



ARTICLE



<https://doi.org/10.1057/s41599-025-06455-7>

OPEN

When the action is “hybrid”–ethical challenges of the emerging technologies in the operating room

Luca Valera^{1,2}, María Jesús Irarrázaval³ & Mauricio Gabrielli³✉

The paper addresses some ethical issues arising from the interaction between surgeons and emerging technologies in the Operating Room (OR). We argue that introducing new technologies capable of performing some functions performed by the surgeon “autonomously” may transform the doctor-patient relationship, which is no longer direct, but “mediated” by autonomous devices. On the one hand, the patient is inserted in a “hybrid” environment, in which decision-making is hybrid as well: many of the actions take place almost independently of the presence (or action) of a human being. On the other, the surgeon is constantly under observation and his/her responsibility may be partially reduced. All of this implies new ethical reflections and challenges to the healthcare environment. To explain these changes and the need for a new ethical outlook, we consider three cases where emerging technologies have a crucial and sometimes controversial role in the OR: (1) the case of audio and panoramic video recording in the OR; (2) the case of the impact of artificial intelligence on surgical decision making; (3) the case of robotic-assisted tele-surgery. The new actions carried out through emerging technologies in the OR imply, thus, a change in our ethical outlook. To do so, we must first rethink the paradigm of technology we are using: following the post-phenomenological reflections, we argue that technologies are not just the “means” we use but the environments we interact with. This entails the idea of a “hybrid” responsibility and, therefore, a different assessment of the possible consequences of the actions.

¹Department of Philosophy, Universidad de Valladolid, Valladolid, Spain. ²Centre for Bioethics, Pontificia Universidad Católica de Chile, Santiago, Chile.

³Department of Digestive Surgery, Pontificia Universidad Católica de Chile, Santiago, Chile. ✉email: mgabrielli@ucchristus.cl

Introduction. Emerging technologies in the operating room and new ethical challenges

The introduction of highly complex technologies in the Operating Room (OR) is generating high expectations at the technical level (Andras et al., 2020) but also raising some ethical questions (Datteri and Tamburrini, 2009) concerning both the use of these technologies and the new scenarios they are generating. Moreover, we must consider that the use of robotic surgery devices is becoming increasingly frequent, with more than one million surgeries performed with robots each year (Goldberg, 2023). Such quantitative growth could imply a qualitative change in many aspects of medical practice.

Indeed, these devices are challenging our traditional ethics since they concern new conceptions of “autonomy, responsibility, [and] distributive justice” (Datteri and Tamburrini, 2009). Even though this is not new, as Ewing et al. pointed out, during the last few years something has changed: these technologies are not simply “instruments to enhance the performance of and extend the capabilities of the hand” (Ewing et al., 2004), but something that, in most cases, goes (or should go) beyond human control (Ewing et al., 2004). This paper will discuss some ethical challenges and problems emerging from these technologies in the OR. Moreover, we will outline a possible hermeneutic paradigm to interpret the interaction with these devices, which can substantially change the physician/patient relationship.

To do so, we need a preliminary classification of those devices in order to offer more straightforward ethical considerations. Thus, we may use two kinds of classifications: (1) a classification based on their use; and (2) a categorization centered on their independence from human interventions. The first one is more readily accepted, as it depends on the context of their use and the main aim of the medical team. Following this taxonomy, we may identify robotic devices for surgery, diagnosis, rehabilitation, prosthetics, assistance to disabled and elderly people, and so on (Andras et al., 2020). Concerning the second criterion, which relies on the functions implemented by the robots regardless of human intervention, we can distinguish three main categories: (1) controlled systems (i.e., those entirely depending on human actions and translating them into precise movements); (2) semi-automatic systems (i.e., those constraining human movements); and (3) automatic systems (i.e., those performing an activity directly after being programmed by a human operator) (Moustris et al., 2011). In this paper, we will mainly focus on robots used for surgical purposes in the OR: two main kinds of robotic surgeries emerge, that is, “robotic-assisted surgery” (via controlled and semi-automatic systems) and “autonomous robotic surgery” (via automatic systems) (O’Sullivan et al., 2019). A distinction is needed here: autonomous is not the same as automatic. Indeed, “automatic behaviors are completely predictable, as they follow well-established theories, either deterministic or probabilistic. [...] An autonomous system, by contrast, is able to make large adaptations to a change in external conditions by planning its tasks” (Attanasio et al., 2021, 652). In this sense, autonomous systems come significantly closer to human capabilities, while automatic systems do not.

Anyway, a preliminary definition is needed to offer a proper ethical evaluation: what are we talking about when we refer to “emerging technologies”? Even more, concerning the subject of this paper, what features do surgical robots have, and what kind of relationship can we engage with them? In brief, we are wondering about the technology paradigm we are employing and, consequently, whether anything changes in the OR through the introduction of these technologies. We will try to answer these questions in the following sections.

