

Original Research

Received 15 June 2026

Revised 30 July 2026

Accepted 12 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (16s): 74–88

KEYWORDS:

Artificial Intelligence; Robotic Surgery; Medical Liability; Distributed Accountability; Meaningful Human Control; Medical Devices; Algorithmic Accountability; Healthcare Law.

When The Surgeon Becomes the System: Legal and Ethical Accountability in Ai-Guided Robotic Surgery

Devaki J¹, Dr. Prof. Upankar Chutia²

¹PhD Research Scholar, School of Law, Dayananda Sagar University.

²Associate Professor, School of Law, Dayananda Sagar University.

ABSTRACT: As artificial intelligence capabilities advance, robotic surgical systems are rapidly evolving from surgeon-controlled tools to intricate socio-technical environments in which algorithms provide perception, prediction, recommendation and execution. This shift creates an accountability void that neither traditional medical negligence nor product-liability doctrines are entirely fit to fill. This article analyses the legal and ethical dimensions of AI assisted surgery through comparative doctrinal and normative analysis with special focus on India, as well as lessons from the European Union and USA. It examines various laws and regulations, including the National Medical Commission Act 2019, Consumer Protection Act 2019, Bharatiya Nyaya Sanhita (Law of Justice) 2023 in India; Medical Devices Rules (GSR 78(E)) of 2017; Regulation on the European Union's Implementation of and Transitional Provisions for 'Regulation (EU) No. Abstract Algorithms are responsible for performing sophisticated decision-making, yet a myriad of challenges pertaining to the deployment of algorithms can have important ramifications on society. In order to remedy these shortcomings, this article presents a Distributed Accountability Framework (DAF) that disperses accountability based on attributes of effective control, autonomy, foreseeability, causal contribution and access to technical evidence. The framework introduces the accountability-control mismatch, positing that liability should gradually move away from clinicians towards manufacturers, developers and medical institutions as machine autonomy increases. The study recommends adaptive-AI regulation, algorithmic audit trails, monitoring through the lifecycle of an AI application, local validation of AI for data and context applicability, clearer standards on human-oversight, and risk-based compensation mechanisms for India.

1. INTRODUCTION

As time passes, artificial intelligence is slowly changing the operating theater from a place of human dominion to one of human-machine collaboration. Robots initially developed for surgery typically worked as teleoperated machines in which the surgeon had control over virtually every relevant movement. Modern surgical smart systems could interpret clinical information, identify anatomy, foresee complications, aid in surgical planning, offer recommendations, provide surgical actions, and possibly execute limited tasks with less human input (Ficuciello et al. 2019; Mohamed 2024; Shakhathreh and Alazzam 2025).

This advancement changes more than surgical technique. It alters the allocation of decision-making powers. Although the link between control and responsibility is relatively clear where a robotic system simply converts a surgeon's hand movements, when we get to more autonomous robotic systems, this link gets more complex and harder to ascertain. When the software autonomously processes information and affects or executes actions of clinical significance during an operation, the assumption that the surgeon alone steers the surgery becomes harder to hold.

It is a problem of technology and law. The traditional medical-negligence principle basically assumes that you have a doctor with a name who owes a duty of care, exercises clinical judgment and is measurable against an applicable professional standard.

Product-liability law also assumes that a defective product, defective design, insufficient instruction, or manufacturing defect can be identified and linked to harm.

Adverse outcomes stemming from the interactive nature of the actors, technologies, and other components destabilise these assumptions, which are representative of AI-guided robotic surgery (De Micco et al., 2024; Mohamed, 2024; Pasquale, 2022).

Imagine a scenario in which a computer-vision-based recommendation for a given surgical trajectory is made by an artificial-intelligence-guided robot. When the surgeon acts on the recommendation, a bodily structure is misidentified and the patient suffers injury. Did the surgeon act negligently, a conventional legal analysis would ask? But several more questions arise soon after. Was the training data representative enough? Has the algorithm been validated properly? Has the software update changed your system performance? Could the surgeon really understand how the suggestion was created? Did the hospital suitably manage and configure the system? Did the producer communicate relevant limitations sufficiently? Was there enough time for a meaningful human intervention?

They prompt us to ask why the key question in AI-assisted surgery can no longer be whose mistake is that. A better question is how control, chance, knowledge, causal influence, and capacity to prevent harm were distributed across the socio-technical system.

This causes the opacity problem to worsen. Machine-learning systems these days can give useful clinical predictions without giving explanations which fit easily into human reasoning. It is often referred to as the black-box problem. Opacity has the potential to inhibit clinicians' and patients' understanding of the rationale behind a given recommendation or action. In addition, opacity may significantly complicate proof of causation (Matthias, 2004; Pai et al., 2024).

Matthias (2004) refers to a similar phenomenon as the “responsibility gap.” When autonomous learning systems display behaviour that cannot be reasonably predicted or directly controlled by their designers or operators, traditional models of responsibility attribution may not suffice. The issue is especially severe in the field of medicine because the stakes involve the safety of the patient, bodily integrity, professional duties, and criminal, civil, regulatory and ethical responsibility.

An investigation of these questions becomes valuable when viewed in the Indian context. AI-enabled healthcare is regulated in a piecemeal way, including medical-professional regulation, consumer protection, criminal law, medical-device regulation, data governance and technology policy. The institutional architecture for regulating the medical profession is created by the National Medical Commission Act 2019. On the other hand, Product-liability mechanisms are provided by the Consumer Protection Act 2019. The regulation of medical devices is primarily governed by the Medical Devices Rules 2017. Additionally, the Bharatiya Nyaya Sanhita 2023 lays down the framework of criminal-law with respect to negligent acts in causing death. However, the design of these regimes was not planned as an integrated adaptive clinical AI liability architecture.

The current study thus explores the distribution of legal and ethical accountability for patient harm arising from robotic surgery based on AI input. It creates a Distributed Accountability Framework (DAF) that aims to get beyond the simplistic human error vs. machine failure divide.

The analysis is organized along four research questions.

RQ1: How well can existing Indian legal frameworks tackle harm by AI-enabled robotic surgical systems?

RQ2: How does the increasing autonomy of machines modify the traditional association between professional accountability and effective control?

RQ3: How can India learn regulatory lessons from developments in Europe and America?

