponding Author:

Karleen F Giannitrapani, MPH, PhD
Department of Primary Care and Population Health
Stanford University
3180 Porter Drive
Palo Alto, CA 94305
United States
Email: karleen@stanford.edu

Abstract

Background: Despite the promising potential of AI in the perioperative context, the rapid pace of development and diverse implementation warrant a thorough review to consolidate existing knowledge, identify gaps, and assess the use of trustworthiness principles in AI integration into the perioperative period for patients with serious illness.

Objective: The purpose of this study was to address deficiencies in the perioperative AI literature by elucidating the extent to which discussions of equity, ethics, and safety are incorporated, thereby establishing a foundation for the development of robust ethical guidelines for the safe and effective integration of AI in health care.

Methods: We searched PubMed, Embase, CENTRAL (Cochrane Central Register of Controlled Trials), and Scopus for studies published from 2010 to July 2024. We included studies that reported patient functional outcomes, occurred in the perioperative period (30 d before and up to 90 d after surgery), incorporated AI integration, and included patients with serious illness (defined as malignancy, advanced organ failure, frailty, dementia or neurodegenerative disease, or stroke). To ensure reliability and minimize bias, 2 independent reviewers screened all studies at the title or abstract and full-text stages; conflicts were resolved through team consensus. The abstraction form was developed iteratively and was tested through pilot abstractions. Any discrepancies identified during data extraction were resolved through discussion and consensus among the reviewers. The ROBINS-I (Risk of Bias Tool in Nonrandomized Studies of Interventions) tool was used to assess quality. Abstraction and risk assessment were conducted through a blinded, independent dual-review process. A narrative review was compiled from the identified studies.

Results: Of the 10,980 papers identified through the database searches, this review yielded 81 papers that met the inclusion criteria. Analysis of AI implementation strategies revealed foundational efforts toward equitable access, with 6 studies providing open-access tools and several more designing models with simple inputs suitable for low-resource settings (17 studies). Seven studies mentioned their commitment to transparency (eg, publishing code) to enhance safety and trust. However, significant ethical deficiencies persist, particularly regarding input data, as only 2 studies explicitly addressed racial or ethnic disparities, and concerns about lack of sample diversity (16 studies) and the omission of socially relevant features (5 studies) were frequently noted as limitations.

Conclusions: Machine learning for predictive analytics and other types of AI tools for surgical outcomes offer significant potential but require adherence to trustworthiness and safety principles to be clinically viable. Future research should prioritize adherence to guidelines for equity, ethics, and safety, conduct prospective studies, incorporate more external validation of

AI models, and facilitate transparent monitoring and reporting of model performance to build clinician and patient trust and encourage broader health care system adoption.

JMIR Perioper Med 2026;9:e89936; doi: [10.2196/89936](https://doi.org/10.2196/89936)

Keywords: AI; surgery; perioperative; systematic review

Introduction

Rationale

In the perioperative period, patients with serious illnesses such as malignancies, organ failure, or frailty face a diminished quality of life, and this can be exacerbated during the perioperative period, when patients face additional morbidity and mortality [1-3]. However, AI offers a transformative approach to address these challenges in the perioperative period [4]. By analyzing complex patient data, AI—specifically through methods like machine learning (ML), natural language processing (NLP), and artificial neural networks (ANNs)—can predict risks, optimize treatment plans, and assist in critical surgical decision-making and risk stratification [5]. This technology not only aids clinicians but also holds potential for providing personalized patient support and education [6]. Although these AI tools are already becoming integral to perioperative care, the integration of generative AI is still in its nascent stages.

Although the potential benefits of AI are significant, its integration into health care, particularly for vulnerable patients in the perioperative period, introduces crucial ethical considerations. The core ethical issues include data privacy—ensuring that sensitive patient information remains secure—and algorithmic bias, which can perpetuate and even amplify existing health disparities if the data used to train the AI are not representative [7]. There are also concerns about overreliance on technology, where clinicians may defer too much to AI recommendations without exercising their own critical judgment. The Department of Health and Human Services (HHS) characterizes AI trustworthiness into 3 categories: ethics, equity, and safety [8]. The American Medical Association's (AMA) 2025 governance framework expands trustworthiness by mandating explainable AI tools whose decision logic clinicians can interpret and communicate to patients more transparently [9,10]. The World Health Organization (WHO) emphasizes that trustworthy AI in health care must protect human autonomy, promote safety and well-being, ensure transparency and explainability, foster accountability, advance inclusiveness and equity, and uphold sustainability across the entire AI lifecycle [11]. This is especially important in the perioperative period for seriously ill patients, where the stakes are high, and decisions may need to be made rapidly [12]. A standard for AI ethics in the perioperative period, when patients are highly vulnerable, does not yet exist. It is imperative to establish a robust and appropriate standard to ensure the well-being and safety of patients and to uphold ethical integrity.

The rapid development and diverse implementation of AI in perioperative care have created a compelling need for a thorough review to consolidate existing knowledge, identify

critical gaps, and assess the effectiveness and limitations of integrating perioperative AI for patients with serious illness. Although a growing body of literature validates the performance of various AI models, a significant gap remains in understanding the practical strategies for integrating these validated models into clinical workflows to tangibly enhance efficiency and patient outcomes [13,14]. Although there has been work conducted on perioperative AI integration, most of it has not focused on serious illness, and there is still more to uncover related to the ethical implementation of AI in these spaces [14]. Furthermore, there is a distinct lack of research dedicated to the crucial ethical and safety considerations of AI, as well as an exploration of how AI-enabled team augmentation is currently incorporated into the care environment. Anchoring the concept of perioperative AI trustworthiness within standard guidelines, such as HHS and WHO, reinforces the notion that surgical AI systems must not only achieve technical excellence but also demonstrate compliance with internationally accepted principles of ethical design, governance, and continuous quality assurance.

Objectives

The purpose of this study is to address these deficiencies by systematically reviewing the literature to better elucidate the extent to which current publications incorporate equity, ethics, and safety discussions. This research also aims to provide a comprehensive understanding of the current landscape of AI in perioperative care, establish a foundation for future research, and contribute to informing the development of robust, ethical guidelines to ensure the safe and effective integration of AI in health care.

Methods

Protocol and Registration

Our review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA) Statement and EQUATOR (Enhancing the Quality and Transparency of Health Research) guidelines [15,16]. We registered our protocol with the PROSPERO (International Prospective Register of Systematic Reviews) database under the registration number CRD42024608387 [17].

Eligibility Criteria

This review included peer-reviewed publications from 2010 to July 2024 that met the following criteria, organized by the population, intervention, comparator, outcomes, timing, and setting (PICOTS) framework (Table 1). *Inclusion criteria* were patients with serious illness over the age of 18, integration of AI in the perioperative period, reporting any patient quality outcomes, and being peer-reviewed and published in English (for feasibility). Studies that solely

focused on radiology or pathology data were excluded from this study, as AI in these fields has been well characterized [18,19]. Studies that only used logistic regression, which is a traditional statistical modeling approach, were also uniformly excluded from this study for feasibility. Serious illness, as operationalized in this research, included the following categories: malignancy, advanced organ failure (eg, end-stage cardiac, pulmonary, renal, or hepatic disease), frailty with dementia or neurodegenerative disease (progressive, incurable conditions with functional decline, such as dementia or Parkinson disease), and stroke with significant residual impairment [1,2]. AI is defined as the field encompassing machine learning and generative AI, enabling systems to learn from data and generate new content, respectively [20-22]. The perioperative period was defined as the period 30 days before and up to 90 days after surgery [23]. Patient quality outcomes were defined as both patient-reported measures of quality of life and surgical factors such as mortality, length of hospital stay, and postoperative complications [24,25]. Conversely, biological intermediaries, such as radiology scans and blood biomarker values, were excluded from the definition of patient quality outcomes [26].

Table 1. Population, intervention, comparator, outcomes, timing, and setting (PICOTS) eligibility criteria.

PICOTS	Eligibility criteria
Population	Inclusion • Patients with serious illness • Adults (18 y and older) Serious illness: • Malignancy ◦ Any type of malignant cancer. Benign cancers will not be included • Advanced organ failure ◦ Must be irreversible (unless patient gets an organ transplant) and chronic (no acute processes such as infection) ◦ Common examples: end-stage heart failure, COPD^a, cirrhosis, end-stage renal disease • Frailty and dementia or neurodegenerative disease ◦ Chronic, irreversible, progressive, and deteriorating conditions with no cure ◦ Multiple chronic conditions, significant functional impairments, and an increased risk of mortality ◦ Common examples: dementia, Parkinson disease • Stroke with significant residual impairment ◦ Major stroke and have experienced significant and persistent neurological deficits
Intervention	Inclusion • Describe the integration of AI in the perioperative period Exclusion • AI is only used for diagnostic radiology
Comparator	• N/A^b
Outcomes	Inclusion • Patient quality outcomes reported in the study ◦ Patient-reported outcomes relating to quality of life ◦ Patient-demonstrated outcomes (eg, arm curl, chair stand, back scratch, chair sit, reach, walk test, overall cognition, memory, etc)
Timing	• Interventions with any follow-up period
Study characteristics	Inclusion • Peer-reviewed, published manuscript • English only (for feasibility)

^aCOPD: chronic obstructive pulmonary disease.

^bN/A: not applicable.

Information Sources and Search

Our comprehensive literature search encompassed terms related to AI, the perioperative period, and serious illness, as detailed in [Multimedia Appendix 1](#). The search was created through an interactive process using prior published search terms, as well as collaboration with experts in AI (CK, SB). We conducted searches across PubMed, Embase, Scopus, and the Cochrane Central Register of Controlled Trials (CENTRAL) for publications from 2010 to July 2024, yielding an initial 10,980 records, which were reduced to 8942 after removing duplicates.