In the following, we will introduce some preliminary ethical challenges concerning robotic surgery (“Surgical robots and emerging technologies in the OR: Some relevant changes”),

presenting three cases where the ethical assessment is particularly needed (“APVR in the OR”–“The Robotic-Assisted Surgery (RAS)”; finally, we will outline some considerations about the concept of autonomy it is currently used, as well as the paradigm of technology employed, and introduce a possible new one, which would allow us to address the current ethical challenges and pose new questions about future scenarios (“Emerging technologies are our environment” and “What kind of autonomy?”).

This paper advances three central claims. First, operating room (OR) technologies should be conceptualized as environments that actively shape human actions and relationships, rather than as mere tools. Second, this reconceptualization necessitates viewing responsibility in the OR as hybrid or distributed across multiple human and non-human agents, rather than attributing it solely to individuals. Third, implementing this perspective in practice requires the development of governance mechanisms capable of addressing the ethical and legal complexities of technologically mediated medicine. In advancing these claims, we build on but go beyond the frameworks of Floridi and Verbeek by operationalizing the notion of technological environments and linking it to concrete regulatory strategies.

Finally, a methodological note is needed. This paper adopts a conceptual ethics approach. Rather than presenting empirical data, it develops a normative analysis informed by selected illustrative cases (e.g., robotic-assisted surgery, audio/panoramic video recording, and AI-based decision-support systems). This perspective enables us to clarify key conceptual distinctions, identify the primary ethical and legal challenges raised by these technologies, and propose ethical frameworks for their responsible implementation.

Surgical robots and emerging technologies in the OR: some relevant changes

When we outline an ethical assessment of medical practice –e.g., a surgery– we usually consider different features of the action: the physician’s aim; the patient’s aim; the circumstances (i.e., the means used, place, time, the possible consequences...); and the action itself. All these features concern the “old” idea of medicine, where the physician/patient relationship was direct and “immediate.” However, the issue completely changes if we consider modern medicine –or at least a part of it, i.e., High-Tech Medicine. The means used to develop an action (e.g., a surgery) is no longer just an “unresponsive” tool; in this sense, it cannot be entirely dependent on humans. While current surgical robots are strictly teleoperated by physicians and do not act independently, scholars have pointed out that future developments may enable them to “act autonomously,” at least in certain circumscribed tasks (Capelli et al., 2023). In this prospective sense, one can imagine surgical actions performed cooperatively by two agents: the physician and the robot. Such scenarios allow us to speak of human “mediated” actions (and relationships) (Verbeek, 2016).¹ This consideration implies at least three consequences, which open up three different scenarios.

First, the paradigm shift described above infers a new form of ethically assessing human responsibilities concerning the consequences of the actions (in our case, the surgery): “What happens if an autonomous robot commits a surgical error? Many people could be held responsible in a court of law, but who should be?” (O’Sullivan et al., 2019). The idea of hybrid (or distributed) actions and responsibilities clearly emerges (Floridi, 2015). We are not assuming that robots are moral agents, yet: we are only stating that the (partial) independence of these devices generates a complex scenario where moral agency itself should be “understood as a fundamentally hybrid affair” (Verbeek, 2014;

Valera, 2022), thus generating an “intricate web of reciprocity” instead of a “linear chain” of causality (Jonas, 1979). In this regard, the physician/patient relationship changes dramatically since the patient is simultaneously relating with the robot and the medical team. At most, we may state that the robot is a significant part of the medical team itself. Here, we are concerned with semi-autonomous robots (Ficuciello et al., 2019).

A second related point is the effective independence of the robots, as suggested. Due to Machine Learning (ML), these devices may evade human control, emancipating them from the master/slave control mode, which enables surgeons to control the entire procedure (Ficuciello et al., 2019). In this regard, the physician may (partially) lose control over the robot, which should create and suggest new opportunities and alternative procedures. In this case, the autonomous robot is the main subject of the surgical procedure, and the surgeon is quite irrelevant in the course of action. As Ma et al. (2020, Fig. 1) suggest, at present, this is only a hypothetical scenario, even though “the intersection of ML and robotics-derived ‘big data’ is a rapidly evolving area of study” (Ma et al. 2020), implying quick changes and further developments.

Lastly, the surgeon assumes a new role in the OR. Not only does he/she cooperate with a technological device—e.g., the robot—but he/she can also be controlled and supervised by the devices themselves. This is the case of Audio and Panoramic Video Recording (APVR), where the medical team is constantly surveilled by cameras and microphones. This has some benefits, obviously, but it also creates ethical and legal concerns (Gabrielli et al., 2021), mainly concerning the ethics of surveillance and the legal features regarding Big Data. Indeed, the AI society can be defined as the panoptic society (Elliott, 2021).

All these challenges are transfiguring the OR environment and the different relationships between the individuals there. Both the patients and the surgeons are involved in a hybrid environment characterized by multiple interactions. Decision-making is hybrid as well: most actions (i.e., assisted surgical procedures, image analysis for navigation, evaluation, diagnosis, or treatment decisions) take place almost independently of the presence (or action) of a human being. This fact implies some emerging ethical challenges. To show them, we introduce three examples of technologies embedded in the OR.