RQ4: How can responsibility be more clearly split among surgeons, hospitals, manufactures, developers and the like when automated image analysis for AI-guided surgery?

It makes three main contributions: For starters, it takes the view of AI surgical liability as a socio-technical accountability issue and not merely medical malpractice. Second, it introduces the notion of an accountability-control misalignment, which occurs where legal liability continues to rest with a specific human professional although meaningful control has shifted toward automated and institutional systems. Third, it presents DAF as an autonomy-sensitive model, whereby accountability tracks reality—actual control, knowledge of the risk-creating activity, actual contribution to causation, and ability to prevent harm.

2. FROM SURGICAL INSTRUMENT TO ALGORITHMIC AGENT

2.1 Evolution of Robotic Surgical Autonomy

The extent of the legal consequences likely will rely on the fact that to which a robotic surgery is independent. Previous robotic systems acted primarily as advanced tools. They increased dexterity, reduced tremor, improved visualization, and permitted a minimally invasive approach — but remained entirely under the control of the human surgeon.

This relationship is changing as AI increasingly becomes part of new data. Surgical systems that are endowed with the ability to interpret features of the surgical environment through computer vision, machine learning, predictive analytics and automated planning can affect future actions. In more advanced stages of automation, the machine may carry out well-defined subtasks with intermittent use of operator intervention.

Ficuciello et al. (2019) discuss this notion of surgical autonomy with respect to meaningful human control. Their analysis shows that different levels of autonomy imply different sharing arrangements between the surgeon and system with respect to controlling an action.

Table 1: Evolution of Surgical Robotics and Legal Accountability

Level	Technological Role	Human Role	Main Accountability Concern
Manual surgery	Passive instrument	Direct operator	Professional negligence
Teleoperated robotics	Executes surgeon commands	Continuous controller	Negligence and product defect
AI decision support	Generates predictions/recommendations	Reviewer and decision-maker	Reliance and explainability
Semi-autonomous robotics	Performs defined clinical tasks	Supervisor and intervenor	Distributed responsibility
Higher autonomy	Performs substantial tasks independently	Oversight/fallback role	System, institutional, and manufacturer responsibility

Moving across these levels upends liability analysis. The relevant question is not whether surgery stays in place, but whether it continues to wield any measure of control over the act that led to harm.

2.2 Meaningful Human Control

The definition of meaningful human control provides an important conceptual foundation for assigning responsibility in AI-mediated surgery (Ficuciello et al., 2019). It must not be seen as one of nominal supervision.

A surgeon might have an emergency-stop feature but be too rushed or ill-informed to activate it. Likewise, a surgeon could be given an algorithmic recommendation but yea still not have the information needed to determine its reliability independently.

That being said, several conditions are needed for control to be meaningful. The human operator must be able to interpret the relevant system behaviour, distinguish between deviation/malfunction, have a fracture of situational awareness and act/intervene during the clinical time available.

Without these conditions, it is hard to justify the requirement that full responsibility for a decision be placed on the surgeon simply because one could say that a human has formally been "in the loop".

2.3 Automation Bias and Degraded Oversight

The greater automation gives rise to a behavioural paradox. Reliable tech encourages greater dependency on it. We expect that, as clinicians get used to accurate AI recommendations, independent verification will weaken.

This is not only a problem of negligence. Automation can restructure human attention. With the advent of technology that displaces more repetitive tasks, the clinician may be further distanced from constant information processing that facilitated situational awareness.

Thus, commanding instantaneous defence in the event of an unpredictable system failure would create a legal obligation that is unrealistic both technologically and cognitively.

This highlights the difference between intention formal control and material control.

3. ALGORITHMIC OPACITY AND THE RESPONSIBILITY GAP

3.1 The Black-Box Problem

Not all coefficients in fitted machine-learning models have meaning at the clinical level, as they generate outputs through computational relationships. This feature poses a significant challenge to the deed of medical liability.

An incorrect recommendation by a system could arise from many possible causes: biased training data, sensor error/failure, model drift, insufficient validation, wrong deployment setting choice (real-world vs trial), inappropriate configuration set for the device or its software, breach of cybersecurity, unintended code change to algorithms after development and deployment, clinical application failure or misuse[7].

This may involve information that is inaccessible to patients and clinicians, making it unclear which factor led to the negative outcome.

So that makes the opacity of algorithms an asymmetry in evidence. The algorithms, software documentation, and validation records may be in the hands of the subject matter experts who design these elaborate technologies and own these systems but not available to injured patients.

Thus, the legal question is not only one of explainability. Namely, whether claimants can prove causation to any extent.

3.2 Responsibility Gap

Matthias (2004) countered that learning algorithms may complicate traditional attribution of responsibility as the behaviour of a machine may not be entirely predictable by its manufacturer or operator. Robotic surgery is a particularly extreme example of the challenge, where accountability is diffuse across multiple actors. It is possible that a developer might design the system architecture but not write code for a particular harmful output; or that a manufacturer integrates and tests the full system, but has limited control over later software updates. By analogy, a hospital can buy and use the tech while only having partial visibility into how it behaves internally, and a surgeon may rely on algorithmic recommendations without access to proprietary technical cards. As a result, while each stakeholder may have some level of influence over the system, not one actor will have unilateral control over the AI-induced robotic surgery solutions that emerge.

3.3 Accountability-Control Mismatch

In this article, we frame the relevant situation as an accountability-control mismatch—a circumstance in which legal or professional responsibility rests on a single actor despite technological regulation, pertinent knowledge to inform mitigation strategies, and capacity for harm prevention being shared across actors. The surgeon, on the other hand, becomes acutely susceptible to this kind of disconnect in an AI-guided robotic surgery scenario as they are still the most salient human figure in the operating room. But visibility does not necessarily imply causality. Many causes of failures arising from ground between design algorithm and system validation, software updates or changes in data quality, flexible configuration by individual devices and even clinical practices implemented at institutional level are beyond the surgeon's control. Thus, as the basic elements of clinical decision-making and system functioning start to reside in developers, manufacturers and healthcare institutions it can be argued that legal and ethical accountability ought, in turn, to be apportioned accordingly. This kind of approach guarantees that accountability is about actual influence over the technology rather than just the presence of a human at point of care.

