
**International Journal of Innovative Technologies in Social Science**

e-ISSN: 2544-9435

**Operating Publisher**
**SciFormat Publishing Inc.**
ISNI: 0000 0005 1449 8214

2734 17 Avenue SW,
Calgary, Alberta, T3E0A7,
Canada
+15878858911
editorial-office@sciformat.ca

| | |
|---|---|
| **ARTICLE TITLE** | THE SURGEON AND THE ALGORITHM: TRUST, AUTONOMY, AND RESPONSIBILITY IN AI-DRIVEN SURGICAL DECISION-MAKING — A SYSTEMATIC LITERATURE REVIEW |
| **DOI** | https://doi.org/10.31435/ijitss.2(50).2026.5510 |
| **RECEIVED** | 04 March 2026 |
| **ACCEPTED** | 16 June 2026 |
| **PUBLISHED** | 30 June 2026 |
| **LICENSE** | The article is licensed under a **Creative Commons Attribution 4.0 International License.** |

# THE SURGEON AND THE ALGORITHM: TRUST, AUTONOMY, AND RESPONSIBILITY IN AI-DRIVEN SURGICAL DECISION-MAKING — A SYSTEMATIC LITERATURE REVIEW

***Jakub Marzec*** *(Corresponding Author, Email: kubamarzec7@gmail.com)*
*University Clinical Hospital No. 1 of Pomeranian Medical University in Szczecin, Poland*
*ORCID ID: 0009-0007-8336-7260*

***Filip Kopacki***
*University Clinical Hospital No. 2, Pomeranian Medical University in Szczecin, Poland*
*ORCID ID: 0009-0005-7897-8222*

***Daniel Jaskot***
*Independent Public Provincial Integrated Hospital in Szczecin, Szczecin, Poland*
*ORCID ID: 0009-0004-8678-9838*

***Weronika Szczeblewska***
*Independent Public Provincial Integrated Hospital in Szczecin, Szczecin, Poland*
*ORCID ID: 0009-0007-3956-622X*

***Jakub Świech***
*University Clinical Hospital No. 2, Pomeranian Medical University in Szczecin, Poland*
*ORCID ID: 0009-0001-1732-4162*

***Marta Kamrowska***
*University Clinical Hospital No. 2, Pomeranian Medical University in Szczecin, Poland*
*ORCID ID: 0009-0001-1570-0144*

***Katarzyna Apanasewicz***
*J. Strus Multidisciplinary Municipal Hospital in Poznań, Poland*
*ORCID ID: 0009-0001-0765-5398*

***Natalia Kita***
*University Clinical Hospital No. 2, Pomeranian Medical University in Szczecin, Poland*
*ORCID ID: 0009-0001-2277-6198*

***Natalia Sztenc***
*Pomeranian Medical University, Szczecin, West Pomerania, Poland*
*ORCID ID: 0009-0006-4307-7958*

***Patrycja Broniszewska***
*University Clinical Hospital No. 2, Pomeranian Medical University in Szczecin, Poland*
*ORCID ID: 0009-0000-6577-307X*

**ABSTRACT**

The rapid integration of artificial intelligence (AI) into surgical practice is transforming clinical decision-making processes, raising critical questions regarding trust, autonomy, and responsibility in the operating room. This systematic literature review aims to synthesize current research on AI-driven surgical decision-making, with a particular focus on human–AI interaction and its ethical, cognitive, and legal implications.

The review was conducted in accordance with PRISMA guidelines. A comprehensive search of Scopus, Web of Science, PubMed, and Google Scholar identified a total of 312 records published between 2015 and 2025. After the removal of duplicates and screening based on predefined inclusion and exclusion criteria, 47 studies were included in the final analysis. These studies were categorized into three main thematic domains: trust in AI systems, autonomy and human oversight, and responsibility and accountability.

The findings indicate that AI technologies significantly enhance predictive accuracy, intraoperative support, and clinical decision-making efficiency (Gao et al., 2021; Shao et al., 2023). However, their increasing integration introduces challenges related to automation bias, reduced clinical autonomy, and ambiguity in legal responsibility ((O'Sullivan et al., 2019; Morris et al., 2023)). Trust in AI emerges as a multidimensional construct influenced by transparency, explainability, and user experience (Lopes et al., 2025).

The review highlights the need for improved explainability, standardized validation frameworks, and clearly defined accountability structures. It also identifies key gaps in the literature, particularly the lack of empirical studies on real-time human–AI decision-making and long-term impacts on surgical training. Future research should adopt interdisciplinary approaches to ensure the safe, ethical, and effective integration of AI into surgical practice.

**KEYWORDS**

Artificial Intelligence; Surgical Decision-Making; Trust in AI; Medical Ethics; Clinical Autonomy; Human–AI Interaction; Surgical Robotics

**CITATION**

Jakub Marzec, Filip Kopacki, Daniel Jaskot, Weronika Szczeblewska, Jakub Świech, Marta Kamrowska, Katarzyna Apanasewicz, Natalia Kita, Natalia Sztenc, Patrycja Broniszewska. (2026) The Surgeon and the Algorithm: Trust, Autonomy, and Responsibility in AI-Driven Surgical Decision-Making — a Systematic Literature Review. *International Journal of Innovative Technologies in Social Science*. 2(50). doi: 10.31435/ijitss.2(50).2026.5510



## 1. Introduction

The integration of artificial intelligence (AI) into healthcare systems has emerged as one of the most significant technological developments in contemporary medicine. In recent years, advances in machine learning, predictive analytics, and data-driven decision-support systems have increasingly influenced clinical practice, particularly in high-stakes environments such as surgery (Cobianchi et al., 2022; Morris et al., 2023). Surgical decision-making, traditionally grounded in clinical expertise, experiential knowledge, and situational awareness, is now being augmented—and in some cases challenged—by AI-driven technologies.

AI applications in surgery span multiple stages of care, including preoperative planning, intraoperative guidance, and postoperative outcome prediction (Gao et al., 2021; Hsu et al., 2023). These systems are designed to enhance precision, reduce variability, and support evidence-based decision-making. For example, machine learning models have demonstrated the ability to predict surgical complications and mortality with increasing accuracy, thereby supporting risk stratification and clinical planning (Shao et al., 2023). However, the growing reliance on AI systems introduces new challenges that extend beyond technical performance and into the domains of cognition, ethics, and governance.

One of the most critical issues associated with AI integration in surgery is **trust**. Trust in AI systems significantly influences their adoption and effective use in clinical practice (Brandenburg et al., 2025). Surgeons must rely on algorithmic recommendations while simultaneously maintaining clinical judgment and responsibility for patient outcomes. The opacity of many AI models, often described as the "black box" problem, can undermine confidence and complicate decision-making processes (Lopes et al., 2025). At the same time, excessive reliance on AI may lead to automation bias, in which clinicians accept algorithmic outputs without sufficient critical evaluation.

Another key issue concerns the concept of **autonomy**. The increasing involvement of AI in clinical decisions challenges the traditional role of the surgeon as the primary decision-maker. While AI systems are typically intended to function as supportive tools, their growing influence raises concerns about the potential erosion of professional autonomy (St John et al., 2024). Moreover, AI may also affect **patient autonomy**, particularly in the context of informed consent and shared decision-making, where the role of algorithmic recommendations may not be fully transparent to patients (Teasdale et al., 2024).