Selection of Sources for Evidence

Our study team advisors included a PhD expert in team science, health care systems design, and health services research (KG), a PhD expert in patient engagement, organizational behavior, and health services research (RR), and a PhD and registered nurse expert in AI and the perioperative period (CK). Screening involved 4 co-authors (BM, RR, IR, and MO), each independently reviewing studies at the title or abstract and full-text stages, with reviewers blinded to each other’s assessments. Discrepancies during title or abstract screening were primarily resolved by a designated “gold

standard” reviewer (BM or KG), with occasional recourse to majority consensus. At the full-text screening stage, exclusion reasons were systematically recorded, and all conflicts were adjudicated solely by the “gold standard” reviewer (BM or KG). If a study did not specify what percentage of their sample size was diagnosed with malignancy compared to benign pathology, the corresponding author was contacted. If the information was not given within 3 weeks, the study was rejected. Three studies were rejected because of this. Furthermore, any systematic reviews fulfilling the inclusion criteria were used to identify additional relevant studies, which were then incorporated into the title or abstract screening phase. The study selection process, including the number of studies at each stage, was documented using Covidence software to generate a PRISMA flow diagram [27].

Data Charting Process

The abstraction form was created through an iterative process (Multimedia Appendix 2). This form collected data on participant characteristics, including demographics such as gender, age, and race or ethnicity, as well as the category of serious illness (malignancy, advanced organ failure, frailty or dementia, or stroke). Information regarding the AI intervention was abstracted, including the AI’s purpose in the study (prognostication, screening, decision support, etc), the type of AI used (generative AI or ML), and the specific AI interface or tool. Information related to the validity of the AI model was not abstracted, since the primary objective of this review was to better understand equity, ethics, and safety discussions within the studies, and the focus was on implementation of the AI model rather than its validity. Additionally, the form captured whether the study addressed AI equity, ethics, and safety, and how AI was used within the health care setting, such as a team member or a prompting tool. Data abstraction for the included studies was conducted using Google Sheets, with 2 reviewers per paper. Reviewers (either BM, RR, IR, or CK) independently abstracted each paper, were blinded to each other’s work, and resolved all abstraction conflicts through a consensus discussion.

Risk of Bias and Quality Assessment

The ROBINS-I tool was used during the data extraction phase to evaluate the quality of each included paper [28], since it allows for risk of bias stratification for nonrandomized studies that examine the effect of an intervention on an outcome. The research team assessed potential biases across the following domains: confounding, selection of participants into the study, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of the reported result. This systematic assessment involved 2 independent reviewers (among BM, RR, IR, and CK), who evaluated the risk of bias across multiple domains. Disagreements between the 2 reviewers were resolved through discussion and consensus.

Synthesis of Results

Given the substantial heterogeneity observed across included studies, stemming from variations in AI implementation, use, and reported patient quality-of-life outcomes, a narrative synthesis of the abstracted data was conducted using a grouping method. We synthesized studies based on malignancy type, AI type, as well as the incorporation of AI equity, ethics, and safety. Studies that included equity, ethics, and/or safety related to AI, as defined in the sections below, were included in this review, and details on these topics were systematically extracted from the manuscripts. The type of AI used was categorized based on whether the individual tools were supervised ML, unsupervised ML, supervised deep learning, unsupervised deep learning, NLP, or linear programming. The definitions used are listed in the sections below. The validity of AI was not considered as part of this analysis. For studies that included AI equity, ethics, and safety, the quotes were extracted and open-coded with dual review until saturation of codes was reached (RR, BM) [29,30]. Then the quotes were consolidated into appropriate themes through open coding and dual review. AI equity, ethics, and safety were defined in the following section.

Data Items

The following definitions were used to guide the identification and extraction of data related to AI equity, ethics, and safety:

- AI equity: The fair and just distribution of the benefits and opportunities afforded by AI technologies across all populations, regardless of race, ethnicity, gender, socioeconomic status, geographic location, or other social determinants of health, while also mitigating any potential harms or biases that could exacerbate existing health disparities [31]
- AI ethics: Ensuring that the development, implementation, and use of AI adhere to the fundamental ethical principles of beneficence, nonmaleficence, autonomy, justice, and explainability, while mitigating potential risks, such as biased screening data, to promote equitable and trustworthy health care outcomes [31]
- AI safety: The development and implementation of measures and practices that ensure that AI systems used in medical settings are reliable and robust and function as intended, minimizing the risk of harm to patients, clinicians, or the health care system [32]

Definitions of AI

The following definitions were used to categorize the types of AI identified in the included studies:

- Generative AI focuses on creating new content, such as text, images, or synthetic patient data. It learns the underlying patterns and structures of the data it is trained on and then uses that knowledge to generate entirely new content that is similar in style and characteristics to the training data [33].
- ML algorithms learn patterns from data without explicit programming, enabling them to make predictions or decisions [34].

- NLP focuses on enabling computers to understand, interpret, and generate human language [35].

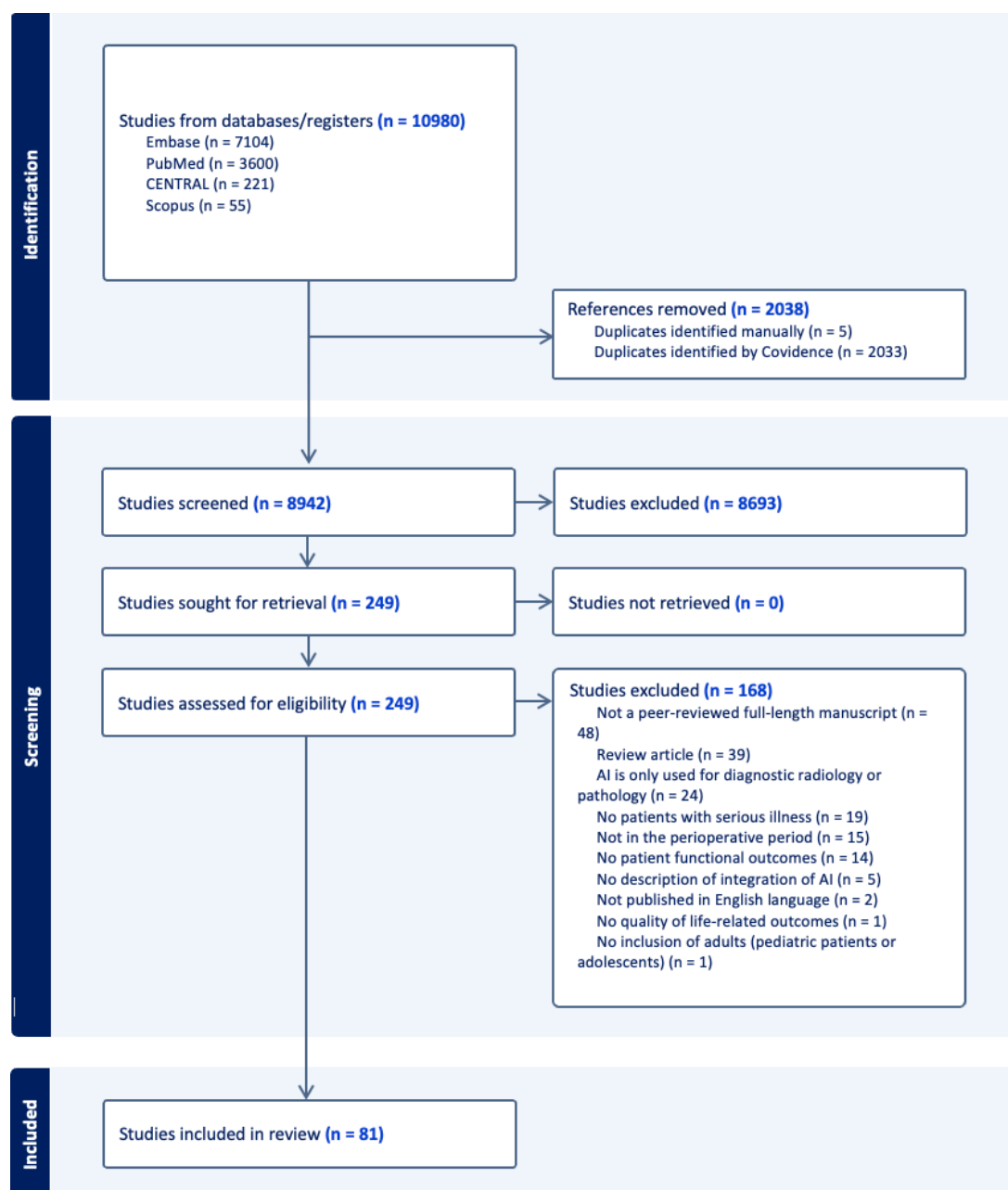
Results

Selection of Sources of Evidence

A total of 10,980 records were considered for review by the research teams. After 2038 (18.6%) papers were removed

as duplicates, 8942 were screened at the title or abstract stage. Of 8942 papers, 8687 (97.1%) of those were excluded, and 249 were screened at the full-text stage. Ultimately, 81 studies were eligible for review, as presented in Figure 1 [36-116]. Overall, the included studies were assessed as having low-to-moderate risk when evaluated with the ROBINS-I tool, suggesting that the overall quality of the included papers was high.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA) flowchart.



Characteristics of Sources of Evidence

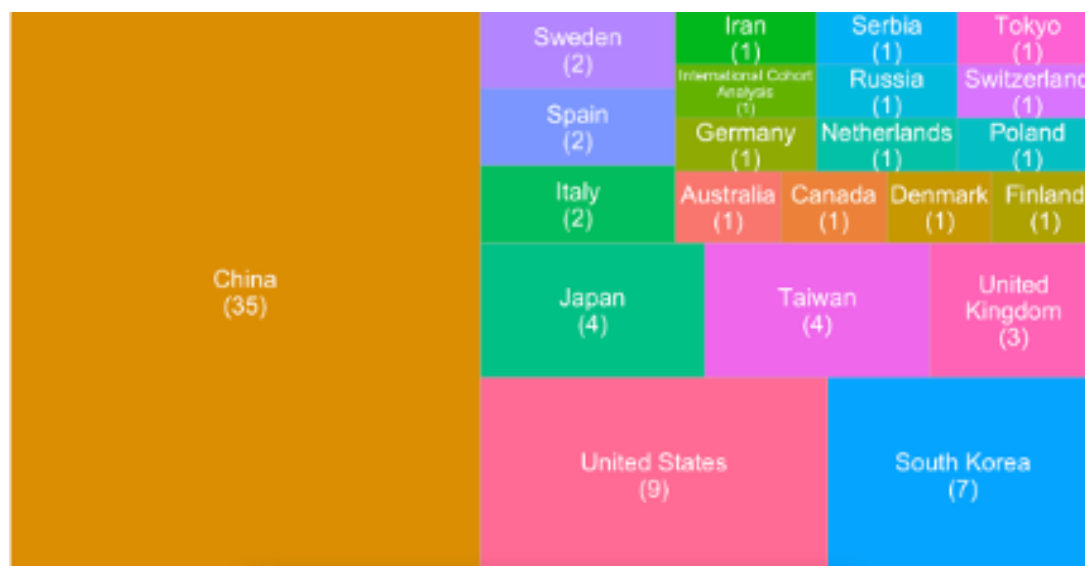
Across all studies, there was an overall total of 714,253 participants. A majority of the studies were published after 2022, with 59 studies published during or after 2022 and 22 studies published prior to 2022. A majority of the studies were published in China (35 studies), with the United States

(9 studies) and South Korea (7 studies) having the next highest numbers of publications, as shown in Figure 2. Of the 81 included studies, 80 (98.8%) focused on patients with malignancy, suggesting a growing interest and importance of AI integration within oncology specifically. The most common forms of cancer in the included papers were colorectal cancer (15 studies), hepatic cancer (11 studies),

and head and neck cancer (9 studies). One study focused on patients with liver failure requiring surgery, without a primary malignancy in all patients [109]. A majority of studies used AI for prognostication (73 studies), with 4 using it for risk scores, 3 for treatment recommendations, and 1 using AI for decision support. All studies used ML as their form of AI and included different models, including support

vector machine (SVM), least absolute shrinkage and selection operator (LASSO) regression, and extreme gradient boosting (XGBoost). Of the included papers, 75 were retrospective studies, and 6 were prospective cohort studies. A comprehensive overview of relevant study characteristics can be found in [Multimedia Appendices 3 and 4](#).