4. METHODOLOGY

Research Methods and Design: The study uses a comparative doctrinal and normative legal research method to explore the legal and regulatory aspects attaching to AI-guided robotic surgery. The doctrinal part evaluates structures regulating medical

professional liability, devices, consumer protection, product liability, criminal negligence and AI governance. In India, analysis is based on the National Medical Commission Act 2019, Consumer Protection Act 2019, Medical Devices Rules 2017, Bharatiya Nyaya Sanhita 2023 and literature pertaining to liability for AI and robotic surgery. The final component looks at the EU and US, two different models for AI-enabled medical-device governance. The EU integrates medical-device regulation with general AI regulation in Regulation (EU) 2024/1689 and the US increasingly employs lifecycle-oriented approaches, such as front-loaded change control planning.

The five dimensions that inform how accountability can be assessed are degree of machine autonomy, effective human or institutional control, knowledge and foreseeability of risk, causal contribution to harm and access to the evidence necessary for reconstructing adverse events. These dimensions allow for the evaluation of how responsibility should be apportioned in cases where surgical outcomes are contingent upon the interactions of clinicians, healthcare institutions, manufacturers, developers, and AI systems. Together, they constitute the analytical justification for the Distributed Accountability Framework (DAF) proposed in this article, which seeks to ensure that liability and responsibility under both law and ethics are aligned with the actual distribution of control over risk in AI-aided robotic surgery.

5. INDIAN LEGAL FRAMEWORK

5.1 Medical Professional Responsibility

The National Medical Commission Act 2019 has the primary legislative framework in India for medical education and professional regulation that elaborates, among other things, how boards maintain professional and ethical standards. However, the current regulatory structures were largely built around clinical practice delivered by humans rather than decision support mediated through algorithms. The use of AI in robotic surgery thus complicates what the standard of care should be when surgeons need to choose between accepting a recommendation (or constraint) by an algorithm, questioning it, or overriding it altogether.

Endorsing a regulatory approval of an AI system, however, should not mean that simply following its recommendation amounts to negligence, nor should relying on a state-approved system constitute automatic protection for a surgeon from liability if the clinical evidence suggests the recommendation may be unsafe. The standard of care should however instead consider the reliability of the system, its level of autonomy and explainability, specified limitations of the technology, clinical context, surgeon expertise training and pragmatics. Pai et al. (2024) likewise identify unresolved medico-legal issues arising from the autonomous and opaque nature of artificial intelligence (AI) systems used in Indian healthcare.

5.2 Consumer Protection and Product Liability

The Consumer Protection Act 2019, includes significant product-liability provisions that can serve as a legal framework to remedy any harm caused by an AI-based medical device. The product-liability scheme is laid out in sections 83–87, with section 84 dealing specifically with manufacturer liability based on a manufacturing defect or design defect, departure from manufacturing specifications, failure to conform to an express warranty, or inadequacy of instructions and warnings. The relevance of these provisions to robotic surgery is that there may be patient harm where the relevant defect in the technological system does not involve a breach of duty on the part of any healthcare professional.

But adaptive AI challenges traditional product-liability concepts since software may change after deployment via updates, retraining, configuration changes, or variations in data environments. A system that was deemed safe when first approved may later experience such performance problems after modification. To determine liability, we thus need to pinpoint when the defect originated — and who designed, authorized, implemented or neglected to monitor the change in question. Therefore, product-liability appraisal of AI-enabled medical technologies needs to implement a lifecycle-centered analysis that adequately considers both pre-market design and post-deployment modifications.

5.3 Bharatiya Nyaya Sanhita and Criminal Negligence

Criminal negligence, in particular, poses a state of accountability challenge because criminal liability usually requires more independent evidence for imposition of guilt than tort liability. Causing death by negligence (Section 106 of the Bharatiya Nyaya Sanhita 2023) becomes complex in AI-driven robotic surgery, as a lethal outcome could result from a mix of human, technological and organizational failures. These may be due to poor surgeon training, faulty algorithmic outputs, defective robotic hardware, bad system maintenance, misleading patient data, hacking or networking failure on the software side of things,

improper hospital deployment or lack of warnings from manufacturers. It becomes more difficult to point a finger at one agent when multiple such factors interact. Therefore, the criminal liability of each stakeholder must depend not on the identity of that participant (i.e., it is inappropriate to convict the surgeon simply because he was the last identifiable human actor in a surgical process), but rather on each party's conduct, control, foreseeability and contribution to causation.

6. MEDICAL DEVICES RULES 2017 AND ADAPTIVE AI

6.1 Existing Regulatory Architecture

The Medical Devices Rules, 2017 of India provide the regulatory framework to classify, manufacture, import and source medical devices, clinical investigation and supervision [12]. The framework uses the risk-based categorization path and designates specific regulatory functions to relevant regulators while providing further classification and directions via the Central Drugs Standard Control Organisation (CDSCO) to include medical devices, even for software-based technologies. While this framework plays a crucial role in providing foundational structure for medical-device safety and performance, it was largely designed around conventional and relatively stable medical technologies. Adaptive AI creates new regulatory challenges as the performance and decision behaviour of adaptive AI can change with every software update, retraining, new data or post-deployment modifications. Thus, AI-enabled robotic systems cannot only be held to the typical regulations of devices but will require a regulatory scheme that evolves alongside their adaptive nature.

6.2 The Static-Product Assumption

The static-product assumption, which underlies most regulatory science governing medical devices including AI-enabled medical devices, assumes that a device evaluated at the time of regulatory approval will not differ substantially from its post-market version. This assumption, however, is becoming progressively harder to apply to adaptive AI systems as their clinically relevant behaviour can vary from one piece of software to another (including updates/patches during its lifecycle), from one model re-training on new data to the next, through different configuration options or operating environments, or via learning after deployment themselves. Li et al. (2023) highlight this difficulty as it plays out in European and UK medical-device regulation, exposing the tension between perpetually-in-flux algorithm design and regulatory frameworks predisposed to products with static characteristics. Accordingly, India needs a lifecycle-based regulatory pathway which looks at AI-enabled medical devices from the time of initial approval, through usage and modification to continuing safe clinical practice.