Equally important is the issue of **responsibility and accountability**. In situations where AI contributes to clinical decision-making, determining responsibility for adverse outcomes becomes significantly more complex. Existing medico-legal frameworks remain largely centered on human actors and are not fully adapted to the realities of AI-assisted decision-making (O'Sullivan et al., 2019; Morris et al., 2023). This creates uncertainty regarding liability, which may act as a barrier to the adoption of AI technologies in surgical practice (O'Sullivan et al., 2019; Morris et al., 2023).

Given these emerging challenges, there is a clear need for a comprehensive synthesis of existing research that examines the intersection of AI, surgical decision-making, and the associated ethical and cognitive dimensions. Although previous studies have explored specific aspects of AI in surgery, there remains a lack of integrative analyses that address trust, autonomy, and responsibility within a unified framework.

The aim of this article is therefore to provide a systematic review of the current literature on AI-driven surgical decision-making, focusing on the interrelated dimensions of trust, autonomy, and responsibility. By identifying key trends, challenges, and gaps in the literature, this study seeks to contribute to a deeper understanding of human–AI interaction in surgical contexts and to inform future research, clinical practice, and policy development. This review uniquely integrates ethical, cognitive, and legal dimensions into a unified framework of AI-driven surgical decision-making.

## 2. Methodology

This study was conducted as a **systematic literature review** aimed at synthesizing current research on artificial intelligence (AI) in surgical decision-making, with a particular focus on trust, autonomy, and responsibility. The review protocol was developed in accordance with the **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)** guidelines to ensure transparency, reproducibility, and methodological rigor.

### 2.1. Data Sources and Search Strategy

A comprehensive literature search was performed across four major academic databases:

- **Scopus**
- **Web of Science**
- **PubMed/MEDLINE**
- **Google Scholar** (for supplementary search and citation tracking)

The search covered studies published between **January 2015 and December 2025**, reflecting the period of rapid development and implementation of AI technologies in surgical practice.

The search strategy combined controlled vocabulary and free-text terms related to artificial intelligence, surgery, and decision-making. The primary search string used was:

("artificial intelligence" OR "machine learning" OR "deep learning" OR "explainable AI" OR "clinical decision support")

AND (surgery OR surgical OR surgeon OR "operative care")

AND ("decision-making" OR autonomy OR trust OR responsibility OR accountability OR liability OR ethics)

Additional searches were conducted using specialty-specific keywords (e.g., "robotic surgery," "colorectal surgery," "orthopedic surgery," "emergency surgery") to ensure broad coverage of relevant domains.

### 2.2. Eligibility Criteria

Studies were included if they met the following criteria:

1. **Peer-reviewed journal articles**
2. Focus on **AI applications in surgical contexts** or **AI-assisted surgical decision-making**
3. Address at least one of the core thematic dimensions:
   - trust
   - autonomy
   - responsibility
   - accountability
   - explainability
4. Published in **English**
5. Published between **2015 and 2025**

Studies were excluded if they:

- focused exclusively on non-surgical medical specialties without relevance to surgical decision-making
- were purely technical studies lacking clinical, ethical, or decision-making implications
- consisted solely of editorials or opinion pieces without analytical contribution
- were duplicate records or lacked full-text availability

### 2.3. Study Selection Process

The study selection process followed the PRISMA framework and consisted of four stages:

1. **Identification**:
2. A total of **312 records** were identified through database searching and additional sources.
3. **Screening**:
4. After removing **68 duplicate records**, **244 studies** remained for title and abstract screening.
5. **Eligibility**:
6. Following screening, **97 full-text articles** were assessed for eligibility based on the inclusion criteria.
7. **Inclusion**:
8. A final set of **47 studies** was included in the qualitative synthesis.

This process ensured a structured and transparent selection of studies relevant to the research objectives.

### 2.4. Data Extraction

Data were systematically extracted from each included study using a standardized framework. The following information was collected:

- author(s) and year of publication
- country or region
- surgical specialty
- study design (e.g., survey, cohort study, review)
- type of AI system or methodology
- clinical or decision-making context
- key findings
- relevance to trust, autonomy, or responsibility

The extracted data were organized into three primary analytical domains:

1. **Trust in AI-driven surgical decision-making**
2. **Autonomy and human oversight**
3. **Responsibility, accountability, and liability**

### 2.5. Data Synthesis

Given the heterogeneity of study designs, methodologies, and outcome measures, a **quantitative meta-analysis was not feasible**. Instead, a **narrative thematic synthesis** approach was adopted.

This approach allowed for the integration of findings across diverse study types, including empirical research, surveys, qualitative studies, and ethical analyses. The synthesis focused on identifying recurring themes, conceptual frameworks, and areas of convergence and divergence within the literature.

### 2.6. Quality Considerations

Although a formal risk-of-bias assessment was not conducted due to the diversity of study designs, only **peer-reviewed articles** were included to ensure a minimum level of scientific quality. Additionally, preference was given to studies published in high-impact journals and those with clear methodological descriptions.

### 2.7. Methodological Limitations

This review has several limitations. First, the inclusion of only English-language studies may have introduced language bias. Second, the reliance on selected databases may have resulted in the omission of relevant studies published elsewhere. Third, the heterogeneity of included studies limits direct comparability of findings.

Despite these limitations, the use of a structured PRISMA-based approach and clearly defined inclusion criteria enhances the reliability and transparency of the review.

## 3. Literature Review

### 3.1. Definitions and Foundations (Revised Version)

Artificial intelligence (AI) in healthcare broadly refers to computational systems capable of performing tasks that typically require human cognitive functions, including pattern recognition, prediction, classification, and decision support (Cobianchi et al., 2022; Morris et al., 2023). In clinical contexts, AI is most commonly associated with machine learning, deep learning, natural language processing, and computer vision, all of which are increasingly integrated into diagnostic and therapeutic processes. In surgery, these technologies are particularly relevant due to the complexity, uncertainty, and high-risk nature of clinical decision-making (Gao et al., 2021).

In surgical practice, AI should be understood not as a singular technology but as a **multifaceted ecosystem of tools** supporting different phases of perioperative care. These include systems for preoperative risk stratification, intraoperative guidance, robotic assistance, and postoperative outcome prediction (Hsu et al., 2023; Kinross et al., 2025). The integration of these tools into surgical workflows reflects a broader transition toward data-driven and algorithm-supported medicine.

A central concept in this review is **AI-driven surgical decision-making**, defined as a process in which algorithmic systems contribute to, inform, or influence clinical decisions in surgical care. Importantly, current AI applications are predominantly **decision-support systems** rather than fully autonomous agents (Shao et al., 2023). However, even as supportive tools, these systems can significantly shape decision-making processes by prioritizing certain variables, generating probabilistic predictions, and transforming complex clinical data into simplified outputs.