Figure 2. Country representation of included studies by frequency.



Results of Individual Sources of Evidence and Synthesis of Results

AI Equity

Consideration of Model Use in Low-Resource Settings (n=17)

Seven study teams addressed the potential for their models to be deployed in low-resource settings. Several study teams explained that model inputs were intentionally selected for their ease of acquisition in standard patient care settings (as opposed to data points that may require costly equipment to obtain or advanced analytical software to calculate) [46,59,84,91]. Two study teams specifically called out the low-cost nature of their model development [84,109], and one referenced their model's convenience and potential for widespread distribution relative to comparable existing options [103]. One study team originally stated that their ideal model application would be an electronic health record (EHR)-integrated tool, but included a statement about the potential for the creation of a standalone mobile application for use in low-resource settings that may not have robust information technology infrastructure [45]. These discussions reflect an effort to mitigate structural inequities in technological capacity and underscore the importance of context-aware model design for advancing AI equity in diverse care environments.

Distribution of an Open-Access Tool (n=6)

Six of the studies included in our sample described the creation of a web-based tool that would allow clinicians

outside their organizations to implement the model and generate insights for their own patients [41,61,6278, 86, 91], reinforcing a commitment to equitable access to these emerging technologies.

Design Considerations for Non--AI-Expert Clinicians (n=4)

Usability for frontline clinicians, regardless of their expertise in AI model development, was explicitly discussed in 4 studies. These authors described human-centered design choices, such as clear instructions, intuitive interfaces, and simplified input-output mechanisms to facilitate real-world implementation of these tools and to ease communication between patients and their providers [41,84,116]. One study focused on how their tool could be embedded easily into clinical workflows by creating publicly available software that could be incorporated into an EHR for individualized risk prediction [45]. These features reflect a growing recognition that equitable deployment of AI depends not only on tool accuracy but also on ensuring that all clinicians, regardless of AI fluency, can effectively incorporate these models into patient care.

Consideration of the Cost Burden to the Patient Resulting From AI Decision Output (n=1)

One study explicitly addressed the potential downstream cost implications of AI-enabled decision-making for patients: Qiao et al [89] cautioned clinicians to interpret model outputs in a manner that avoids unnecessary or overly burdensome interventions, particularly in cases where the AI-generated

risk assessments might influence decisions around costly diagnostic or treatment options. This finding highlights an emerging but underdeveloped area of AI equity: namely, the affordability of care pathways shaped by algorithmic outputs. Without attention to the financial consequences of model-informed decisions, there is a risk of exacerbating disparities among patients already facing economic hardship.

AI Ethics

Lack of Patient Sample Diversity (n=16)

Many studies were conducted at a single site and/or involved a retrospective study design, inherently introducing the potential for selection bias and limiting generalizability. This was frequently noted via a brief acknowledgment in the limitations sections of the papers included in our review. However, 16 study teams discussed concerns about selection bias and lack of generalizability more robustly. Our findings indicate that authors primarily relied on 3 techniques to reflect on concerns about lack of patient sample diversity: calling out the specific methodological issues leading to these worries, explaining specific techniques used to address limitations in diversity, and identifying the need for further exploration beyond the scope of the existing work to further validate their models and thereby address these ethical considerations.

Six study teams mentioned specific considerations driving their concerns regarding a lack of representativeness of the study population, ranging from a lack of geographic and demographic diversity in the sample [43,77,89], limited sample sizes [44], and lack of external validation in additional populations [66], to the overrepresentation of patient data from highly resourced settings (eg, hospitals with infrastructure to uphold NSQIP (National Surgical Quality Improvement Program) reporting requirements [61] or from exclusively tertiary care settings [68]).

Other authors mentioned intentional steps taken to enhance the diversity of their analytic sample, including the intentional inclusion of patients from multiple institutions [41, 80], strict study recruitment screening guidelines designed to address concerns about population homogeneity [58], and the intentional selection of a patient population that is currently underrepresented in clinical trials [90]. In the manuscript, 1 study team included an explanation of the technique they used to address class-imbalance problems, leading to improved data quality and representativeness of real-world settings [88]. A handful of other studies mentioned concerns about lack of generalizability more generally but also included a call for future work with more diverse populations, inherently implying the need to address ethical considerations of model dissemination [45,50,53,92].

Omission of Socially and Clinically Relevant Features Limiting Model Accuracy (n=5)

Concerns about missing or incomplete data elements were discussed in a small subset of studies, primarily in the context of limitations in model performance. For example, several study teams highlighted the absence of key variables, such

as potentially relevant clinical characteristics, smoking status, insurance status, and other socioeconomic factors, which have the potential to meaningfully impact model performance [47, 57,74].

One study team mentioned that various methodological options are available to impute missing values, which could help avoid the loss of important information [86]. Relatedly, Kadomatsu et al [60] noted that the inclusion of subjective variables (ie, those requiring clinician judgment) may introduce variability or distortion in the dataset, further complicating model accuracy and reliability.

Discussion of Health Disparities (n=2)

Only 2 studies included in our review explicitly discussed racial and ethnic disparities in the context of model development or evaluation, despite the clear relevance of these considerations. Osman et al [86] specifically mentioned differences in mortality rates across racial and ethnic groups that are observed in national trends; they highlighted that their model performed consistently across datasets with differing baseline characteristics, thereby improving faith in the model's robustness. Verma [101] acknowledged that their model's output identified differences in risks across racial and ethnic groups and called out the need to both understand the mechanism of action behind this finding and to prevent the model from exacerbating existing disparities.

AI Safety

Removing the AI Black Box for Transparency (n=7)

AI safety included clarity and transparency about how the algorithm used the data to generate output and provided insight into what is sometimes thought of as the "black box." Certain studies emphasized the importance of sharing how the AI algorithm generated results through what they termed a "white box," since they felt it provides highly interpretable results [55]. This has been echoed by the AMA's emphasis on explainability and clinician interpretability. Others have emphasized transparency by publishing their code publicly so that others can easily see how the AI model was trained to produce results, and so that the results can be reproduced in other settings [78]. Other studies have argued that increased insight into how AI created the results was important for transparency and to assist surgeon decision-making [79].

AI Considers Additional Factors That May Influence Patient Care (n=6)

Although AI can be a powerful tool, it cannot be used in every setting without clinician input, and the final decision regarding patient care should not be made solely by AI. In addition to clinical factors that affect patient outcomes, there are certain socioeconomic factors that can significantly impact a patient's care that are not fully accounted for in current iterations of AI [60]. Multiple studies have emphasized that although there is great benefit to incorporating AI into serious illness perioperative care, the provider should still make the final decision regarding the patient's care [43,65].

Physical Safety From Additional Testing

Harnessing AI algorithms may also pose a risk of physical harm to patients, threatening AI safety. Liu et al [77] showed that, although there were benefits to incorporating AI, patients were exposed to high radiation from positron emission tomography (PET) scans and incurred increased costs without significant benefit.

AI-Enabled Team Augmentation

Despite growing attention toward the consideration of AI as true “team members,” none of the included studies used this lens in the narrative discussion of how the tools were being included in workflows [117]. However, the majority of papers described how the AI tool under consideration could be used to improve perioperative care by prompting a provider, team member, or patient/caregiver to take a tangible action.

There were 2 studies in which the AI model specifically prompted the provider in terms of patient care. Wang et al [104] used their model to provide personalized recommendations and solutions that the physicians could then act on as part of the study. Zhang et al [113] crafted a study design where their AI model assisted in distinguishing gallbladder carcinoma from xanthogranulomatous cholecystitis in the preoperative setting, thus directly impacting surgical decision-making in the perioperative period. Seventy-seven of the other included studies did not explicitly use their AI model to prompt providers, team members, or patients/caregivers, but described how their model could be incorporated in health care settings to prompt others and lead to real-time action contributing positively to patient care in the perioperative setting.

Discussion

Summary of Evidence

Our study systematically examined the extent to which scholarly work on AI in perioperative care for patients with serious illness explicitly incorporates discussions of equity, ethics, and safety. Only low-to-moderate risk studies, per the ROBINS-I risk of bias tool, were included in this review. However, the ultimate goal of these applications is to use AI-enabled team augmentation that contributes positively to clinician workflow, improves decision-making speed, and enhances overall patient safety. Although numerous studies have focused on the statistical and methodological validation of AI tools in this setting, our findings reveal a significant gap: key ethical considerations—particularly concerning vulnerable patient populations—remain largely unaddressed [14]. Specifically, we noted that limitations in sample diversity and generalizability, often framed by authors in purely methodological terms, carry profound ethical implications for justice and beneficence. By risking unequal performance across diverse patient subgroups and obscuring systemic disparities through the omission of socially and clinically relevant features, these limitations directly challenge the principles required for fair and trustworthy decision-making in real-world clinical applications. Although

this study focused primarily on oncology, many core tenets related to the ethics of integrating AI can be extrapolated to other types of serious illnesses, such as advanced organ failure, frailty, dementia, neurodegenerative diseases, and stroke. Therefore, our analysis underscores the critical need for a paradigm shift, moving the conversation beyond technical statistical performance to demand the integration of rigorous, equity-focused analyses into both the validation and the broader implementation planning of AI tools used in the sensitive domain of perioperative patient care by explicitly including equity, ethics, and safety within the implementation of perioperative AI.