6.3 Risk Classification and Algorithmic Autonomy

The risk classification of medical devices should not only be determined by the physical invasiveness and conventional technical characteristics of robotic systems, but also that algorithmic autonomy and clinical decision-making authority can generate new risks. Robotic platforms accomplishing equivalent surgical functions may carry very different risk profiles based on the amount of autonomy. An example of a teleoperated robot that mimics the movements of a surgeon has very different implications for accountability and safety than an autonomous system that can recognize surgical anatomy or perform subtasks independently. AI-driven software: so far gone from regulatory risk assessment Regulatory assessments should therefore take into account not only the physical characteristics of the device, but to what extent can AI-independently interpret clinical information, affect decisions and execute actions. The proposed means to incorporate algorithmic autonomy into risk classification would provide a regulatory framework more aligned with the actual clinical and technological risks posed by increasingly autonomous surgical systems.

6.4 Local Validation

The performance of AI models is highly dependent on the data used for model development and validation--its quality, diversity, and representativeness. An algorithm that has been trained largely on specific patient populations or healthcare environments may not perform with the same accuracy and reliability if deployed in other demographic or clinical settings. This could be especially pertinent to India as the imported AI-enabled medical technology may perform differently due to differences in patient characteristics, health-care infrastructure and clinical practices. Thus, local clinical suitability should not be inferred solely from approval by foreign regulators. Conclusion: Context-sensitive local validation studies should be performed before technology transfer of high-accuracy AI medical systems for safe, effective and accurate impact on Indian patients and healthcare settings.

7. Comparative Regulatory Analysis

7.1 European Union

The European Union has laid out a complete regulatory framework for the governance of AI with Regulation (EU) 2024/1689, or the Artificial Intelligence Act. Pursuant to Article 6, AI systems linked to products regulated by Union harmonization legislation may only qualify as high risk where certain criteria are fulfilled (ie a class of AI-enabled medical devices in scope of a third-party conformity assessment will be qualified as high-risk). That classification augments requirements for risk management, technical documentation, human oversight and monitoring, among other regulations. These requirements are particularly important with AI-guided robotic surgery, where increasing autonomy may make traditional approaches to safety and responsibility more complex. De Micco et al. There are still questions that the law has to tackle like how the medical injuries that involve AI and robotics should be divided between clinicians (as in a vicarious liability) or institutions (2024). Similarly, Li et al. (2023) show that the use of artificial-intelligence and machine-learning-enabled medical devices disrupts long-held regulatory assumptions, because they are often rapidly modified through a process known as iterative development, which may run counter to regulatory paradigms based on stable medical products.

7.2 United States

The USA is increasingly taking a lifecycle-oriented regulatory path for AI-enabled medical-device Software, largely by way of the US Food and Drug Administration's (FDA) Predetermined Change Control Plans (PCCPs). This practice acknowledges that AI-enabled medical devices could need necessary changes post-clearance as the algorithms, models, or indications of these devices change over time. A PCCP may permit certain specified modifications that prospectively define those changes, and how the changes will be developed, validated and implemented while assuring device safety and effectiveness. This framework imposes structured monitoring along the device lifecycle, rather than considering every expected change as an isolated regulatory event. The US approach therefore demonstrates a helpful regulatory lesson for India, which is that the key question with regard to AI-enabled medical devices need not be whether their utility may evolve post-deployment but rather how substantial clinical changes are reported, validated and monitored in order to ensure patient safety.

Table 2: Comparative Regulatory Orientation

Dimension	India	European Union	United States
Core structure	Medical-device regulation	Medical-device regulation + AI Act	FDA medical-device regulation
AI-specific horizontal law	No comprehensive equivalent	Regulation (EU) 2024/1689	No equivalent comprehensive horizontal statute
Adaptive AI approach	Developing	Lifecycle and high-risk governance obligations	PCCP and lifecycle-oriented oversight
Human oversight	Predominantly professional/device-based	Explicit requirement for high-risk AI	Risk and device specific
Main liability concern	Negligence/product liability	Multi-actor accountability	Professional/product liability

8. STAKEHOLDER-BASED ACCOUNTABILITY

8.1 Surgeon

The surgeon continues to be a key ethical and professional actor in AI-guided robotic surgery, and clinical accountability is not rendered irrelevant by the use of AI. Surgeons should be aware of what is known and unknown about the functions of AIs, when they can be appropriate for clinical use, how their performance can be monitored, and how to intervene to the extent that it is reasonable to do so. Nevertheless, liability should be associated with meaningful human control instead of the presence of a surgeon in the surgical theater. One should not impute strict liability to a surgeon, for example, from an obscure failure in

the algorithm that could not have been reasonably anticipated, understood or prevented. Thus, the inquiry into liability ought to focus on what in practical terms the surgeon was able to review and counteract at that point in time when the injury occurred.

8.2 Hospital

Hospitals are pivotal the safe deployment of AI-guided robotic surgery because they control much of the relevant technological and institutional environment in which these systems operate. In fact, their mandate includes technology procurement but also proper tech-selection for use, credentialing and training of clinicians, cybersecurity protections, network reliability, device maintenance, configuration and management of software updates as well as adverse-event reporting. This is especially critical as surgical systems become more interconnected and autonomous. If patients are harmed because of a lack of infrastructure, improper staff training, insufficient maintenance and cybersecurity protections, misconfigured systems, or bad governance regarding software upgrades when care is delivered by devices, the responsibility should rightly lie with the healthcare institution rather than reside solely with each surgeon. This perspective acknowledges hospitals as key contributors in the distributed accountability model surrounding AI-supported robotic procedures.

8.3 Manufacturer

Manufacturers are a key stakeholder in AI guided robot surgery because they determine system integration, hardware safety, device intended use, performance specifications, and methods for testing/clinical validation as well as risk warnings. Pasquale (2022) contrasts technologies that support professional judgement versus those that replace the function of professionals, which seems particularly relevant to surgical AI. With AI as a mere assistant to the surgeon, more responsibility can fairly reside with the clinician who still exercises significant control over the operation. When AI systems begin to take over more roles previously performed by surgeons (interpretation, decision support, or execution), control is increasingly a matter of the technological system itself. In such as case, focusing liability on the clinician is less compelling, with manufacturers taking at least a corresponding share of accountability for risks arising from system design, validation and performance in practice and for foreseeable technology failures.