Traditionally, surgical decision-making has been grounded in three core elements: clinical knowledge, technical expertise, and professional judgment. Surgeons rely not only on formal evidence but also on tacit knowledge developed through experience and situational awareness. The introduction of AI introduces a fourth dimension—**algorithmic inference**—which differs fundamentally from human reasoning. Unlike human judgment, which is often transparent and explainable, AI outputs are frequently derived from complex statistical models that may not be easily interpretable (Lopes et al., 2025). This shift has important implications for how decisions are understood, justified, and evaluated in clinical practice.

A key conceptual distinction in the literature is between **assistance, augmentation, and autonomy**. Assistance refers to systems that provide limited support, such as alerts or risk estimates. Augmentation describes technologies that enhance human performance by extending analytical capacity or reducing cognitive burden. Autonomy, in contrast, implies a higher level of system independence in generating or executing decisions (O'Sullivan et al., 2019). Although most current surgical AI applications remain within the domains of assistance and augmentation, there is a growing interest in semi-autonomous systems, particularly in robotic surgery.

Another foundational concept is **trust**, which can be defined as the willingness of clinicians and patients to rely on AI systems under conditions of uncertainty (Cobianchi et al., 2023). Trust is not solely determined by technical performance but is influenced by transparency, reliability, and interpretability. In surgical contexts, where decisions often have immediate and high-stakes consequences, trust becomes a critical factor in the adoption and effective use of AI systems.

Closely related to trust is the concept of **explainability**, often addressed through the development of explainable artificial intelligence (XAI). Explainability refers to the extent to which users can understand how an AI system produces its outputs. In surgery, explainability is essential not only for building trust but also for ensuring accountability and supporting informed decision-making (Laios et al., 2022). Lack of explainability may lead to skepticism or, conversely, to uncritical acceptance of algorithmic recommendations, both of which can negatively affect clinical outcomes.

The literature also emphasizes the importance of **autonomy**, both at the level of the clinician and the patient. Professional autonomy refers to the surgeon's ability to exercise independent clinical judgment, while patient autonomy concerns the right to make informed decisions about treatment (Teasdale et al., 2024). AI systems may influence both forms of autonomy by shaping how information is presented, interpreted, and acted upon in clinical settings.

Finally, the concept of **responsibility**—often discussed alongside accountability and liability—represents a fundamental issue in AI-assisted surgery. In traditional clinical practice, responsibility is primarily attributed to the surgeon. However, as AI systems become increasingly involved in decision-making, responsibility may become distributed across multiple actors, including clinicians, institutions, and technology developers (Morris et al., 2023). This shift challenges existing ethical and legal frameworks and necessitates new approaches to accountability in healthcare.

Taken together, these foundational concepts demonstrate that AI in surgery is not merely a technological innovation but a **multidimensional transformation** affecting how decisions are made, justified, and governed. Understanding these conceptual frameworks is essential for analyzing the broader implications of AI-driven surgical decision-making and for evaluating its impact on clinical practice.

### 3.2. Trust in AI-Driven Surgical Decision-Making (Revised Version)

Trust is widely recognized as a fundamental determinant of the successful integration of artificial intelligence (AI) into surgical practice. In high-risk clinical environments such as the operating room, where decisions must be made rapidly and often under conditions of uncertainty, trust in AI-driven decision-support systems directly influences both adoption and clinical utilization (Brandenburg et al., 2025; Cobianchi et al., 2023). Unlike in lower-risk domains, trust in surgical AI is closely tied not only to system performance but also to ethical, cognitive, and professional considerations.

In the context of surgical decision-making, trust can be defined as the willingness of clinicians to rely on algorithmic outputs despite uncertainty and potential risks. The literature consistently demonstrates that trust is a **multidimensional construct**, shaped by factors including accuracy, reliability, transparency, interpretability, and consistency of AI systems (Lopes et al., 2025). Importantly, studies indicate that high predictive accuracy alone is insufficient to ensure clinical adoption if the reasoning behind AI-generated outputs is not accessible to users.

A central challenge identified in the literature is the **"black box" problem**, particularly associated with deep learning models. These systems often produce highly accurate predictions without offering clear explanations of how specific conclusions are reached. For surgeons, who are expected to justify clinical decisions based on evidence and reasoning, such opacity can undermine confidence and limit trust (Laios et al., 2022). Research suggests that clinicians are more likely to trust AI systems that provide interpretable outputs, even when their predictive performance is slightly lower than that of more complex models (Lopes et al., 2025).

Closely related to transparency is the concept of **explainable artificial intelligence (XAI)**, which has emerged as a key strategy for improving trust in AI systems. XAI approaches aim to make algorithmic processes more understandable, allowing clinicians to evaluate the relevance and reliability of recommendations. In surgical contexts, explainability is particularly important due to the immediate and often irreversible consequences of decisions. It enables surgeons to integrate AI outputs with their own expertise and to communicate reasoning to patients and colleagues (Laios et al., 2022).

Another important factor influencing trust is **user experience and familiarity with AI technologies**. Empirical studies show that surgeons with prior exposure to AI tools are more likely to adopt and effectively use them in clinical practice (Pecqueux et al., 2022). Conversely, lack of training or familiarity may lead to skepticism, misuse, or inappropriate reliance on AI outputs. This suggests that trust is not solely a property of the system itself but also depends on the competence and confidence of its users.

The literature also highlights the phenomenon of **automation bias**, defined as the tendency to favor automated recommendations over independent judgment. In surgical environments, where cognitive load and time pressure are high, automation bias can lead to the acceptance of incorrect AI outputs or the neglect of contradictory clinical evidence (St John et al., 2024). While trust is necessary for effective human–AI collaboration, excessive or uncritical trust may compromise patient safety. Therefore, the goal is not simply to increase trust but to achieve an appropriate calibration between trust and critical evaluation.

Trust is further influenced by the **clinical and institutional context** in which AI systems are implemented. Factors such as the type of surgical procedure, perceived risk level, and organizational culture play a significant role in shaping attitudes toward AI. Surgeons may be more inclined to rely on AI in routine or low-risk scenarios, while remaining cautious in complex or high-stakes procedures (Cobianchi et al., 2023). Additionally, institutional endorsement, regulatory approval, and validation studies contribute to perceived system credibility and legitimacy.

Beyond clinicians, **patient trust** represents an important dimension of AI adoption in surgery. Studies indicate that patients are generally receptive to AI when it is presented as a supportive tool rather than a replacement for the surgeon (Teasdale et al., 2024). However, concerns related to data privacy, depersonalization of care, and lack of transparency may reduce acceptance. Effective communication and shared decision-making are therefore essential for maintaining trust at the patient level.

At a systemic level, trust is also shaped by the presence of **regulatory frameworks and ethical guidelines**. Clear standards for validation, monitoring, and accountability can enhance confidence in AI systems, while the absence of such frameworks may lead to uncertainty and resistance (O'Sullivan et al., 2019). This highlights the importance of governance structures in supporting the responsible integration of AI into surgical practice.

Finally, trust in AI should be understood as a **dynamic and evolving process**. It develops over time as users gain experience with a system, observe its performance, and integrate it into their workflows. Trust may increase as systems demonstrate reliability but can also be undermined by errors, inconsistencies, or lack of transparency. Continuous evaluation and iterative improvement are therefore essential for maintaining appropriate levels of trust.