The integration of AI into the perioperative care of seriously ill patients presents unique and critical ethical and safety challenges. As a highly vulnerable population, these patients stand to benefit from AI’s ability to analyze complex data for refined decision-making and personalized care, but they are also at increased risk of the potential for algorithmic bias and data privacy breaches [118]. The application of AI must be carefully considered, distinguishing between a patient-facing chatbot and other leveraged forms of AI that may also improve patient care [119]. Ensuring the safe deployment of AI is paramount and necessitates a multifaceted approach, including clearly defined safety parameters, robust risk mitigation strategies, and integrated workflows with essential human oversight [120,121]. Given the transformative potential of AI to fundamentally reshape perioperative care, ensuring robust ethical standards is paramount to safeguard patient safety and maintain trust in these technologies [122]. Currently, a gold standard for addressing the equity, ethics, and safety of AI in the perioperative period is absent from most research articles, a gap that must be addressed in all future manuscripts to align with emerging guidelines such as the US Department of Health and Human Services’ Trustworthy AI Playbook or the AMA’s 8-Step AI Governance Toolkit [8,10,123]. This review helps define the aspects of AI trustworthiness through the definition mentioned in the *Methods* section and encourages all future papers using AI in the perioperative period to explicitly report these criteria to ensure all AI models are held to a rigorous standard for equity, ethics, and safety.

While no existing studies actively use AI as a direct team member, the opportunity to integrate AI in this capacity remains a burgeoning field. The development of trust within human-AI teams is a complex process, and recent research shows that while smaller human-AI teams are initially seen as less trustworthy than human-only groups, this trust deficit diminishes as the team grows, suggesting that team complexity plays a key role in developing trust [124]. AI-powered systems, including chatbots, large language models (LLMs), and agentic AI, are poised to function as effective members of the health care team, either by partnering with providers to deliver clinical decision support or by engaging with patients to facilitate communication, education, and guidance [6,125]. Although minimal literature exists on integrating AI-enabled team augmentation in health care, and none specifically in the context of surgery, AI has the potential to prompt a

team member, whether a clinician or a patient. This capacity to prompt offers significant potential to enhance surgical care for seriously ill patients through improved risk stratification, personalized education, and optimized monitoring. However, its successful implementation hinges on addressing critical concerns related to algorithmic bias, data privacy, and patient trust, while also tackling practical challenges related to integration, overreliance, and ensuring patient comprehension, particularly within this vulnerable population. In addition, trust calibration between humans and AI improves after exposure and iterative feedback, suggesting that increased engagement with AI by health care teams may lead to more AI-enabled team augmentation [126].

Moving beyond traditional machine learning models, generative AI presents both unique opportunities and significant pitfalls in the perioperative period. Generative AI offers the potential to create synthetic patient data for enhanced preoperative risk modeling and simulation, allowing for more personalized surgical planning in complex cases of serious illness [127]. However, this necessitates careful validation to ensure these synthetic datasets accurately reflect real-world patient variability and avoid perpetuating biases [127]. Although generative AI can produce tailored patient education materials and postoperative instructions, improving patient understanding and adherence, there is a risk of generating inaccurate or misleading information, requiring rigorous oversight and quality control to maintain patient safety—a risk not as prominent as with more static ML models [128]. Furthermore, generative AI can streamline administrative tasks, such as generating discharge summaries and patient follow-up plans—a benefit not yet realized in ML—thereby freeing clinicians to focus on direct patient care. Despite these opportunities, critical ethical concerns remain regarding data privacy and the potential for algorithmic bias to impact resource allocation and patient prioritization, underscoring the need for careful and deliberate implementation.

Limitations

This review should be considered with the following limitations. Our broad definitions, used to capture a comprehensive range of AI applications, may have inadvertently led to the exclusion of relevant studies, as the terminology in this constantly evolving field lacks standardization. There were few studies focusing on the workflow integration of AI, which represents a lack of evidence in this space and demonstrates an opportunity to develop this field of research. Due to feasibility constraints, we were unable to conduct a continuous search, and thus, our findings represent a snapshot of a rapidly developing landscape, potentially missing the most recent publications. A significant limitation in the current literature is the scarcity of data on the practical integration of AI models into clinical workflows and a notable absence of discussion regarding ethics, equity, and safety. Furthermore, our study did not analyze the validity or performance of the AI models themselves,

focusing instead on their application and integration. Finally, a significant number of included studies were published in China (35 studies), and additional studies published in languages other than English may have been excluded from this study. This limited geographic and demographic diversity may affect generalizability and inherently create a risk of algorithmic bias. Future research should prioritize the exploration of AI-enabled team augmentation and analyze clinical workflows to better understand where and how AI is being integrated into the perioperative period across different regions of the world and in different health care settings. Since a significant majority of papers were in the realm of oncology, further research on AI integration should focus on populations with other serious illnesses. There should also be longer-term follow-up of AI models to truly understand their impact and safety concerns, along with studies that focus on external validation, prospective implementation, and real-world workflow evaluation so that stronger conclusions can be drawn about clinical utility [129]. It is also crucial to explore the perspectives of patients with serious illnesses and their clinicians regarding the ethical implications of AI in perioperative care, examining their concerns, expectations, and preferred levels of AI integration to ensure that future implementation plans are patient-centered and ethically sound.

Conclusions

This investigation highlights the considerable potential of AI tools for predictive analytics and other AI tools in surgical outcomes while simultaneously exposing a significant deficit in the explicit reporting of ethical considerations, equity, and safety in current publications in the perioperative period. Among the 81 included papers, 80 of which focused on malignancy, there were significant ethical deficiencies, as very few studies mentioned key factors such as racial or ethnic disparities, sample diversity, or socially relevant features. Achieving the clinical viability of these tools requires moving beyond initial statistical reporting to incorporate rigorous external validation, continuous performance monitoring, and fundamental adherence to principles of trustworthiness and safety. The evidence suggests that unless researchers prioritize equity-focused design and data transparency, the risk of perpetuating structural health disparities within this vulnerable perioperative population remains high. To address these critical issues, it would be beneficial for future research to uphold the gold standard of including validation sets, a practice currently lacking in many published studies, and to demand comprehensive data transparency, including information on patient race and ethnicity. By prioritizing prospective studies and transparent reporting, the research community can build the trust necessary for adoption, aligning AI innovation in perioperative care with the broader HHS goals for improving health outcomes through innovative, fair, safe, and ethical technology.

Acknowledgments

This paper was presented at AcademyHealth's 2025 Annual Research Meeting in Minneapolis, MN, in June 2025. This study would not have been possible without the support of the Stanford University School of Medicine and the VA Center for Innovation to Implementation. Generative AI was not used in any portion of the manuscript writing.

Funding

KFG is supported by a VA Career Development Award (19-075).

Data Availability

The data sets analyzed during this study are publicly available as detailed in the Methods section.

Authors' Contributions

Data analysis: BJM, RLR, IR, CK, MNO, SB, KFG

Data collection: BJM, RLR, IR, MNO, KFG

Project ideation: BJM, RLR, CK, KFG

Overall project oversight: KFG

Writing – original draft: BJM, RLR

Writing – review & editing: BJM, RLR, CK, MNO, SB, KFG

All authors read and approved the final version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[[DOCX File \(Microsoft Word File\), 8 KB-Multimedia Appendix 1](#)]

Multimedia Appendix 2

Abstraction form and ROBINS-I tool.

[[DOCX File \(Microsoft Word File\), 4847 KB-Multimedia Appendix 2](#)]

Multimedia Appendix 3

Included study characteristics.

[[DOCX File \(Microsoft Word File\), 21 KB-Multimedia Appendix 3](#)]

Multimedia Appendix 4

Structured synthesis table of included studies.

[[DOCX File \(Microsoft Word File\), 20 KB-Multimedia Appendix 4](#)]

References

1. Sanders JJ, Curtis JR, Tulskey JA. Achieving goal-concordant care: a conceptual model and approach to measuring serious illness communication and its impact. *J Palliat Med*. Mar 2018;21(S2):S17-S27. [doi: [10.1089/jpm.2017.0459](#)] [Medline: [29091522](#)]
2. Murray SA, Kendall M, Boyd K, Sheikh A. Illness trajectories and palliative care. *BMJ*. Apr 30, 2005;330(7498):1007-1011. [doi: [10.1136/bmj.330.7498.1007](#)] [Medline: [15860828](#)]
3. Makary MA, Segev DL, Pronovost PJ, et al. Frailty as a predictor of surgical outcomes in older patients. *J Am Coll Surg*. Jun 2010;210(6):901-908. [doi: [10.1016/j.jamcollsurg.2010.01.028](#)] [Medline: [20510798](#)]
4. 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](#)]
5. Balch JA, Shickel B, Bihorac A, Upchurch GR, Loftus TJ. Integration of AI in surgical decision support: improving clinical judgment. *Global Surg Educ*. 2024;3(1):56. [doi: [10.1007/s44186-024-00257-2](#)]
6. Lee JC, Hamill CS, Shnayder Y, Buczek E, Kakarala K, Bur AM. Exploring the role of artificial intelligence chatbots in preoperative counseling for head and neck cancer surgery. *Laryngoscope*. Jun 2024;134(6):2757-2761. [doi: [10.1002/lary.31243](#)] [Medline: [38126511](#)]
7. Morley J, Machado CCV, Burr C, et al. The ethics of AI in health care: a mapping review. *Soc Sci Med*. Sep 2020;260:113172. [doi: [10.1016/j.socscimed.2020.113172](#)] [Medline: [32702587](#)]
8. Trustworthy AI (TAI) playbook executive summary. U.S. Department of Health & Human Services; 2021. URL: <https://www.hhs.gov/sites/default/files/hhs-trustworthy-ai-playbook-executive-summary.pdf>? [Accessed 2026-09-05]

- JMIR Perioper Med 2026 | vol. 9 | e89936 | p. 11
(page number not for citation purposes)