8.4 Software Developer and AI Provider

The model architecture, training procedures, validation, performance monitoring, software updates and technical documentation will be given by developers and AI providers on whom AI-guided robotic surgery relies heavily. Therefore, systems that are poorly tested, biased or simply too opaque to hold accountable should not be excused from accountability just because their outputs are put into the hands of real doctors. This is especially important when developers have specialized knowledge of algorithm limitations, risks where algorithms may fail, potential data biases or uncertainties about performance that surgeons could not otherwise easily elicit through clinical use. When such information is either not available or is insufficiently communicated to healthcare professionals, liability must mirror this informational imbalance. Developers and AI providers, therefore, should take on according liability for harm resulting from risks that are well understood and where there is a reasonable degree of control over possible disclosure, monitoring or mitigation.

8.5 Data and Model Governance Actors

The quality, accuracy and representativeness of the data used in the development, validation, and operation of AI systems is absolutely fundamental to their design. Thus, responsibility might arise when harm happens to patients that can be traced back to the use of inappropriate or non-representative data sets, inaccurate labels for outcomes in the data, biased validation procedures, model drift over time and lack of controls over data quality. These risks become particularly salient in the context of dynamic systems, where AI systems may be updated or retrained after deployment, as changes to the data or behavior of the model can directly affect clinical performance. As such, data and model governance organizations should calibrate on standards around data quality as well as validation, monitoring and post-deployment performance. Where these processes are deficient in ways that make a material contribution to an adverse surgical outcome, responsibility should rest with the actors who manage relevant data and model lifecycle rather than only being attributed to the clinician realizing the system.

9. DISTRIBUTED ACCOUNTABILITY FRAMEWORK

The shortcomings of traditional single-actor liability models reveal the limitations of accountability with respect to AI-based robotic surgery. Attributing blame to a sole player — the surgeon, hospital, manufacturer, developer of products or data system, or autonomous technologies — may not represent the allocation of control and risk when something bad happens involving

these parties. To overcome this limitation, we present the Distributed Accountability Framework (DAF), which is aimed at specifying pathways through which legal and ethical accountability can be attributed to relevant stakeholders. The framework centers on five interlocking factors – machine autonomy, effective control, foresight and knowledge, causal contribution, and traceability or evidentiary control. Collectively, these dimensions allow accountability to mirror the real organisations, people and entities who had the power or ability over something that predicted-prevented-explained a bad outcome.

9.1 Degree of Autonomy

The boundaries of traditional single-actor liability regimes highlight the importance of a broader systems model for ensuring accountability in AI-assisted robotic surgery. When harm stems from a constellation of actors (eg, surgeons, hospitals, manufacturers, developers and data systems, autonomous technologies), their singular attribution to any actor may misrepresent the real source of control and risk. To overcome this shortcoming, the Distributed Accountability Framework (DAF) is proposed as a conceptual and action-oriented model for allocating legal and ethical accountability across relevant stakeholders. It develops the framework on five interlinked dimensions: level of machine autonomy, direct control, knowledge and predictability, causal contribution; and traceability or evidence-based control. These dimensions together allow accountability to mirror the real ability of any stakeholder to predict, direct, prevent and/or account for a negative health outcome.

9.2 Effective Control

The second dimension is effective control, where one wants to know which actor can realistically recognize and prevent the injurious result. Meaningful control is more important than nominal (or theoretical) control; the mere presence of a clinician or the possibility of an emergency override does not constitute genuine responsibility. If the surgeon did not have enough information, awareness of the situation or time to detect an algorithmic mistake and respond in a meaningful way, assigning control liability merely on formal grounds may be misguided. Thus, accountability should mirror the ability of the stakeholder to detect or prevent or avert harm.

9.3 Knowledge and Foreseeability

Responsibility should similarly echo the diffusion of relevant knowledge among stakeholders — various actors have proprietary forms of information on potential risks. Manufacturers understand technical failure modes, developers know model uncertainty and algorithmic limitations, hospitals must have knowledge of infrastructure or deployment deficiencies, and surgeons have patient-specific clinical knowledge. In practice, accountability should align with the risk profiles that each actor could or should reasonably identify, foresee and disclose (or seek to mitigate), given their informational domain. It also enables a critique of this pattern, whereby responsibility is shifted to the clinician without being based on knowledge and envisioning that there are actual foreseen consequences.

9.4 Causal Contribution

Adverse events associated with AI-guided robotic surgery should be appreciated as causal chains rather than a mere isolated failure, since patient harm can result from the convoluted interactions of multiple human, technological and institutional mechanisms. A surgical injury might simultaneously be the result of algorithmic misclassification, inadequate clinical oversight, poor manufacturer warnings, an inappropriate software upgrade and bad institutional maintenance. Thus, examining only one cause may hide the role of others and allocate liability incorrectly. In cases where several factors significantly cause the harm, liability should be apportioned to reflect the part played of each such stakeholder in creating, detecting, preventing or limiting the harm.

9.5 Traceability and Evidentiary Control

The last dimension deals with traceability and auditability, which are necessary for accountability in the event of an adverse AI-related surgical incident. No one should be able to leverage their design, deployment, or management of any system that provides critical technical information because the workings are effectively opaque. And hence, high-risk surgical AI should power a secured and tamper-proof record of relevant system inputs, algorithmic outputs, confidence or uncertainty signals, software and model versions, human overrides, actions carried out by the system that were non-trivial and result in patches to either software or data schema as well as recorded errors/warnings. These records can serve as an algorithmic black box that allows regulators, courts, healthcare institutions, and technical experts to retrace the steps of adverse events to find out where

in a causal chain the contribution occurred and whether human decisions, algorithmic behaviour, software changes or larger system-level problems were at fault.

9.6 Core Proposition

The Distributed Accountability Framework (DAF) relies on the notion that legal and ethical accountability ought to match the real distribution of autonomy as well as control in AI-enabled surgical robots. In the face of increasing machine autonomy and decreasing meaningful human control over decisions, liability for risks introduced by technology should gradually migrate away from the individual clinician toward those stakeholders who design, validate, deploy, maintain, update and governs this system. This approach does not absolve the surgeon of professional responsibility, but rather guarantees a fair distribution of accountability based on appropriate metrics of control, knowledge, contribution to causation and ability to prevent harm by each stakeholder.