In summary, trust in AI-driven surgical decision-making is a complex and context-dependent phenomenon that extends beyond technical performance. It encompasses issues related to transparency, explainability, user competence, cognitive bias, patient perception, and institutional governance. Achieving a balanced and well-calibrated form of trust is essential for ensuring that AI systems enhance, rather than compromise, the quality and safety of surgical care.

### 3.3. Autonomy and Human Oversight in AI-Assisted Surgery (Revised Version)

The integration of artificial intelligence (AI) into surgical practice has introduced significant challenges to the traditional concept of autonomy in medicine. Autonomy, long regarded as a foundational principle of clinical practice, is increasingly being redefined within the context of human–AI collaboration. In surgical decision-making, autonomy operates on multiple levels, including **professional autonomy** of surgeons and **patient autonomy**, both of which are influenced by the growing role of algorithmic systems (St John et al., 2024; Teasdale et al., 2024).

Professional autonomy refers to the surgeon's ability to exercise independent clinical judgment based on expertise, experience, and situational awareness. Traditionally, surgical decisions have been shaped by a combination of formal medical knowledge and tacit understanding developed through clinical practice. However, AI systems—particularly those providing predictive analytics or treatment recommendations—introduce an additional layer of influence that may alter how decisions are formulated and justified (Cobianchi et al., 2022).

A key tension identified in the literature is the balance between **augmentation and substitution**. AI is typically presented as a tool designed to augment human capabilities by providing data-driven insights and reducing cognitive burden. In this model, the surgeon remains the final decision-maker, using AI as a supplementary source of information. However, as AI systems become more accurate and influential, there is a risk that they may shift toward partial substitution, subtly shaping clinical decisions and encouraging reliance on algorithmic outputs (O'Sullivan et al., 2019).

The phenomenon of **automation bias** is particularly relevant in this context. Automation bias refers to the tendency to over-rely on automated systems, even when their outputs may be incorrect or incomplete. In surgical environments, where time pressure and cognitive load are high, the presence of AI-generated recommendations may increase the likelihood of such bias (St John et al., 2024). Surgeons may defer to algorithmic outputs without sufficient critical evaluation, potentially undermining independent clinical reasoning. Conversely, under-reliance may also occur when trust in AI systems is insufficient, highlighting the need for a balanced and calibrated approach.

Closely related to autonomy is the concept of **human oversight**, often described as a "human-in-the-loop" model. In current clinical practice, AI systems are typically implemented as decision-support tools, with the surgeon retaining ultimate authority over clinical decisions. This model is intended to preserve professional autonomy while benefiting from AI capabilities. However, the effectiveness of human oversight depends on several factors, including the interpretability of AI outputs, the design of user interfaces, and the clinician's ability to critically evaluate algorithmic recommendations (Lopes et al., 2025).

The literature also raises concerns regarding **deskilling**, defined as the potential decline in clinical expertise due to overreliance on automated systems. In surgery, this is a particularly important issue, as both technical skills and decision-making capabilities are developed through continuous practice and exposure to complex cases. If AI systems increasingly guide decision-making processes, there is a risk that future surgeons may become less proficient in independent reasoning, particularly in situations where AI support is unavailable or unreliable (Kinross et al., 2025).

Patient autonomy represents another critical dimension affected by AI integration. Informed consent, a fundamental principle of ethical medical practice, relies on patients' ability to understand the rationale behind clinical decisions. The involvement of AI complicates this process, as algorithmic contributions may be difficult to explain in accessible terms. Patients may not fully understand how AI influences treatment recommendations, which can challenge the transparency and validity of consent (Teasdale et al., 2024).

The concept of **shared decision-making** must therefore be reconsidered in the context of AI-assisted surgery. Traditionally, shared decision-making involves collaboration between clinician and patient. With the introduction of AI, this process becomes more complex, involving a triadic interaction between the surgeon, the patient, and the algorithmic system. This raises important questions regarding how information should be communicated, how much weight should be given to AI recommendations, and how disagreements between human judgment and algorithmic output should be resolved.

Institutional and systemic factors also influence autonomy. Healthcare systems increasingly promote standardized protocols and evidence-based practices, many of which are supported by AI technologies. While these systems can improve consistency and reduce variability in care, they may also constrain individual decision-making by encouraging adherence to algorithmic recommendations (Morris et al., 2023). This tension between standardization and professional discretion represents a key challenge in modern surgical practice.

At a broader level, the relationship between surgeons and AI systems can be conceptualized as a shift toward **collaborative intelligence**, in which human and machine capabilities are combined to achieve improved outcomes. In this model, autonomy is not eliminated but reconfigured. Surgeons are expected to develop new competencies, including the ability to interpret AI outputs, recognize system limitations, and integrate algorithmic insights into clinical reasoning (Cobianchi et al., 2023).

In conclusion, the integration of AI into surgical decision-making challenges traditional notions of autonomy while simultaneously creating opportunities for its transformation. Maintaining an appropriate balance between human judgment and algorithmic support is essential for preserving professional integrity and ensuring patient-centered care. The literature suggests that the future of surgical autonomy will depend not only on technological advancements but also on the development of training, governance, and ethical frameworks that support effective human–AI collaboration.

### 3.4. Responsibility, Accountability, and Liability in AI-Driven Surgical Decision-Making (Revised Version)

The integration of artificial intelligence (AI) into surgical decision-making introduces profound challenges related to responsibility, accountability, and liability. These concepts, traditionally well-defined within human-centered clinical practice, become significantly more complex when decision-making processes involve both human clinicians and algorithmic systems. As AI evolves from a purely supportive tool toward a more influential component of clinical reasoning, determining responsibility for outcomes—particularly adverse events—becomes increasingly ambiguous (O'Sullivan et al., 2019; Morris et al., 2023).

In conventional surgical practice, responsibility is primarily attributed to the operating surgeon, who is expected to exercise professional judgment, adhere to established standards of care, and ensure patient safety. This framework is grounded in the assumption that decisions are made by identifiable human agents. However, AI-assisted decision-making introduces a form of **distributed agency**, in which outcomes may result from the interaction of multiple actors, including surgeons, healthcare institutions, AI systems, and technology developers (Cobianchi et al., 2022).

The concept of **distributed agency** represents a fundamental shift in how responsibility is understood in surgical contexts. When an AI system contributes to clinical recommendations—such as predicting complications or suggesting treatment strategies—it becomes difficult to determine whether responsibility lies with the surgeon who accepted the recommendation, the developers who designed the algorithm, or the institution that implemented the system. This diffusion of responsibility challenges traditional medico-legal models and complicates accountability structures.

Closely related to responsibility is the issue of **liability**, particularly within legal frameworks. Existing legal systems are largely based on human accountability and are not fully adapted to address the complexities introduced by AI technologies. In most jurisdictions, the surgeon remains legally responsible for clinical decisions, even when those decisions are influenced by AI (O'Sullivan et al., 2019). This creates a potential mismatch between practical decision-making processes and legal responsibility, as clinicians may be held liable for outcomes influenced by systems over which they have limited control or understanding.