31. Abràmoff MD, Tarver ME, Loyo-Berrios N, et al. Considerations for addressing bias in artificial intelligence for health equity. *NPJ Digit Med.* Sep 12, 2023;6(1):170. [doi: [10.1038/s41746-023-00913-9](https://doi.org/10.1038/s41746-023-00913-9)] [Medline: [37700029](https://pubmed.ncbi.nlm.nih.gov/37700029/)]
32. Ratwani RM, Bates DW, Classen DC. Patient safety and artificial intelligence in clinical care. *JAMA Health Forum.* Feb 2, 2024;5(2):e235514. [doi: [10.1001/jamahealthforum.2023.5514](https://doi.org/10.1001/jamahealthforum.2023.5514)] [Medline: [38393719](https://pubmed.ncbi.nlm.nih.gov/38393719/)]
33. Oluwagbenro MB. Generative AI: definition, concepts, applications, and future prospects. *Techrxiv.* Preprint posted online on Jun 4, 2024. [doi: [10.36227/techrxiv.171746875.59016695/v1](https://doi.org/10.36227/techrxiv.171746875.59016695/v1)]
34. Badillo S, Banfai B, Birzele F, et al. An introduction to machine learning. *Clin Pharma Therapeutics.* Apr 2020;107(4):871-885. [doi: [10.1002/cpt.1796](https://doi.org/10.1002/cpt.1796)] [Medline: [32128792](https://pubmed.ncbi.nlm.nih.gov/32128792/)]
35. Martinez AR. Natural language processing. *WIREs Computational Stats.* May 2010;2(3):352-357. [doi: [10.1002/wics.76](https://doi.org/10.1002/wics.76)]
36. Amparore D, De Cillis S, Alladio E, et al. Development of machine learning algorithm to predict the risk of incontinence after robot-assisted radical prostatectomy. *J Endourol.* Aug 2024;38(8):871-878. [doi: [10.1089/end.2024.0057](https://doi.org/10.1089/end.2024.0057)] [Medline: [38512711](https://pubmed.ncbi.nlm.nih.gov/38512711/)]
37. Cai LQ, Yang DQ, Wang RJ, Huang H, Shi YX. Establishing and clinically validating a machine learning model for predicting unplanned reoperation risk in colorectal cancer. *World J Gastroenterol.* Jun 21, 2024;30(23):2991-3004. [doi: [10.3748/wjg.v30.i23.2991](https://doi.org/10.3748/wjg.v30.i23.2991)] [Medline: [38946868](https://pubmed.ncbi.nlm.nih.gov/38946868/)]
38. Chen D, Afzal N, Sohn S, et al. Postoperative bleeding risk prediction for patients undergoing colorectal surgery. *Surgery.* Dec 2018;164(6):1209-1216. [doi: [10.1016/j.surg.2018.05.043](https://doi.org/10.1016/j.surg.2018.05.043)] [Medline: [30033185](https://pubmed.ncbi.nlm.nih.gov/30033185/)]
39. Choi N, Kim Z, Song BH, et al. Prediction of risk factors for pharyngo-cutaneous fistula after total laryngectomy using artificial intelligence. *Oral Oncol.* Aug 2021;119:105357. [doi: [10.1016/j.oraloncology.2021.105357](https://doi.org/10.1016/j.oraloncology.2021.105357)] [Medline: [34044316](https://pubmed.ncbi.nlm.nih.gov/34044316/)]
40. Costantino A, Sampieri C, Pace GM, et al. Development of machine learning models for the prediction of long-term feeding tube dependence after oral and oropharyngeal cancer surgery. *Oral Oncol.* Jan 2024;148:106643. [doi: [10.1016/j.oraloncology.2023.106643](https://doi.org/10.1016/j.oraloncology.2023.106643)] [Medline: [38006688](https://pubmed.ncbi.nlm.nih.gov/38006688/)]
41. Cui Y, Shi X, Qin Y, et al. Establishment and validation of an interactive artificial intelligence platform to predict postoperative ambulatory status for patients with metastatic spinal disease: a multicenter analysis. *Int J Surg.* May 1, 2024;110(5):2738-2756. [doi: [10.1097/JS9.0000000000001169](https://doi.org/10.1097/JS9.0000000000001169)] [Medline: [38376838](https://pubmed.ncbi.nlm.nih.gov/38376838/)]
42. Du J, Yang J, Yang Q, Zhang X, Yuan L, Fu B. Comparison of machine learning models to predict the risk of breast cancer-related lymphedema among breast cancer survivors: a cross-sectional study in China. *Front Oncol.* 2024;14:1334082. [doi: [10.3389/fonc.2024.1334082](https://doi.org/10.3389/fonc.2024.1334082)] [Medline: [38410115](https://pubmed.ncbi.nlm.nih.gov/38410115/)]
43. Fourman MS, Siraj L, Duvall J, et al. Can we use artificial intelligence cluster analysis to identify patients with metastatic breast cancer to the spine at highest risk of postoperative adverse events? *World Neurosurg.* Jun 2023;174:e26-e34. [doi: [10.1016/j.wneu.2023.02.064](https://doi.org/10.1016/j.wneu.2023.02.064)] [Medline: [36805503](https://pubmed.ncbi.nlm.nih.gov/36805503/)]
44. Fu Y, Shi W, Zhao J, et al. Prediction of postoperative health-related quality of life among patients with metastatic spinal cord compression secondary to lung cancer. *Front Endocrinol (Lausanne).* 2023;14:1206840. [doi: [10.3389/fendo.2023.1206840](https://doi.org/10.3389/fendo.2023.1206840)] [Medline: [37720536](https://pubmed.ncbi.nlm.nih.gov/37720536/)]
45. Ganguli R, Franklin J, Yu X, Lin A, Heffernan DS. Machine learning methods to predict presence of residual cancer following hysterectomy. *Sci Rep.* Feb 17, 2022;12(1):2738. [doi: [10.1038/s41598-022-06585-x](https://doi.org/10.1038/s41598-022-06585-x)] [Medline: [35177700](https://pubmed.ncbi.nlm.nih.gov/35177700/)]
46. Ganguli R, Franklin J, Yu X, Lin A, Lad R, Heffernan DS. Machine learning models to prognose 30-day mortality in postoperative disseminated cancer patients. *Surg Oncol.* Sep 2022;44:101810. [doi: [10.1016/j.suronc.2022.101810](https://doi.org/10.1016/j.suronc.2022.101810)] [Medline: [36088867](https://pubmed.ncbi.nlm.nih.gov/36088867/)]
47. Ghaith AK, Ghanem M, Zamanian C, et al. Using machine learning to predict 30-day readmission and reoperation following resection of supratentorial high-grade gliomas: an ACS NSQIP study involving 9418 patients. *Neurosurg Focus.* Jun 2023;54(6):E12. [doi: [10.3171/2023.3.FOCUS22652](https://doi.org/10.3171/2023.3.FOCUS22652)] [Medline: [37552633](https://pubmed.ncbi.nlm.nih.gov/37552633/)]
48. Guo SW, Shen J, Gao JH, et al. A preoperative risk model for early recurrence after radical resection may facilitate initial treatment decisions concerning the use of neoadjuvant therapy for patients with pancreatic ductal adenocarcinoma. *Surgery.* Dec 2020;168(6):1003-1014. [doi: [10.1016/j.surg.2020.02.013](https://doi.org/10.1016/j.surg.2020.02.013)] [Medline: [32321665](https://pubmed.ncbi.nlm.nih.gov/32321665/)]
49. Han IW, Cho K, Ryu Y, et al. Risk prediction platform for pancreatic fistula after pancreatoduodenectomy using artificial intelligence. *World J Gastroenterol.* Aug 14, 2020;26(30):4453-4464. [doi: [10.3748/wjg.v26.i30.4453](https://doi.org/10.3748/wjg.v26.i30.4453)] [Medline: [32874057](https://pubmed.ncbi.nlm.nih.gov/32874057/)]
50. He Y, Luo L, Shan R, et al. Development and validation of a nomogram for predicting postoperative early relapse and survival in hepatocellular carcinoma. *J Natl Compr Canc Netw.* Dec 20, 2023;22(1D):e237069. [doi: [10.6004/jnccn.2023.7069](https://doi.org/10.6004/jnccn.2023.7069)] [Medline: [38118280](https://pubmed.ncbi.nlm.nih.gov/38118280/)]
51. Hoek AG, van Oort S, Mukamal KJ, Beulens JWJ. Alcohol consumption and cardiovascular disease risk: placing new data in context. *Curr Atheroscler Rep.* Jan 2022;24(1):51-59. [doi: [10.1007/s11883-022-00992-1](https://doi.org/10.1007/s11883-022-00992-1)] [Medline: [35129737](https://pubmed.ncbi.nlm.nih.gov/35129737/)]