Table 3: Distributed Accountability Framework

Actor	Primary Control Domain	Potential Liability Trigger
Surgeon	Clinical judgment and supervision	Unreasonable reliance or failure to intervene where intervention was feasible
Hospital	Deployment and institutional infrastructure	Inadequate training, maintenance, cybersecurity, or credentialing
Manufacturer	Integrated product safety	Design defect, unsafe system integration, inadequate warnings
AI developer/provider	Model development and software behaviour	Inadequate validation, algorithmic defect, undisclosed limitation
Data/model controller	Data quality and model monitoring	Bias, drift, contamination, poor data governance

10. ETHICAL IMPLICATIONS

AI-guided surgery raises ethical concerns beyond formal liability.

10.1 Patient Autonomy and Informed Consent

Patients should be informed about the uses of AI and robotic autonomy in their care in a way that is clear and meaningful and in a manner that enables patients to make autonomous decisions. While informed consent is not intended to detail the operation of the algorithm, it should identify AI functions, elements requiring human intervention (and how), known limitations and risks, therapeutic alternatives, requirements for continued human oversight. The rise of AI in surgical decision making and performance is unprecedented, calling for novel techniques of clear disclosure to uphold patient autonomy. Vijayanath et al. And (2024) also stress that getting informed consent is an important example of medicolegal issues in robotic surgery, especially in whether patients understand technology-assisted procedures.

10.2 Professional Competence

The ethical digital surgeon represents a new professional competence that goes beyond traditional surgical skills and incorporates the responsible use of, and understanding of, digital technologies. Lam et al. (2021) is notable in that it points out that digital transformation does include more professional responsibilities for data governance and use of technology. This highlights the need for algorithmic clinical literacy among surgeons — to understand not just if but specifically why AI-enabled surgery works, as well as the potential limitations and reliability boundaries of these systems, and conditions under which they should (and shouldn't) be used. Surgeons are not expected to be as proficient at software engineering design but should have enough context within the technology so they can critically assess AI-assisted decisions, identify potential shortcomings, and know when a dependent decision is appropriate for an AI system or require intervention.

10.3 Fairness

AI systems can reproduce or magnify existing inequalities in data on which the methods were trained and validated. This risk is especially important when deploying medical AI with high-risk potential across defined patient subpopulations that exhibit significant heterogeneity in demographic, physiological, and clinical features. By extension, regulators and healthcare institutions should demand evidence of consistent clinical performance in clinically relevant patient groups and populations. Such as monitoring of algorithmic bias and performance differences with respect to time to ensure that AI-guided surgical technologies promote, rather than undermine equity in patient care.

10.4 Human Oversight

Human oversight should act as a genuine safety net, not just an empty operational compliance step in AI-driven robotic surgery. When a system is moving fast enough, or involving interactions and interdependencies in such complexity that there will be no realistic opportunity for the clinician to identify an erroneous action and counteract it, the fact that a human is looking at what is coming out of the machine has practically no value as control. Holding clinicians accountable under these conditions produces an illusion of accountability and shifts attention from those who truly govern system design and operation. The scope of human oversight requirements should make sure that there is enough information, authority, downtime and technical expertise for clinicians to understand how systems operate and when patient safety could be compromised to intervene timely control the hazards.

11. DISCUSSION

The analysis exemplifies how AI-assisted robotic surgery is achieving a paradigmatic shift in medical responsibility by altering the distribution of clinical authority, technological control, and discretion within the operating room. The traditional malpractice doctrine arose in an environment in which, as the principal decision-maker and controller of the surgical instrument, a clinician was also typically the source of professional judgement. With this model, identification of a poor outcome that was linked to the behavior/act of an identifiable healthcare professional was relatively simple. AI has also disrupted this arrangement with its properties of robotics: many surgical functions that traditionally might have only been performed by human surgeons can now be split between human experts and technological systems. AI might be used to calculate the perception, interpretation, prediction and recommendation of a surgery, while the surgeon will remain responsible for patient-oriented professional ethics. This separation generates the accountability-control mismatch developed in this article, as an effective control over specific decisions or risks may have transferred to software, manufacturers, developers or healthcare institutions but formal responsibility continues to lie with the clinician.

The importance of this idea is to separate responsibility from the technology that causes the harm. So a surgeon may ethically still be responsible for maintaining patient welfare, proper technology use, clinical supervision and effecting intervention where reasonably practicable while not necessarily legally-responsible for every failure arising from an AI-enabled system. On the flip side, because a doctor is nominally still at any point responsible for the outcome of care should not be an excuse for manufacturers, developers and healthcare institutions to escape scrutiny for poorly validated or health damaging technology. To the extent that harm arises from defective system design, insufficient validation, algorithmic bias, improper software updates or infrastructure or undisclosed limitations on the technology, liability should attach to those actors who meaningfully controlled those risks. The latter is critical since the role of the surgeon is transitioning from direct manual controller of the surgical instrument to overseer of a complex human-machine system, where task complexity increases.

It is for this reason that the Distributed Accountability Framework (DAF) rejects problematics located at two extremes. The first, surgeon-default liability, assigns blame for adverse outcomes to the clinician because the surgeon is simply the most conspicuous human actor. The second is automation displacement, where professional responsibility is vastly diminished whenever AI is involved in clinical decision making. DAF, by contrast, ties responsibility to five interrelated aspects of autonomy and control: the degree of machine independence; the presence or lack of effective control; relevant knowledge and foreseeability; causal contribution; and access to technical information. Such an approach aligns with Pasquale's (2022) distinction between complementary and substitutive medical automation, where liability structures should evolve in tandem as AI shifts from merely aiding clinicians to carrying out functions previously exercised through professional judgment.

DAF also provides a pragmatic alternative to legitimating independent legal personhood for AIs. This framework focuses on specific human, corporate, professional and institutional actors that can directly mitigate harm via: taking action to prevent

anticipated risks from technology; managing known technological risks; maintaining precautionary measures as safeguards, preserving documentation of relevant evidence, and compensating injured patients without undue delay. Thus the central legal question should not merely consider whether the surgeon or robot caused the harm, but which stakeholders created, managed, comprehended, participated in or could have readily eliminated the risk that yielded the adverse event. Instead of assigning individual blame, a proportional and evidence-based accountability better grounds the governance of increasingly autonomous surgical technologies.