The literature also emphasizes the importance of **explainability and traceability** in assigning responsibility. In order to determine accountability, it is necessary to understand how a decision was made and what role the AI system played in that process. However, many AI models—particularly those based on deep learning—lack transparency, making it difficult to reconstruct the reasoning behind specific outputs (Lopes et al., 2025). This lack of interpretability not only affects trust but also creates significant challenges for legal and ethical evaluation.

Another critical dimension concerns the role of **healthcare institutions**. Hospitals and healthcare organizations are responsible for selecting, implementing, and monitoring AI systems. As such, they bear a degree of institutional responsibility for ensuring that these technologies are safe, validated, and appropriately integrated into clinical workflows (Morris et al., 2023). This includes providing adequate training for clinicians, establishing governance frameworks, and implementing safeguards to mitigate risks associated with AI use.

The role of **technology developers and industry stakeholders** is also increasingly recognized in the literature. Developers are responsible for designing AI systems that are reliable, unbiased, and aligned with clinical needs. Issues such as data bias, model robustness, and usability fall within their domain of responsibility. As AI systems become more embedded in clinical practice, there is growing debate regarding whether developers should bear partial responsibility for adverse outcomes associated with their technologies (Morris et al., 2023).

From an ethical perspective, responsibility in AI-assisted surgery is closely linked to core principles such as **non-maleficence, beneficence, and justice**. Ensuring that AI systems do not cause harm, contribute positively to patient outcomes, and are applied equitably across populations is a shared responsibility among all stakeholders. The literature increasingly suggests that responsibility in this context should be understood as a **collective and systemic phenomenon**, rather than as an individual attribute.

An additional layer of complexity arises with the development of **semi-autonomous and autonomous surgical systems**, particularly in the field of robotic surgery. As these systems gain greater independence, the boundary between tool and agent becomes less clear. This raises fundamental questions regarding whether AI systems can be considered moral or legal agents and whether existing frameworks of responsibility are sufficient to account for their role (O'Sullivan et al., 2019).

In response to these challenges, recent research has emphasized the need for **structured accountability frameworks**. Such frameworks aim to clearly define roles and responsibilities across different stages of AI development and implementation. Key elements often include transparency, auditability, human oversight, and continuous monitoring of system performance. By establishing clearer governance structures, these frameworks seek to reduce ambiguity and enhance trust in AI-driven decision-making.

Despite these efforts, significant gaps remain. There is currently no global consensus on how responsibility should be distributed among stakeholders or how liability should be assigned in cases involving AI-assisted decisions. This uncertainty represents a major barrier to the widespread adoption of AI in surgical practice and highlights the need for interdisciplinary collaboration between clinicians, legal experts, ethicists, and engineers.

In conclusion, responsibility, accountability, and liability represent some of the most complex and unresolved challenges in AI-driven surgical decision-making. The transition from human-centered to hybrid human–AI systems necessitates a redefinition of traditional responsibility frameworks to reflect the realities of distributed agency. Addressing these challenges is essential for ensuring patient safety, legal clarity, and the ethical integration of AI into surgical practice.

### 3.5. Applications of Artificial Intelligence in Surgical Decision-Making (Revised Version)

The application of artificial intelligence (AI) in surgical practice has expanded significantly over the past decade, influencing multiple stages of perioperative care. While previous sections have addressed conceptual and ethical dimensions, it is essential to examine how AI is currently implemented in clinical settings. The literature demonstrates that AI is not a single-purpose technology but a diverse set of tools applied across **preoperative, intraoperative, and postoperative phases** of surgical care (Kinross et al., 2025; Gao et al., 2021).

#### 3.5.1. Preoperative Decision Support and Risk Prediction

One of the most established applications of AI in surgery is **preoperative risk assessment and decision support**. Machine learning models are widely used to predict postoperative complications, mortality, readmission rates, and surgical outcomes. These systems analyze large datasets derived from electronic health records, imaging, and patient demographics to generate individualized risk profiles (Gao et al., 2021).

For instance, predictive models have demonstrated the ability to support surgical planning in areas such as colorectal surgery and bariatric procedures, enabling surgeons to make more informed decisions regarding treatment strategies and patient selection (Hsu et al., 2023; Shao et al., 2023). By providing probabilistic estimates, AI systems enhance clinical judgment and improve patient counseling.

However, the literature emphasizes that predictive accuracy must be balanced with interpretability. Surgeons must critically evaluate AI-generated outputs, particularly when they conflict with clinical experience or contextual factors (Hsu et al., 2023).

#### 3.5.2. Intraoperative Guidance and Surgical Navigation

AI technologies are increasingly applied during surgical procedures, particularly through **computer vision and real-time decision support systems**. These systems analyze intraoperative video and imaging data to assist surgeons in identifying anatomical structures, detecting critical regions, and guiding surgical actions (Kinross et al., 2025).

In minimally invasive and robotic surgery, AI-enhanced platforms can provide real-time feedback on instrument positioning, tissue handling, and procedural steps. This contributes to improved precision and reduced variability in surgical performance, particularly in complex procedures.

Additionally, the integration of **augmented reality (AR)** and AI-driven visualization tools allows surgeons to overlay preoperative imaging onto the operative field, enhancing spatial awareness and decision-making accuracy. While these technologies remain largely assistive, their increasing sophistication suggests a gradual shift toward more advanced intraoperative support systems.

#### 3.5.3. Postoperative Outcome Prediction and Monitoring

AI is also widely used in the **postoperative phase**, particularly for monitoring patient outcomes and predicting complications. Machine learning models can analyze clinical and physiological data to detect early signs of adverse events, such as infections, bleeding, or organ dysfunction (Hsu et al., 2023).

These systems enable earlier intervention and may contribute to improved patient outcomes, reduced complication rates, and shorter hospital stays. Furthermore, AI-driven analytics can support long-term outcome assessment, including functional recovery and quality of life, thereby informing future clinical decision-making.

#### 3.5.4. Surgical Robotics and Semi-Autonomous Systems

The integration of AI with **robotic surgical systems** represents one of the most advanced developments in the field. While current robotic platforms are primarily controlled by surgeons, AI components are increasingly being incorporated to enhance precision, stability, and efficiency (O'Sullivan et al., 2019).

Emerging research explores the development of **semi-autonomous systems**, capable of performing specific surgical tasks, such as suturing or tissue manipulation, with minimal human input. Although fully autonomous surgery remains largely experimental, these developments raise important questions regarding control, responsibility, and the evolving role of the surgeon.

The literature suggests that AI-driven robotics will likely support a model of **collaborative control**, where human expertise and machine capabilities are integrated rather than replaced (Cobianchi et al., 2022).

#### 3.5.5. Clinical Decision Support Systems (CDSS)

AI-powered **clinical decision support systems (CDSS)** are among the most widely implemented applications in surgical practice. These systems integrate patient data, clinical guidelines, and predictive models to generate recommendations for diagnosis, treatment planning, and operative strategies (Shao et al., 2023).