52. Hong Y, Li Y, Ye M, Yan S, Yang W, Jiang C. Identifying an optimal machine learning model generated circulating biomarker to predict chronic postoperative pain in patients undergoing hepatectomy. *Front Surg*. 2023;9:1068321. [doi: [10.3389/fsurg.2022.1068321](https://doi.org/10.3389/fsurg.2022.1068321)] [Medline: [36684250](#)]
53. Huang J, Xie X, Wu H, et al. Development and validation of a combined nomogram model based on deep learning contrast-enhanced ultrasound and clinical factors to predict preoperative aggressiveness in pancreatic neuroendocrine neoplasms. *Eur Radiol*. Nov 2022;32(11):7965-7975. [doi: [10.1007/s00330-022-08703-9](https://doi.org/10.1007/s00330-022-08703-9)] [Medline: [35389050](#)]
54. Huang Y, Chen H, Zeng Y, Liu Z, Ma H, Liu J. Development and validation of a machine learning prognostic model for hepatocellular carcinoma recurrence after surgical resection. *Front Oncol*. 2020;10:593741. [doi: [10.3389/fonc.2020.593741](https://doi.org/10.3389/fonc.2020.593741)] [Medline: [33598425](#)]
55. Ishizaki T, Mazaki J, Enomoto M, et al. Predictive modelling for high-risk stage II colon cancer using auto-artificial intelligence. *Tech Coloproctol*. Mar 2023;27(3):183-188. [doi: [10.1007/s10151-022-02685-y](https://doi.org/10.1007/s10151-022-02685-y)] [Medline: [36031650](#)]
56. Ivanics T, Nelson W, Patel MS, et al. The Toronto postliver transplantation hepatocellular carcinoma recurrence calculator: a machine learning approach. *Liver Transpl*. Apr 2022;28(4):593-602. [doi: [10.1002/lt.26332](https://doi.org/10.1002/lt.26332)] [Medline: [34626159](#)]
57. Jeon Y, Kim YJ, Jeon J, et al. Machine learning based prediction of recurrence after curative resection for rectal cancer. *PLoS One*. 2023;18(12):e0290141. [doi: [10.1371/journal.pone.0290141](https://doi.org/10.1371/journal.pone.0290141)] [Medline: [38100485](#)]
58. Jin F, Liu W, Qiao X, Shi J, Xin R, Jia HQ. Nomogram prediction model of postoperative pneumonia in patients with lung cancer: a retrospective cohort study. *Front Oncol*. 2023;13:1114302. [doi: [10.3389/fonc.2023.1114302](https://doi.org/10.3389/fonc.2023.1114302)] [Medline: [36910602](#)]
59. Jung JO, Crnovrsanin N, Wirsik NM, et al. Machine learning for optimized individual survival prediction in resectable upper gastrointestinal cancer. *J Cancer Res Clin Oncol*. May 2023;149(5):1691-1702. [doi: [10.1007/s00432-022-04063-5](https://doi.org/10.1007/s00432-022-04063-5)] [Medline: [35616729](#)]
60. Kadomatsu Y, Emoto R, Kubo Y, et al. Development of a machine learning-based risk model for postoperative complications of lung cancer surgery. *Surg Today*. Dec 2024;54(12):1482-1489. [doi: [10.1007/s00595-024-02878-y](https://doi.org/10.1007/s00595-024-02878-y)] [Medline: [38896280](#)]
61. Karabacak M, Margetis K. A machine learning-based online prediction tool for predicting short-term postoperative outcomes following spinal tumor resections. *Cancers (Basel)*. Jan 28, 2023;15(3):812. [doi: [10.3390/cancers15030812](https://doi.org/10.3390/cancers15030812)] [Medline: [36765771](#)]
62. Karhade AV, Thio QCBS, Ogink PT, et al. Development of machine learning algorithms for prediction of 30-day mortality after surgery for spinal metastasis. *Neurosurgery*. Jul 1, 2019;85(1):E83-E91. [doi: [10.1093/neuros/nyy469](https://doi.org/10.1093/neuros/nyy469)] [Medline: [30476188](#)]
63. Kim K, Han Y, Jeong S, et al. Prediction of postoperative length of hospital stay based on differences in nursing narratives in elderly patients with epithelial ovarian cancer. *Methods Inf Med*. Dec 2019;58(6):222-228. [doi: [10.1055/s-0040-1705122](https://doi.org/10.1055/s-0040-1705122)] [Medline: [32349156](#)]
64. Kinoshita F, Takenaka T, Yamashita T, et al. Development of artificial intelligence prognostic model for surgically resected non-small cell lung cancer. *Sci Rep*. Sep 21, 2023;13(1):15683. [doi: [10.1038/s41598-023-42964-8](https://doi.org/10.1038/s41598-023-42964-8)] [Medline: [37735585](#)]
65. Klén R, Salminen AP, Mahmoudian M, Syvänen KT, Elo LL, Boström PJ. Prediction of complication related death after radical cystectomy for bladder cancer with machine learning methodology. *Scand J Urol*. Oct 2019;53(5):325-331. [doi: [10.1080/21681805.2019.1665579](https://doi.org/10.1080/21681805.2019.1665579)] [Medline: [31552774](#)]
66. Kuo PJ, Wu SC, Chien PC, et al. Artificial neural network approach to predict surgical site infection after free-flap reconstruction in patients receiving surgery for head and neck cancer. *Oncotarget*. 2018;9(17):13768-13782. [doi: [10.18632/oncotarget.24468](https://doi.org/10.18632/oncotarget.24468)] [Medline: [29568393](#)]
67. Kuwayama N, Hoshino I, Mori Y, Yokota H, Iwatate Y, Uno T. Applying artificial intelligence using routine clinical data for preoperative diagnosis and prognosis evaluation of gastric cancer. *Oncol Lett*. 2023;26(5):499. [doi: [10.3892/ol.2023.14087](https://doi.org/10.3892/ol.2023.14087)] [Medline: [37854867](#)]
68. Laios A, De Freitas DLD, Saalmink G, et al. Stratification of length of stay prediction following surgical cytoreduction in advanced high-grade serous ovarian cancer patients using artificial intelligence; the Leeds L-AI-OS score. *Curr Oncol*. Nov 23, 2022;29(12):9088-9104. [doi: [10.3390/curroncol29120711](https://doi.org/10.3390/curroncol29120711)] [Medline: [36547125](#)]
69. Laios A, De Oliveira Silva RV, Dantas De Freitas DL, et al. Machine learning-based risk prediction of critical care unit admission for advanced stage high grade serous ovarian cancer patients undergoing cytoreductive surgery: the Leeds-Natal score. *J Clin Med*. Dec 24, 2021;11(1):87. [doi: [10.3390/jcm11010087](https://doi.org/10.3390/jcm11010087)] [Medline: [35011828](#)]
70. Lee IC, Huang JY, Chen TC, et al. Evolutionary learning-derived clinical-radiomic models for predicting early recurrence of hepatocellular carcinoma after resection. *Liver Cancer*. 2021;10(6):572-582. [doi: [10.1159/000518728](https://doi.org/10.1159/000518728)] [Medline: [34950180](#)]

71. Lee KS, Jang JY, Yu YD, et al. Usefulness of artificial intelligence for predicting recurrence following surgery for pancreatic cancer: retrospective cohort study. *Int J Surg*. Sep 2021;93:106050. [doi: [10.1016/j.jisu.2021.106050](https://doi.org/10.1016/j.jisu.2021.106050)] [Medline: [34388677](#)]
72. Lee W, Park HJ, Lee HJ, et al. Preoperative data-based deep learning model for predicting postoperative survival in pancreatic cancer patients. *Int J Surg*. Sep 2022;105:106851. [doi: [10.1016/j.jisu.2022.106851](https://doi.org/10.1016/j.jisu.2022.106851)] [Medline: [36049618](#)]
73. Li R, Zheng Z, Yang L, et al. Development of a machine learning algorithm to forecast the likelihood of postoperative neurological complications in patients with parotid tumors. *Ear Nose Throat J*. May 28, 2024;1455613241258648. [doi: [10.1177/01455613241258648](https://doi.org/10.1177/01455613241258648)] [Medline: [38804648](#)]
74. Li Y, Wu JH, Li CP, et al. Multidimensional characteristics, prognostic role, and preoperative prediction of peritoneal sarcomatosis in retroperitoneal sarcoma. *Front Oncol*. 2022;12:950418. [doi: [10.3389/fonc.2022.950418](https://doi.org/10.3389/fonc.2022.950418)] [Medline: [36387243](#)]
75. Lin B, Chen F, Wu M, Li C, Lin L. Machine learning models for prediction of postoperative venous thromboembolism in gynecological malignant tumor patients. *J Obstet Gynaecol Res*. Jul 2024;50(7):1175-1181. [doi: [10.1111/jog.15960](https://doi.org/10.1111/jog.15960)] [Medline: [38689519](#)]
76. Lin V, Tsouchnika A, Allakhverdiiev E, et al. Training prediction models for individual risk assessment of postoperative complications after surgery for colorectal cancer. *Tech Coloproctol*. Aug 2022;26(8):665-675. [doi: [10.1007/s10151-022-02624-x](https://doi.org/10.1007/s10151-022-02624-x)] [Medline: [35593971](#)]
77. Liu CT, Peng YH, Hong CQ, et al. A nomogram based on nutrition-related indicators and computed tomography imaging features for predicting preoperative lymph node metastasis in curatively resected esophagogastric junction adenocarcinoma. *Ann Surg Oncol*. Aug 2023;30(8):5185-5194. [doi: [10.1245/s10434-023-13378-7](https://doi.org/10.1245/s10434-023-13378-7)] [Medline: [37010663](#)]
78. Liu R, Wu S, Yu HY, et al. Prediction model for hepatocellular carcinoma recurrence after hepatectomy: machine learning-based development and interpretation study. *Heliyon*. Nov 2023;9(11):e22458. [doi: [10.1016/j.heliyon.2023.e22458](https://doi.org/10.1016/j.heliyon.2023.e22458)] [Medline: [38034691](#)]
79. Lopez-Lopez V, Morise Z, Albaladejo-González M, et al. Explainable artificial intelligence prediction-based model in laparoscopic liver surgery for segments 7 and 8: an international multicenter study. *Surg Endosc*. May 2024;38(5):2411-2422. [doi: [10.1007/s00464-024-10681-6](https://doi.org/10.1007/s00464-024-10681-6)] [Medline: [38315197](#)]
80. Lou SJ, Hou MF, Chang HT, et al. Machine learning algorithms to predict recurrence within 10 years after breast cancer surgery: a prospective cohort study. *Cancers (Basel)*. Dec 17, 2020;12(12):3817. [doi: [10.3390/cancers12123817](https://doi.org/10.3390/cancers12123817)] [Medline: [33348826](#)]
81. Ma Y, Tan B, Wang S, Ren C, Zhang J, Gao Y. Influencing factors and predictive model of postoperative infection in patients with primary hepatic carcinoma. *BMC Gastroenterol*. Apr 12, 2023;23(1):123. [doi: [10.1186/s12876-023-02713-7](https://doi.org/10.1186/s12876-023-02713-7)] [Medline: [37046206](#)]
82. Mai RY, Lu HZ, Bai T, et al. Artificial neural network model for preoperative prediction of severe liver failure after hemihepatectomy in patients with hepatocellular carcinoma. *Surgery*. Oct 2020;168(4):643-652. [doi: [10.1016/j.surg.2020.06.031](https://doi.org/10.1016/j.surg.2020.06.031)] [Medline: [32792098](#)]
83. Masum S, Hopgood A, Stefan S, Flashman K, Khan J. Data analytics and artificial intelligence in predicting length of stay, readmission, and mortality: a population-based study of surgical management of colorectal cancer. *Discov Oncol*. 2022;13(1):11. [doi: [10.1007/s12672-022-00472-7](https://doi.org/10.1007/s12672-022-00472-7)] [Medline: [35226196](#)]
84. Mazaki J, Katsumata K, Ohno Y, et al. A novel prediction model for colon cancer recurrence using auto-artificial intelligence. *Anticancer Res*. Sep 2021;41(9):4629-4636. [doi: [10.21873/anticancer.15276](https://doi.org/10.21873/anticancer.15276)] [Medline: [34475091](#)]
85. Obrzut B, Kusy M, Semczuk A, Obrzut M, Kluska J. Prediction of 10-year overall survival in patients with operable cervical cancer using a probabilistic neural network. *J Cancer*. 2019;10(18):4189-4195. [doi: [10.7150/jca.33945](https://doi.org/10.7150/jca.33945)] [Medline: [31413737](#)]
86. Osman MH, Mohamed RH, Sarhan HM, et al. Machine learning model for predicting postoperative survival of patients with colorectal cancer. *Cancer Res Treat*. Apr 2022;54(2):517-524. [doi: [10.4143/crt.2021.206](https://doi.org/10.4143/crt.2021.206)] [Medline: [34126702](#)]
87. Paro A, Hyer MJ, Tsilimigras DI, et al. Machine learning approach to stratifying prognosis relative to tumor burden after resection of colorectal liver metastases: an international cohort analysis. *J Am Coll Surg*. Apr 1, 2022;234(4):504-513. [doi: [10.1097/XCS.0000000000000094](https://doi.org/10.1097/XCS.0000000000000094)] [Medline: [35290269](#)]
88. Pera M, Gibert J, Gimeno M, et al. Machine learning risk prediction model of 90-day mortality after gastrectomy for cancer. *Ann Surg*. Nov 1, 2022;276(5):776-783. [doi: [10.1097/SLA.0000000000005616](https://doi.org/10.1097/SLA.0000000000005616)] [Medline: [35866643](#)]
89. Qiao W, Wang Y, Luo C, et al. Development of preoperative and postoperative models to predict recurrence in postoperative glioma patients: a longitudinal cohort study. *BMC Cancer*. Feb 28, 2024;24(1):274. [doi: [10.1186/s12885-024-11996-2](https://doi.org/10.1186/s12885-024-11996-2)] [Medline: [38418976](#)]