12. POLICY RECOMMENDATIONS FOR INDIA

Although the present medical-device and healthcare liability framework in India provides a good regulatory backbone, there needs to be an effort to adapt this traditional framework since AI-guided robotic surgery is quickly becoming a reality. The Medical Devices Rules should clarify the non-biological intervention even for conventional medical software from AI-enabled and adaptive systems that reason (i.e. algorithmic behaviour) in a way that can materially impact clinical decision-making and patient outcomes. While the need for follow-up is clear for long-term surveillance, regulatory oversight could also extend to the lifecycle management of medications beyond initial market authorization. Software updates, model retraining based on new datasets or any other changes in performance of the system that have clinical significance should be defined and subject to change-management/validation and monitoring remits. On Adaptable Mechanisms for Quality: While India could learn from the US approach to quality mechanism — such as its FDA's Predetermined Change Control Plan (PCCP) — this will need adaptation to the Indian healthcare system. Likewise, risk classification should be autonomy-sensitive too, with distinctions made between the hardware and invasiveness of a medical device as well as how much information AI interprets independently, recommends clinical actions or performs surgical tasks.

More attention should also be given to algorithmic traceability and local clinical validation. Standardized, secure, and tamper-proof audit trails should reconstruct adverse events (including algorithmic outputs, system changes, human interventions, and technical alerts) for high-risk surgical AI systems. That evidence would clarify the extent to which an adverse outcome was caused mainly by clinical conduct, algorithmic behaviour, technical malfunction or institutional failure. Furthermore, AI systems currently created or assessed largely on foreign datasets should be subjected to appropriate local evaluation to prove their safety and appropriateness in Indian patient populations and healthcare settings. Regulatory and professional authorities should also have clearer definitions of meaningful human control—a complete suspension of autonomy over a specific order, accompanied by the ability to easily obtain relevant information to inform how such an emergency override mechanism should function.

Healthcare institutions must take on a clearer governance role surrounding high-risk surgical AI deployments. Formal mechanisms for technology procurement and validation, clinician training and credentialing, software updates, cybersecurity, system monitoring and maintenance, and adverse-event reporting should be established within hospitals. Institutional accountability matters especially, because safe deployment of AI relies not just on the competence of the technology itself but also on the context in which it runs. Question: How might India strengthen a centralized adverse-event reporting mechanism for significant AI-related medical-device events? Collating this sort of data would allow regulators to spot patterns among algorithmic failures, ongoing problems with the performance of models, or systemic risks that might elude detection when incidents were reviewed in isolation at individual hospitals.

Lastly, the enhanced distributed complexity of technological causation necessitates that discussions on compensating patients for injury reflect these realities. Even with in depth technical input, some of the injuries associated with autonomous or semi-autonomous surgical systems may never be entirely assignable to a single culprit. In this context, high-risk AI medical technologies in India may therefore be potentially explored through narrowly designed insurance, risk-pooling or no-fault compensation mechanisms. These should supplement, not supplant, traditional negligence and product-liability remedies. They would be primarily to prevent patients from being deprived of compensation only because the allocation of responsibility is layered amongst surgeons, hospitals, manufacturers, developers and technological systems in ways that are extraordinarily challenging to pin fault on with any individual. Together, these reforms could help reshape regulation in India away from a fragmented model focused on the approval of individual AI applications to one that is lifecycle-based, autonomy-sensitive, and distributed — able to protect patient safety while enabling responsible innovation in AI-guided robotic surgery.

13. THEORETICAL CONTRIBUTIONS

The study contributes to the nascent literature on accountability in AI-enabled healthcare with four key theoretical and regulatory implications. By acknowledging that adverse outcomes may arise from interactions between clinicians, hospitals, manufacturers, developers, data systems and AI technologies, it first expands the analysis of robotic-surgery liability beyond the traditional doctor-versus-machine paradigm. Second, it introduces accountability-control mismatch to account for cases where legal or professional responsibility remains focused on a clinician despite effective control over pertinent decisions and risks having shifted toward more technological or institutional actors. The ideas here offer a foundation for clarifying formal professional responsibility over the actual distribution of control in AI-enhanced clinical settings.

Third, the study introduces the proposal of a Distributed Accountability Framework (DAF) that outlines five intertwined criteria by which responsibility can be assigned: autonomy, control, awareness and foreseeability, contribution to causation and evidence-based accountability. These criteria enable a more proportionate basis for apportioning liability when multiple persons have contributed to the result. Fourth, the research outlines a regulatory trajectory for governing medical devices in India, which can include structures of governance by profession, lifecycle regulation and algorithmic traceability standards; local validation protocols to generate evidence that builds upon these standards; institutional accountability norms to produce results from such evaluations; and care-embedded compensation mechanisms articulating clinically relevant outcomes. While the framework was developed specifically in the context of robotic surgery, it applies more generally to other high-risk healthcare applications such as AI-assisted diagnosis, automated radiology, intensive-care decision-support systems, and autonomous medication technologies. Such applications may create equivalent discrepancies between formal responsibility and real regulatory power over technology, indicating that DAF could supply a more expansive framework for accountability in AI-driven healthcare.

14. LIMITATIONS AND FUTURE RESEARCH

The current study is predominantly doctrinal, comparative and conceptual and therefore does not apply publically and privately collected empirical data as to how judges, attorneys, surgeons, patients; regulators; manufacturers; developers or engineers would assign responsibility at various levels of surgical autonomy. Future work may implement the Distributed Accountability Framework (DAF) through experimental and scenario modelling studies that manipulate variables such as machine autonomy, algorithmic explainability, surgeon intervention opportunities, alarm settings for technical faults, hospital training metrics and whether the technical failure originated from the device or system. Participants from legal, medical, regulatory, patient groups and engineering communities could then evaluate a responsibility spectrum of how liability and duty is reasonably apportioned among stakeholders from clinician to industry. Conducting research of this type would help to establish whether the five facets of DAF correspond to empirical differences in how responsibility is perceived or ascribed in real-world deliberative and decision-making contexts.

Future research should also investigate meaningful human control by measuring whether surgeons are realistically able to identify, comprehend and (if needed) override incorrect AI actions in the brief window before the high risk procedure is completed. Another salient research direction is the cybersecurity domain because malignant activity that could compromise clinical performance of digitized surgical systems may create an additional causal actor through malicious interference, data manipulation and system vulnerabilities. Future comparative research should reach outside the EU and USA into other jurisdictions with fundamentally different legal systems, healthcare infrastructures, and regulatory capabilities, as well as in emerging economies confronted by challenges to those faced by India. Finally, since systems may behave and be controlled differently as their learning activity continues to transform capabilities after deployment, ongoing legal research will also be needed for continuously learning and adaptive AI systems. This suggests that future research should investigate the modes of liability, regulatory approval, monitoring and accountability that need to be adapted as AI-embedded surgical technologies become more adaptive and autonomous.