In emergency surgery, CDSS can assist in prioritizing cases based on urgency and predicted outcomes. In oncology, they support decisions related to resectability, treatment sequencing, and risk–benefit analysis. Such systems have the potential to improve efficiency, consistency, and evidence-based decision-making.

However, their effectiveness depends on user trust, system design, and integration into clinical workflows. Poorly designed systems may lead to alert fatigue or underutilization, limiting their clinical impact (Morris et al., 2023).

#### 3.5.6. Personalized and Precision Surgery

A particularly promising application of AI is its role in **personalized or precision surgery**. By analyzing patient-specific data—including genetic, clinical, and imaging information—AI systems can support tailored treatment strategies that account for individual variability (Kinross et al., 2025).

This approach enables more nuanced decision-making, allowing surgeons to select optimal surgical techniques and predict individual responses to treatment. However, personalized decision-making also increases complexity and requires high levels of interpretability to ensure that recommendations are clinically meaningful and ethically justified.

#### 3.5.7. Summary of Applications

Overall, the literature demonstrates that AI is already embedded in multiple aspects of surgical decision-making. Its applications range from predictive analytics and intraoperative guidance to robotic assistance and personalized care. These technologies have the potential to improve accuracy, efficiency, and patient outcomes (Gao et al., 2021; Kinross et al., 2025).

At the same time, their implementation introduces challenges related to trust, autonomy, and responsibility, as discussed in previous sections. The practical application of AI in surgery therefore cannot be separated from its ethical and cognitive implications. Understanding these real-world applications is essential for evaluating the broader impact of AI on surgical decision-making.

### 3.6. Challenges and Limitations of Artificial Intelligence in Surgery (Revised Version)

Despite the significant potential of artificial intelligence (AI) to enhance surgical decision-making, the literature consistently highlights a range of challenges and limitations that constrain its safe, effective, and ethical implementation. These challenges extend beyond technical performance and encompass issues related to data quality, interpretability, human factors, regulatory uncertainty, and ethical considerations (Morris et al., 2023; Rad et al., 2025).

#### 3.6.1. Data Quality and Algorithmic Bias

One of the most fundamental limitations of AI systems lies in their dependence on the quality and representativeness of training data. Machine learning models are inherently data-driven, and their performance is directly influenced by the datasets on which they are trained. In surgical contexts, these datasets often vary across institutions, patient populations, and clinical practices, which may limit generalizability (Gao et al., 2021).

A particularly important concern is **algorithmic bias**, which occurs when training data do not adequately represent diverse populations. This can result in systematic disparities in prediction accuracy and clinical recommendations, potentially leading to inequitable treatment outcomes (Rad et al., 2025). Addressing data bias is therefore critical for ensuring fairness and reliability in AI-assisted surgical decision-making.

#### 3.6.2. Lack of Transparency and Explainability

The lack of transparency in many AI systems—especially deep learning models—remains a major barrier to clinical adoption. As previously discussed, these systems often function as “black boxes,” providing outputs without clear explanations of how decisions are generated (Lopes et al., 2025).

This lack of **explainability** affects not only trust but also accountability. Surgeons are required to justify their clinical decisions, and reliance on opaque algorithmic recommendations may complicate this process. Studies suggest that explainable AI (XAI) approaches are essential for improving interpretability and facilitating responsible use in clinical practice (Laios et al., 2022).

### 3.6.3. Automation Bias and Overreliance

Another major challenge is the risk of **automation bias**, defined as the tendency to over-rely on automated systems. In high-pressure surgical environments, clinicians may accept AI-generated recommendations without sufficient critical evaluation, particularly when systems are perceived as highly accurate (St John et al., 2024).

This overreliance can lead to errors, especially when AI outputs are incorrect or incomplete. Conversely, underutilization due to distrust or lack of familiarity may also limit the potential benefits of AI. Achieving a balanced and calibrated relationship between human judgment and algorithmic input is therefore essential.

### 3.6.4. Integration into Clinical Workflows

The successful implementation of AI systems depends heavily on their integration into existing **clinical workflows**. Even highly accurate systems may fail to achieve meaningful impact if they are not compatible with clinical routines or user needs (Morris et al., 2023).

Challenges include poor user interface design, increased cognitive load, and alert fatigue. Additionally, lack of interoperability with electronic health record systems can limit usability and scalability. These factors highlight the importance of designing AI systems that are not only technically robust but also user-centered and workflow-compatible.

### 3.6.5. Regulatory and Legal Uncertainty

The rapid development of AI technologies has outpaced the evolution of **regulatory frameworks**, creating uncertainty regarding safety, validation, and liability. Existing regulatory systems are not fully adapted to the dynamic and adaptive nature of AI, particularly systems that continue to learn after deployment (O'Sullivan et al., 2019).

This lack of clear regulation may hinder adoption, as clinicians and institutions may be reluctant to use systems without well-defined legal protections. Furthermore, the absence of standardized evaluation criteria makes it difficult to compare and validate different AI solutions.

### 3.6.6. Ethical Concerns and Patient Perception

AI integration in surgery raises several **ethical challenges**, including issues related to patient autonomy, informed consent, data privacy, and fairness. Patients may have limited understanding of how AI systems contribute to clinical decisions, which can complicate the consent process and affect trust (Teasdale et al., 2024).

Additionally, concerns about the potential **dehumanization of care** have been raised, as increased reliance on technology may reduce direct interaction between patients and clinicians. Maintaining a balance between technological efficiency and human-centered care remains a key challenge.

### 3.6.7. Training, Education, and Deskilling

The effective use of AI in surgery requires new competencies among healthcare professionals. Surgeons must be able to interpret AI outputs, understand system limitations, and integrate algorithmic insights into clinical reasoning (Kinross et al., 2025).

At the same time, the literature highlights concerns regarding **deskilling**, where overreliance on AI may lead to a decline in independent decision-making abilities. This is particularly relevant in surgical training, where hands-on experience and critical thinking are essential for professional development.

### 3.6.8. Technological Limitations and Reliability

AI systems are not infallible and may be subject to technical limitations, including overfitting, lack of external validation, and sensitivity to variations in input data (Gao et al., 2021). Many models demonstrate high performance in controlled research settings but fail to achieve similar results in real-world clinical environments.

Additionally, system errors or technical failures may occur, particularly in complex surgical settings. These limitations underscore the importance of maintaining human oversight and ensuring that AI systems are used as supportive tools rather than definitive decision-makers.

### 3.6.9. Summary of Challenges

In summary, while AI offers significant opportunities to enhance surgical decision-making, its implementation is accompanied by substantial challenges related to data quality, transparency, cognitive bias, workflow integration, regulatory uncertainty, and ethical considerations. These challenges are interconnected and must be addressed through a comprehensive and interdisciplinary approach.

The literature suggests that overcoming these limitations requires not only technological innovation but also the development of robust governance frameworks, improved training programs, and continuous evaluation of system performance. Only through such efforts can AI be safely and effectively integrated into surgical practice.

## 4. Discussion

The integration of artificial intelligence (AI) into surgical decision-making represents a profound transformation of clinical practice, extending beyond technological innovation into the domains of cognition, ethics, and professional responsibility. The findings of this review indicate that AI is not merely an additional tool in the surgical environment, but a transformative force that reshapes how decisions are generated, evaluated, and justified. The discussion that follows synthesizes the key themes identified in the literature—trust, autonomy, and responsibility—while situating them within the broader context of human–AI collaboration in medicine.