90. Qin L, Liang Z, Xie J, et al. Development and validation of machine learning models for postoperative venous thromboembolism prediction in colorectal cancer inpatients: a retrospective study. *J Gastrointest Oncol*. Feb 28, 2023;14(1):220-232. [doi: [10.21037/jgo-23-18](https://doi.org/10.21037/jgo-23-18)] [Medline: [36915444](https://pubmed.ncbi.nlm.nih.gov/36915444/)]
91. Sammour T, Cohen L, Karunatilake AI, et al. Validation of an online risk calculator for the prediction of anastomotic leak after colon cancer surgery and preliminary exploration of artificial intelligence-based analytics. *Tech Coloproctol*. Nov 2017;21(11):869-877. [doi: [10.1007/s10151-017-1701-1](https://doi.org/10.1007/s10151-017-1701-1)] [Medline: [29080956](https://pubmed.ncbi.nlm.nih.gov/29080956/)]
92. Shahriarirad R, Meshkati Yazd SM, Fathian R, Fallahi M, Ghadiani Z, Nafissi N. Prediction of sentinel lymph node metastasis in breast cancer patients based on preoperative features: a deep machine learning approach. *Sci Rep*. Jan 16, 2024;14(1):1351. [doi: [10.1038/s41598-024-51244-y](https://doi.org/10.1038/s41598-024-51244-y)] [Medline: [38228684](https://pubmed.ncbi.nlm.nih.gov/38228684/)]
93. Shao S, Liu L, Zhao Y, Mu L, Lu Q, Qin J. Application of machine learning for predicting anastomotic leakage in patients with gastric adenocarcinoma who received total or proximal gastrectomy. *J Pers Med*. Jul 29, 2021;11(8):748. [doi: [10.3390/jpm11080748](https://doi.org/10.3390/jpm11080748)] [Medline: [34442391](https://pubmed.ncbi.nlm.nih.gov/34442391/)]
94. Shen Y, Huang LB, Lu A, Yang T, Chen HN, Wang Z. Prediction of symptomatic anastomotic leak after rectal cancer surgery: a machine learning approach. *J Surg Oncol*. Feb 2024;129(2):264-272. [doi: [10.1002/jso.27470](https://doi.org/10.1002/jso.27470)] [Medline: [37795583](https://pubmed.ncbi.nlm.nih.gov/37795583/)]
95. Shi HY, Tsai JT, Chen YM, Culbertson R, Chang HT, Hou MF. Predicting two-year quality of life after breast cancer surgery using artificial neural network and linear regression models. *Breast Cancer Res Treat*. Aug 2012;135(1):221-229. [doi: [10.1007/s10549-012-2174-6](https://doi.org/10.1007/s10549-012-2174-6)] [Medline: [22836876](https://pubmed.ncbi.nlm.nih.gov/22836876/)]
96. Shi X, Cui Y, Wang S, Pan Y, Wang B, Lei M. Development and validation of a web-based artificial intelligence prediction model to assess massive intraoperative blood loss for metastatic spinal disease using machine learning techniques. *Spine J*. Jan 2024;24(1):146-160. [doi: [10.1016/j.spinee.2023.09.001](https://doi.org/10.1016/j.spinee.2023.09.001)] [Medline: [37704048](https://pubmed.ncbi.nlm.nih.gov/37704048/)]
97. Spelt L, Nilsson J, Andersson R, Andersson B. Artificial neural networks—a method for prediction of survival following liver resection for colorectal cancer metastases. *Eur J Surg Oncol*. Jun 2013;39(6):648-654. [doi: [10.1016/j.ejso.2013.02.024](https://doi.org/10.1016/j.ejso.2013.02.024)] [Medline: [23514791](https://pubmed.ncbi.nlm.nih.gov/23514791/)]
98. Sun LY, Ouyang Q, Cen WJ, Wang F, Tang WT, Shao JY. A model based on artificial intelligence algorithm for monitoring recurrence of HCC after hepatectomy. *Am Surg*. May 2023;89(5):1468-1478. [doi: [10.1177/00031348211063549](https://doi.org/10.1177/00031348211063549)] [Medline: [34894786](https://pubmed.ncbi.nlm.nih.gov/34894786/)]
99. van de Beld JJ, Crull D, Mikhal J, et al. Complication prediction after esophagectomy with machine learning. *Diagnostics (Basel)*. Feb 17, 2024;14(4):439. [doi: [10.3390/diagnostics14040439](https://doi.org/10.3390/diagnostics14040439)] [Medline: [38396478](https://pubmed.ncbi.nlm.nih.gov/38396478/)]
100. van Niftrik CHB, van der Wouden F, Staartjes VE, et al. Machine learning algorithm identifies patients at high risk for early complications after intracranial tumor surgery: registry-based cohort study. *Neurosurgery*. Oct 1, 2019;85(4):E756-E764. [doi: [10.1093/neuros/nyz145](https://doi.org/10.1093/neuros/nyz145)] [Medline: [31149726](https://pubmed.ncbi.nlm.nih.gov/31149726/)]
101. Verma A, Balian J, Hadaya J, et al. Machine learning-based prediction of postoperative pancreatic fistula following pancreaticoduodenectomy. *Ann Surg*. Aug 1, 2024;280(2):325-331. [doi: [10.1097/SLA.0000000000006123](https://doi.org/10.1097/SLA.0000000000006123)] [Medline: [37947154](https://pubmed.ncbi.nlm.nih.gov/37947154/)]
102. Vesovic R, Milosavljevic M, Punt M, et al. The role of the diaphragm in prediction of respiratory function in the immediate postoperative period in lung cancer patients using a machine learning model. *World J Surg Oncol*. Dec 22, 2023;21(1):393. [doi: [10.1186/s12957-023-03278-1](https://doi.org/10.1186/s12957-023-03278-1)] [Medline: [38135875](https://pubmed.ncbi.nlm.nih.gov/38135875/)]
103. Wang K, Tang Y, Zhang F, Guo X, Gao L. Combined application of inflammation-related biomarkers to predict postoperative complications of rectal cancer patients: a retrospective study by machine learning analysis. *Langenbecks Arch Surg*. Oct 13, 2023;408(1):400. [doi: [10.1007/s00423-023-03127-5](https://doi.org/10.1007/s00423-023-03127-5)] [Medline: [37831218](https://pubmed.ncbi.nlm.nih.gov/37831218/)]
104. Wang L, Song D, Wang W, et al. Data-driven assisted decision making for surgical procedure of hepatocellular carcinoma resection and prognostic prediction: development and validation of machine learning models. *Cancers (Basel)*. 2023;15(6):1784. [doi: [10.3390/cancers15061784](https://doi.org/10.3390/cancers15061784)] [Medline: [36980670](https://pubmed.ncbi.nlm.nih.gov/36980670/)]
105. Wen R, Zheng K, Zhang Q, et al. Machine learning-based random forest predicts anastomotic leakage after anterior resection for rectal cancer. *J Gastrointest Oncol*. Jun 2021;12(3):921-932. [doi: [10.21037/jgo-20-436](https://doi.org/10.21037/jgo-20-436)] [Medline: [34295545](https://pubmed.ncbi.nlm.nih.gov/34295545/)]
106. Wu X, Guan Q, Cheng ASK, et al. Comparison of machine learning models for predicting the risk of breast cancer-related lymphedema in Chinese women. *Asia Pac J Oncol Nurs*. 2022;9(12):100101. [doi: [10.1016/j.apjon.2022.100101](https://doi.org/10.1016/j.apjon.2022.100101)] [Medline: [36276882](https://pubmed.ncbi.nlm.nih.gov/36276882/)]
107. Xu Z, Xie Y, Wu L, et al. Using machine learning methods to assess lymphovascular invasion and survival in breast cancer: performance of combining preoperative clinical and MRI characteristics. *J Magn Reson Imaging*. Nov 2023;58(5):1580-1589. [doi: [10.1002/jmri.28647](https://doi.org/10.1002/jmri.28647)] [Medline: [36797654](https://pubmed.ncbi.nlm.nih.gov/36797654/)]
108. Ying T, Borrelli P, Edenbrandt L, et al. Automated artificial intelligence-based analysis of skeletal muscle volume predicts overall survival after cystectomy for urinary bladder cancer. *Eur Radiol Exp*. Nov 19, 2021;5(1):50. [doi: [10.1186/s41747-021-00248-8](https://doi.org/10.1186/s41747-021-00248-8)] [Medline: [34796422](https://pubmed.ncbi.nlm.nih.gov/34796422/)]