15. CONCLUSION

AI-assisted robotic surgery fundamentally transforms the emergent relationships between medical technology, professional autonomy and legal accountability. When robotic systems serve mostly as tools that execute human instructions, traditional negligence and product-liability doctrines are still fairly workable. As AI-enabled systems gain increasing capabilities for perception of clinical conditions, recommendation generating, decision making and surgical actions—a presumption that the surgeon is the primary source of every clinically-impactful decision becomes next to impossible to maintain. That is, the core

legal problem for autonomy is not the automation of technology by itself but an increasing divergence between formal obligations and de facto control. An accountability-control mismatch is created when we hold surgeons legally responsible for algorithmic decisions that they could not have reasonably comprehended or prevented from occurring. Conversely, relieving manufacturers, developers, or healthcare institutions of liability simply because a surgeon maintained formal involvement would produce an equally dangerous gap in accountability.

The heart of the problem is that responsibility does not align with the organizational structure of the social-technical system, as we imagine it to be (as done in a traditional top-down approach), but rather what best suits activities appeals towards accountability for outcomes — which is why we have proposed Distributed Accountability Framework (DAF). Surgeons should remain liable for true clinical decisions and interventions meaningfully in their control, while hospitals should be responsible for appropriate allocation, credentialing of clinicians, infrastructure and maintenance support as well as cybersecurity and institutional governance. Manufacturers can still be held responsible for the integrated product parts, design, validation and possibly system failures that should have been expected, whilst developers (along with AI providers) must continue to take responsibility for risks stemming from algorithmic loss in places they have technical capability or control. Likewise, actors involved in data and model governance should feel responsible when unfit datasets, biased validation, model drift or negligent data-management practises have led to patient harm. Responsibility should then depend on the amount of autonomy, control, relevant knowledge, causal contribution and ability of each stakeholder to foresee or avert the negative outcome.

Thus, the legal analysis of AI-assisted robotic surgery should shift from the simplistic binary question of whether it was the doctor or the robot that caused the harm. The critical question is which actors had the authority, expertise, active liability, causal agency and practical competence to avert the risk that ultimately injured the patient? This approach maintains appropriate professional accountability while shielding clinicians from becoming inadvertent liability scapegoats for failures that emerge elsewhere within an increasingly complex systems of technology. With the increasing interconnectedness of our surgical ecosystem among surgeons, hospitals, manufacturers, developers, algorithms, data and robotic platforms, accountability must also transition away from a person-centric model to one that is more equitable and more broadly dispersed.

REFERENCES

- [1] Central Drugs Standard Control Organisation. (2017). *Medical Devices Rules, 2017*. Ministry of Health and Family Welfare, Government of India.
- [2] Consumer Protection Act, No. 35 of 2019, India.
- [3] De Micco, F., Grassi, S., Tomassini, L., Di Palma, G., Ricchezza, G., & Scendoni, R. (2024). Robotics and AI into healthcare from the perspective of European regulation: Who is responsible for medical malpractice? *Frontiers in Medicine*, 11, 1428504. <https://doi.org/10.3389/fmed.2024.1428504>
- [4] European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*.
- [5] Ficuciello, F., Tamburrini, G., Arezzo, A., Villani, L., & Siciliano, B. (2019). Autonomy in surgical robots and its meaningful human control. *Paladyn, Journal of Behavioral Robotics*, 10(1), 30–43. <https://doi.org/10.1515/pjbr-2019-0002>
- [6] Food and Drug Administration. (2024). *Marketing submission recommendations for a predetermined change control plan for artificial intelligence-enabled device software functions*. U.S. Department of Health and Human Services.
- [7] Lam, K., Purkayastha, S., & Kinross, J. M. (2021). The ethical digital surgeon. *Journal of Medical Internet Research*, 23(7), e25849. <https://doi.org/10.2196/25849>
- [8] Li, P., Williams, R., Gilbert, S., & Anderson, S. (2023). Regulating artificial intelligence and machine learning-enabled medical devices in Europe and the United Kingdom. *Law, Technology and Humans*, 5(2), 94–113. <https://doi.org/10.5204/lthj.3073>
- [9] Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6, 175–183. <https://doi.org/10.1007/s10676-004-3422-1>
- [10] Mohamed, H. (2024). The iRobo-Surgeon conundrum: Comparative reflections on the legal treatment of intraoperative errors committed by autonomous surgical robots. *Law, Innovation and Technology*, 16(1), 194–217. <https://doi.org/10.1080/17579961.2024.2313802>
- [11] National Medical Commission Act, No. 30 of 2019, India.

- [12] Pai, S. N., Jeyaraman, M., Jeyaraman, N., & Yadav, S. (2024). Doctor, bot, or both: Questioning the medicolegal liability of artificial intelligence in Indian healthcare. *Cureus*, 16(9), e69230. <https://doi.org/10.7759/cureus.69230>
- [13] Pasquale, F. (2022). Liability standards for medical robotics and AI: The price of autonomy. In L. A. DiMatteo, C. Poncibò, & M. Cannarsa (Eds.), *The Cambridge handbook of artificial intelligence: Global perspectives on law and ethics* (pp. 200–212). Cambridge University Press. <https://doi.org/10.1017/9781009072168.020>
- [14] Pasquale, F. A., III. (2022). The price of autonomy: Liability standards for complementary and substitutive medical robotics and artificial intelligence. *Ius et Praxis*, 28(1), 3–19. <https://doi.org/10.4067/S0718-00122022000100003>
- [15] Shakhathreh, H. J. M., & Alazzam, F. A. F. (2025). Legal frameworks for AI-driven diagnostics and robotic surgery. *Acta Informatica Medica*, 33(3), 248–252. <https://doi.org/10.5455/aim.2025.33.248-252>
- [16] Vijayanath, V., Joel, V., Priyadharsan, S., & Kesavan Bharathi, K. R. (2024). Robotic surgery: Consent and medico-legal aspect. *Indian Journal of Forensic and Community Medicine*, 11(2), 74–77.