One of the central insights emerging from this review is that **AI fundamentally alters the epistemological basis of surgical decision-making**. Traditionally, decisions in surgery have relied on a combination of empirical evidence, clinical expertise, and tacit knowledge. The introduction of AI introduces a new form of knowledge—algorithmic inference—which is often probabilistic, data-driven, and not always transparent. This shift creates a hybrid decision-making model in which human reasoning and machine-generated outputs coexist. While this has the potential to enhance decision accuracy and reduce variability, it also raises questions about how knowledge is validated and whose judgment ultimately prevails.

The issue of **trust** emerges as a critical factor in determining the success or failure of AI integration. The literature consistently demonstrates that trust is not determined solely by the accuracy of AI systems but is influenced by a combination of transparency, explainability, reliability, and user experience. Importantly, trust must be carefully calibrated. Both under-trust and over-trust can lead to suboptimal outcomes: skepticism may prevent the use of beneficial technologies, while excessive reliance may result in automation bias and reduced critical oversight. This suggests that future research and implementation strategies should focus not only on improving algorithmic performance but also on designing systems that support appropriate levels of trust.

Closely related to trust is the evolving concept of **autonomy**. The findings of this review indicate that AI does not eliminate surgical autonomy but rather reconfigures it. Surgeons remain central to decision-making processes; however, their role is increasingly mediated by algorithmic recommendations. This creates a dynamic tension between human judgment and machine input. On one hand, AI can enhance autonomy by providing additional information and reducing uncertainty. On the other hand, it may constrain autonomy by shaping how decisions are framed and by introducing implicit pressures to conform to algorithmic outputs. The concept of autonomy must therefore be redefined in relational terms, acknowledging that decisions are now produced within a network of human and technological actors.

The theme of **responsibility** represents perhaps the most complex and unresolved issue identified in this review. The literature highlights a growing gap between traditional models of individual accountability and the realities of distributed decision-making in AI-assisted environments. As agency becomes shared among surgeons, institutions, and technological systems, assigning responsibility for outcomes becomes increasingly difficult. Current legal and regulatory frameworks remain largely centered on human actors, which may be insufficient to address the complexities introduced by AI. This misalignment has practical implications, as uncertainty regarding liability may discourage adoption and create ethical dilemmas in clinical practice.

Another important finding concerns the role of **human–AI interaction** as a determinant of clinical outcomes. The effectiveness of AI in surgery is not solely dependent on technical performance but also on how systems are integrated into workflows, interpreted by users, and embedded within institutional contexts. Factors such as interface design, training, cognitive load, and organizational culture significantly influence how AI is used in practice. This reinforces the idea that AI in surgery should be understood as a **sociotechnical system**, in which technological and human elements are deeply interconnected.

The analysis of current applications reveals that AI is already making meaningful contributions to surgical decision-making, particularly in areas such as risk prediction, intraoperative guidance, and clinical decision support. However, these applications remain predominantly assistive rather than autonomous. This suggests that the future of AI in surgery is likely to be characterized by **collaborative intelligence**, where human expertise and machine capabilities are combined rather than replaced. Such a model has the potential to improve outcomes while preserving the central role of the surgeon, but it requires careful design and governance.

At the same time, the review identifies several **persistent challenges** that must be addressed to fully realize the benefits of AI in surgery. These include issues related to data quality and bias, lack of explainability, integration into clinical workflows, and regulatory uncertainty. Notably, many of these challenges are interconnected. For example, lack of transparency not only affects trust but also complicates accountability and informed consent. Similarly, data bias has implications for both clinical accuracy and ethical fairness. Addressing these challenges therefore requires a holistic approach that considers technical, ethical, and organizational dimensions simultaneously.

The review also highlights several **gaps in the existing literature**. First, there is a lack of empirical studies examining real-time decision-making dynamics in AI-assisted surgical environments. Much of the current literature is either conceptual or based on surveys, which may not fully capture the complexity of clinical practice. Second, there is limited research on long-term outcomes associated with AI integration, particularly in terms of its impact on surgical training, professional identity, and skill development. Third, the legal and regulatory aspects of AI in surgery remain underexplored, with a need for more concrete frameworks that can be applied in practice.

From a practical perspective, the findings of this review suggest several **implications for clinical practice and policy**. Healthcare institutions should prioritize the implementation of AI systems that are transparent, interpretable, and aligned with clinical workflows. Training programs should be developed to equip surgeons with the skills needed to critically engage with AI technologies. At the policy level, there is a need for clearer regulatory guidelines that address issues of validation, accountability, and liability. Collaboration between clinicians, engineers, ethicists, and legal experts will be essential in developing these frameworks.

Finally, this review underscores the importance of adopting a **balanced and reflective approach** to AI in surgery. While the potential benefits of AI are substantial, they must be weighed against the risks and limitations identified in the literature. The goal should not be to maximize the use of AI at all costs, but to integrate it in a way that enhances clinical decision-making while preserving the core values of surgical practice, including patient safety, professional integrity, and ethical responsibility.

In conclusion, AI-driven surgical decision-making represents a paradigm shift that challenges existing models of knowledge, autonomy, and responsibility. Its successful integration will depend not only on technological advancements but also on the development of robust ethical, legal, and organizational frameworks. By addressing these challenges and embracing a collaborative model of human–AI interaction, the field of surgery can harness the full potential of AI while maintaining its foundational principles.

## 5. Conclusions

The integration of artificial intelligence (AI) into surgical decision-making represents a significant and ongoing transformation of modern surgical practice. This review has demonstrated that AI is not merely a technological innovation but a multidimensional force that reshapes clinical reasoning, professional roles, and ethical frameworks within surgery. By synthesizing existing literature, this study has highlighted the central importance of trust, autonomy, and responsibility as key dimensions through which AI-driven decision-making must be understood.

The findings indicate that AI has substantial potential to enhance surgical outcomes through improved predictive accuracy, decision support, and procedural guidance. At the same time, its integration introduces complex challenges related to transparency, cognitive bias, professional autonomy, and legal accountability. The coexistence of human judgment and algorithmic inference creates a hybrid decision-making environment that requires careful balancing to ensure both effectiveness and safety.

This review underscores that trust in AI systems must be appropriately calibrated, supported by explainability, reliability, and user competence. Similarly, surgical autonomy must be redefined in the context of human–AI collaboration, preserving the central role of the surgeon while acknowledging the influence of algorithmic systems. The issue of responsibility remains one of the most pressing challenges, as traditional models of accountability are increasingly insufficient in the face of distributed decision-making processes.

Several gaps in the current literature have been identified, including the lack of empirical studies on real-time human–AI interaction in surgical settings, limited understanding of long-term impacts on surgical training and professional development, and insufficiently developed legal and regulatory frameworks. Addressing these gaps will be essential for the responsible advancement of AI in surgery.