109. Zaver HB, Mzaik O, Thomas J, et al. Utility of an artificial intelligence enabled electrocardiogram for risk assessment in liver transplant candidates. *Dig Dis Sci*. Jun 2023;68(6):2379-2388. [doi: [10.1007/s10620-023-07928-y](https://doi.org/10.1007/s10620-023-07928-y)] [Medline: [37022601](https://pubmed.ncbi.nlm.nih.gov/37022601/)]
110. Zeng J, Song D, Li K, Cao F, Zheng Y. Deep learning model for predicting postoperative survival of patients with gastric cancer. *Front Oncol*. 2024;14:1329983. [doi: [10.3389/fonc.2024.1329983](https://doi.org/10.3389/fonc.2024.1329983)] [Medline: [38628668](https://pubmed.ncbi.nlm.nih.gov/38628668/)]
111. Zeng L, Liu L, Chen D, et al. The innovative model based on artificial intelligence algorithms to predict recurrence risk of patients with postoperative breast cancer. *Front Oncol*. 2023;13:1117420. [doi: [10.3389/fonc.2023.1117420](https://doi.org/10.3389/fonc.2023.1117420)] [Medline: [36959794](https://pubmed.ncbi.nlm.nih.gov/36959794/)]
112. Zhang G, Liu X, Hu Y, et al. Development and comparison of machine-learning models for predicting prolonged postoperative length of stay in lung cancer patients following video-assisted thoracoscopic surgery. *Asia Pac J Oncol Nurs*. 2024;11(6):100493. [doi: [10.1016/j.apjon.2024.100493](https://doi.org/10.1016/j.apjon.2024.100493)] [Medline: [38808011](https://pubmed.ncbi.nlm.nih.gov/38808011/)]
113. Zhang W, Wang Q, Liang K, et al. Deep learning nomogram for preoperative distinction between xanthogranulomatous cholecystitis and gallbladder carcinoma: a novel approach for surgical decision. *Comput Biol Med*. Jan 2024;168:107786. [doi: [10.1016/j.compbiomed.2023.107786](https://doi.org/10.1016/j.compbiomed.2023.107786)] [Medline: [38048662](https://pubmed.ncbi.nlm.nih.gov/38048662/)]
114. Zhang Y, Zhou Q, Chen G, Xue S. Early postoperative prediction of the risk of distant metastases in medullary thyroid cancer. *Front Endocrinol (Lausanne)*. 2023;14:1209978. [doi: [10.3389/fendo.2023.1209978](https://doi.org/10.3389/fendo.2023.1209978)] [Medline: [38075078](https://pubmed.ncbi.nlm.nih.gov/38075078/)]
115. Zhao F, Wang P, Yu C, et al. A LASSO-based model to predict central lymph node metastasis in preoperative patients with cN0 papillary thyroid cancer. *Front Oncol*. 2023;13:1034047. [doi: [10.3389/fonc.2023.1034047](https://doi.org/10.3389/fonc.2023.1034047)] [Medline: [36761950](https://pubmed.ncbi.nlm.nih.gov/36761950/)]
116. Zheng CY, Wu J, Chen CS, et al. A scoring model for predicting early recurrence of gastric cancer with normal preoperative tumor markers: a multicenter study. *Eur J Surg Oncol*. Nov 2023;49(11):107094. [doi: [10.1016/j.ejso.2023.107094](https://doi.org/10.1016/j.ejso.2023.107094)] [Medline: [37797381](https://pubmed.ncbi.nlm.nih.gov/37797381/)]
117. Schmutz JB, Outland N, Kerstan S, Georganta E, Ulfert AS. AI-teaming: redefining collaboration in the digital era. *Curr Opin Psychol*. Aug 2024;58:101837. [doi: [10.1016/j.copsyc.2024.101837](https://doi.org/10.1016/j.copsyc.2024.101837)] [Medline: [39024969](https://pubmed.ncbi.nlm.nih.gov/39024969/)]
118. Oh O, Demir G, Ulrich CM. The ethical dimensions of utilizing artificial intelligence in palliative care. *Nurs Ethics*. Jun 2025;32(4):1285-1296. [doi: [10.1177/09697330241296874](https://doi.org/10.1177/09697330241296874)] [Medline: [39551621](https://pubmed.ncbi.nlm.nih.gov/39551621/)]
119. Burry N, Nakagawa S, Blinderman CD. "You are not alone": the allure and limitations of artificial intelligence in serious illness communication. *J Palliat Med*. Jan 2024;27(1):7-9. [doi: [10.1089/jpm.2023.0471](https://doi.org/10.1089/jpm.2023.0471)] [Medline: [38011011](https://pubmed.ncbi.nlm.nih.gov/38011011/)]
120. Sittig DF, Singh H. Recommendations to ensure safety of AI in real-world clinical care. *JAMA*. Feb 11, 2025;333(6):457-458. [doi: [10.1001/jama.2024.24598](https://doi.org/10.1001/jama.2024.24598)] [Medline: [39602298](https://pubmed.ncbi.nlm.nih.gov/39602298/)]
121. Maliha G, Gerke S, Cohen IG, Parikh RB. Artificial intelligence and liability in medicine: balancing safety and innovation. *Milbank Q*. Sep 2021;99(3):629-647. [doi: [10.1111/1468-0009.12504](https://doi.org/10.1111/1468-0009.12504)] [Medline: [33822422](https://pubmed.ncbi.nlm.nih.gov/33822422/)]
122. Guni A, Varma P, Zhang J, Fehervari M, Ashrafi H. Artificial intelligence in surgery: the future is now. *Eur Surg Res*. Jan 22, 2024. [doi: [10.1159/000536393](https://doi.org/10.1159/000536393)] [Medline: [38253041](https://pubmed.ncbi.nlm.nih.gov/38253041/)]
123. Weiner EB, Dankwa-Mullan I, Nelson WA, Hassanpour S. Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice. *PLOS Digit Health*. Apr 2025;4(4):e0000810. [doi: [10.1371/journal.pdig.0000810](https://doi.org/10.1371/journal.pdig.0000810)] [Medline: [40198594](https://pubmed.ncbi.nlm.nih.gov/40198594/)]
124. Georganta E, Ulfert AS. Would you trust an AI team member? Team trust in human-AI teams. *J Occupat Organ Psych*. 2024;97(3):1212-1241. [doi: [10.1111/joop.12504](https://doi.org/10.1111/joop.12504)]
125. Abi-Rafeh J, Henry N, Xu HH, et al. Utility and comparative performance of current artificial intelligence large language models as postoperative medical support chatbots in aesthetic surgery. *Aesthet Surg J*. Jul 15, 2024;44(8):889-896. [doi: [10.1093/asj/sjae025](https://doi.org/10.1093/asj/sjae025)] [Medline: [38318684](https://pubmed.ncbi.nlm.nih.gov/38318684/)]
126. Almutairi M. Teaming in the AI era: AI-augmented frameworks for forming, simulating, and optimizing human teams. *Proc ACM Conf User Model Adapt Pers*. Jun 16, 2025:414-418. [doi: [10.1145/3699682.3727574](https://doi.org/10.1145/3699682.3727574)]
127. Rodler S, Ganjavi C, De Backer P, et al. Generative artificial intelligence in surgery. *Surgery*. Jun 2024;175(6):1496-1502. [doi: [10.1016/j.surg.2024.02.019](https://doi.org/10.1016/j.surg.2024.02.019)] [Medline: [38582732](https://pubmed.ncbi.nlm.nih.gov/38582732/)]
128. Qin S, Chislett B, Ischia J, et al. ChatGPT and generative AI in urology and surgery—a narrative review. *BJUI Compass*. Sep 2024;5(9):813-821. [doi: [10.1002/bco2.390](https://doi.org/10.1002/bco2.390)] [Medline: [39323919](https://pubmed.ncbi.nlm.nih.gov/39323919/)]
129. Jacob C, Brasier N, Laurenzi E, et al. AI for IMPACTS framework for evaluating the long-term real-world IMPACTS of AI-powered clinician tools: systematic review and narrative synthesis. *J Med Internet Res*. Feb 5, 2025;27:e67485. [doi: [10.2196/67485](https://doi.org/10.2196/67485)] [Medline: [39909417](https://pubmed.ncbi.nlm.nih.gov/39909417/)]

Abbreviations

AMA: American Medical Association
ANN: artificial neural network
CENTRAL: Cochrane Central Register of Controlled Trials
EHR: electronic health record
HHS: Health and Human Services
LASSO: least absolute shrinkage and selection operator

LLM: large language model

ML: machine learning

NLP: natural language processing

PET: positron emission tomography

PICOTS: population, intervention, comparator, outcomes, timing, and setting

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols

PROSPERO: International Prospective Register of Systematic Reviews

ROBINS-I: risk of bias tool in nonrandomized studies of interventions

SVM: support vector machine

WHO: World Health Organization

XGBoost: extreme gradient boosting

Edited by Nidhi Rohatgi; peer-reviewed by Huasheng Lv, Lili Zhou, Miloud Chakit, Zhao Liu; submitted 20.Dec.2025; final revised version received 05.Jul.2026; accepted 21.Aug.2026; published 01.Oct.2026

Please cite as:

Maheta BJ, Ross RL, Raspi I, Keny C, Okech MN, Bozkurt S, Giannitrapani KF

AI Trustworthiness in the Perioperative Period for Patients with Serious Illness: Scoping Narrative Review

JMIR Perioper Med 2026;9:e89936

URL: <https://periop.jmir.org/2026/1/e89936>

doi: [10.2196/89936](https://doi.org/10.2196/89936)

© Bhagvat J Maheta, Rachel L Ross, Isabella Raspi, Christina Keny, Marti N Okech, Selen Bozkurt, Karleen F Giannitrapani. Originally published in JMIR Perioperative Medicine (<http://periop.jmir.org>), 01.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Perioperative Medicine, is properly cited. The complete bibliographic information, a link to the original publication on <http://periop.jmir.org>, as well as this copyright and license information must be included.