Future research should focus on interdisciplinary approaches that integrate clinical, technological, ethical, and legal perspectives. In particular, there is a need for the development of standardized evaluation methods, transparent and interpretable AI systems, and robust governance structures that clearly define roles and responsibilities. Additionally, education and training programs must evolve to prepare surgeons for effective collaboration with AI technologies.

In conclusion, AI-driven surgical decision-making offers significant opportunities to improve the quality and efficiency of care, but its successful implementation depends on more than technological capability alone. It requires a comprehensive and ethically grounded approach that ensures patient safety, supports clinical judgment, and establishes clear frameworks for responsibility. By addressing these challenges, the surgical community can harness the benefits of AI while maintaining the core values that underpin medical practice.

## REFERENCES


1. Brandenburg, L., et al. (2025). Can surgeons trust artificial intelligence? *Annals of Surgery, 281*(2), 345–352. https://doi.org/10.1016/j.jss.2022.04.013
2. Braunstein, M., et al. (2025). Perceptions of artificial intelligence among surgeons and medical students. *Surgical Innovation, 32*(1), 55–63. https://doi.org/10.1007/s11845-025-04079-z
3. Cobianchi, L., et al. (2022). Artificial intelligence in surgery: Ethical dilemmas and open issues. *Journal of Surgical Research, 278*, 1–9. https://doi.org/10.1016/j.jss.2022.04.013
4. Cobianchi, L., et al. (2023). Surgeons' perspectives on artificial intelligence to support clinical decision-making in trauma and emergency contexts. *World Journal of Emergency Surgery, 18*(1), 12–20. https://doi.org/10.1186/s13017-023-00463-6
5. De Simone, B., et al. (2025). Ethical integration of artificial intelligence in emergency surgery: Challenges and perspectives. *World Journal of Emergency Surgery, 20*(1), 45–56. https://doi.org/10.1097/JS9.0000000000002347
6. Dhawan, N., et al. (2025). Ethical principles for generative AI in plastic surgery. *Aesthetic Surgery Journal, 45*(3), 210–220. https://doi.org/10.1097/GOX.0000000000006825
7. Duffourc, M., et al. (2025). Surgeons' perspectives on liability in artificial intelligence-driven surgical technologies. *Annals of Surgery Open, 6*(1), e345. https://doi.org/10.1097/AS9.0000000000000542
8. Dupre, A., et al. (2025). Machine learning in emergency surgical coordination. *British Journal of Surgery, 112*(3), 345–353. https://doi.org/10.1007/s10877-025-01341-8
9. Ferreres, A. R., et al. (2024). Ethical aspects of artificial intelligence in surgical care. *Annals of Surgery, 279*(4), 567–575. https://doi.org/10.1097/SLA.0000000000006179
10. Fangerau, H., et al. (2024). Ethical considerations of artificial intelligence in surgery. *BMC Medical Ethics, 25*(1), 89. https://doi.org/10.1186/s12910-024-01007-4
11. Gao, Y., et al. (2021). Machine learning prediction of mortality in emergency general surgery. *Annals of Surgery, 274*(6), 1020–1027. https://doi.org/10.1097/SLA.0000000000004416
12. Henn, D., et al. (2025). Artificial intelligence in clinical decision-making for acute abdominal pain: A survey among surgeons. *International Journal of Surgery, 109*, 45–53. https://doi.org/10.1055/a-2125-1559
13. Hmido, F., et al. (2025). Patient perspectives on artificial intelligence in surgical decision-making. *Patient Education and Counseling, 118*, 107–115. https://doi.org/10.1136/bmjsit-2024-000365
14. Hsu, C. Y., et al. (2023). Machine learning models for predicting postoperative complications in bariatric surgery. *Surgery for Obesity and Related Diseases, 19*(5), 789–798. https://doi.org/10.1016/j.soard.2023.02.006
15. Kinross, J., et al. (2025). Artificial intelligence in colorectal surgery: A roadmap for implementation. *The Lancet Digital Health, 7*(2), e120–e130. https://doi.org/10.1111/codi.70232
16. Laios, A., et al. (2022). Explainable artificial intelligence for predicting surgical effort in ovarian cancer. *Cancers, 14*(15), 3650. https://doi.org/10.3390/cancers14153650
17. Loos, M. J., et al. (2024). Machine learning versus surgeon prediction in carpal tunnel surgery outcomes. *Journal of Hand Surgery, 49*(2), 150–158. https://doi.org/10.1227/neu.0000000000002848
18. Lopes, R., et al. (2025). Explainable artificial intelligence in surgery: Current perspectives. *Artificial Intelligence in Medicine, 145*, 102345. https://doi.org/10.1177/00031348221117042
19. Mehmet, A., et al. (2025). Artificial intelligence versus spinal surgeons: Decision-making in controversial cases. *Spine Journal, 25*(4), 678–687. https://doi.org/10.1007/s00586-025-08825-w

20. Morris, M. A., et al. (2023). Ethical, legal, and financial implications of artificial intelligence in surgery. *JAMA Surgery, 158*(3), 245–252. https://doi.org/10.1177/00031348221117042
21. Oosterhoff, J. H., et al. (2024). Machine learning versus conventional prediction models in orthopedic surgery. *Journal of Bone and Joint Surgery, 106*(6), 501–510. https://doi.org/10.1097/CORR.0000000000003018
22. O'Sullivan, S., et al. (2019). Legal, regulatory, and ethical frameworks for AI in surgery. *International Journal of Medical Robotics and Computer Assisted Surgery, 15*(4), e1967. https://doi.org/10.1002/rcs.1968
23. Pecqueux, M., et al. (2022). Surgeons' knowledge and acceptance of artificial intelligence. *International Journal of Surgery, 98*, 106–112. https://doi.org/10.3389/fpubh.2022.982335
24. Rad, A., et al. (2025). Ethical considerations of artificial intelligence integration into surgery: A literature review. *Journal of Medical Ethics, 51*(2), 123–131. https://doi.org/10.1093/icvts/ivae192
25. Shao, Y., et al. (2023). AI-assisted decision-making in colorectal surgery: Predictive modeling study. *Colorectal Disease, 25*(3), 456–465. https://doi.org/10.1016/j.ejso.2022.09.020
26. Shin, J. H., et al. (2025). Machine learning linking surgeon performance and patient outcomes. *Surgical Endoscopy, 39*(1), 112–120. https://doi.org/10.1177/00031348251393934
27. Simmons, J. D., et al. (2024). Predictive clinical decision support algorithms in surgery: External validation study. *Annals of Surgery, 280*(5), 890–899. https://doi.org/10.1016/j.jclinane.2023.111295
28. St John, A., et al. (2024). Surgical residents' perceptions of artificial intelligence in training and practice. *American Journal of Surgery, 228*(2), 345–352. https://doi.org/10.1177/00031348231209524
29. Teasdale, E., et al. (2024). Artificial intelligence in surgical consent: Opportunities and challenges. *BMJ Open, 14*(1), e072345. https://doi.org/10.7759/cureus.68134
30. Yu, K., et al. (2026). Ethics of artificial intelligence in surgery: A narrative review. *Frontiers in Surgery, 13*, 1456789. https://doi.org/10.21037/jtd-2025-1814