Operational mapping of cases, autonomy levels, and governance. Before doing this, it is worth showing very briefly how current legal frameworks are adapting to surgical robotics and AI. A need for new legislation in this field is emerging, indeed. For example, according to Biasin et al., (2024), the EU’s Medical Device Regulation (MDR) and forthcoming AI Act impose new requirements on AI-based medical systems, focusing on safety, transparency, and human oversight (Ebers, 2024). The complementary EU Machinery Regulation also ensures conformity for robotic machinery (Aboy et al., 2024; Mahler, 2024). In the U.S., Lee et al. (2024) classify FDA-approved surgical robots under the LASR framework, highlighting that most remain at Level 1 (robot assistance), though conditional autonomy is emerging. Moreover, the IDEAL framework, as described by Marcus et al. (2024), provides a staged approach to surgical innovation, accounting for learning curves and long-term outcomes (McCulloch et al., 2024). Yet, as Ludvigsen and Nagaraja (2022) argue, liability laws lag behind, especially regarding cyberattacks and software failures, which blur the line between safety and security. Together, these legal and ethical frameworks shape a complex but evolving regulatory landscape we should consider: indeed, emerging policies must continue adapting alongside these disruptive technological advances.

APVR in the OR

Over the past ten years, APVR in the OR has gained public and industry attention. Surgeries can now be watched later by the surgical patient and used for trainees’ education, training, and development (Walsh et al., 2023). The most attractive aspect is the potential benefit of allowing critical analysis of the performance of healthcare personnel in the OR, which grants objective and tailored actions to every individual involved to improve the outcome concerning the quality of care.

In response to the need to make the OR a more open environment, surgical video recordings in conjunction with artificial intelligence enable the creation of a new tool: the Operating Room Black Box (ORBB), a multiport and synchronized analytic platform that continuously records and collects information from the patient and everything that happens in the OR, regarding the staff’s technical and non-technical performance. Most surgical communities nowadays rely on retrospective analysis of self-reported data from morbidity and mortality, and this kind of analysis is limited by recall bias, low compliance, and lack of details (Jung et al., 2020). In contrast, with the ORBB, a team of experts supported by AI programs analyzes and determines the possible concrete situations to improve the behavior of every person involved in the procedure, be it the surgeon, nurse, or any other staff working in the OR (Goldenberg et al., 2017). This will undoubtedly open up a new universe of possibilities in terms of the ethical and legal implications of these technologies: it should be possible for a machine using AI to determine the mistakes made by the surgeon or, on the contrary, the best way for the surgeon to improve his/her performance during the subsequent surgery. To what extent, thus, will a surgeon have to rely on the analysis of his/her behavior performed by the ORBB? What would happen if the ORBB supported by AI and ML determines that a certain action should be performed, which triggers an error that causes the death of a patient? Should we consider the surgeon the only one responsible for that mistake, or could some degree of responsibility be attributed to the machine? (Verbeek, 2016; Fosch-Villaronga, 2023, 568).

The feasibility of the implementation of the ORBB has been addressed previously by other authors, and they have identified possible barriers (Møller et al., 2023). The principal reasons for staff to decline to participate in this project were concerns about data security, deidentification, and legal issues (Møller et al., 2023). There are lots of concerns over the ethical and legal implications of surgical video recording and the resulting data processing (Gabrielli et al., 2021). The legal aspects have been discussed elsewhere, providing legal frameworks, but the ethical implications have been discussed less (Walsh et al., 2023).

For example, when the staff works in the OR, they talk about the patient and their private lives: this aspect has been shown to improve both the work environment and the overall performance of the surgical team. What happens when the entire team knows they are being surveilled while working? How will this affect their performance? Finally, how will this affect the surgeon’s behavior if he/she knows that the patient may request the complete recording of his/her procedure later on? (Gabrielli et al., 2021). This is known as the “observer effect,” which is defined as a change in normal behavior when individuals are aware they are being observed (Walsh et al., 2023). Nevertheless, some authors compared the ability of video and live observation to promote operating room teamwork and determined that video observations may not be as effective as evaluating live performance (Bui et al., 2018). Maybe because OR staff, after the first cases, forget they are being recorded. Indeed, as Gabrielli et al. (2021) argue, “it is known that this effect typically fades with time, as the subjects get used to being observed, especially if the presence of the observer is not directly visible.” In this sense, the

staff should gradually feel comfortable working in a surveilled environment.

Another important question is who will ultimately own the information generated there. Multiple studies have addressed this topic, and there is still no consensus about data ownership. Some authors argue that the patients own their video recording, while others suggest it should be included in the patient's medical records so the ownership will be of the institution where the surgery was done (Gallant et al., 2022; Xiao et al., 2007; Prigoff et al., 2016). A paramount principle should be that data ownership should be explained during the consent process (Walsh et al., 2023), as Thia et al. (2019) and Turnbull et al. (2014) suggest. Moving on, another question that arises is whether patients should be able to view or have a copy of the surgery recording. Patient ownership and/or holding of this material naturally confers similar responsibilities, as these surgical videos contain data belonging to others (OR personnel) and not only to them. It is different from what occurs in medical records, where the information saved is only about the patient (Walsh et al., 2023). In addition, there is a risk of misinterpretation, considering that the standard is reasonable competence (not perfection) and that there is usually no clear line about the acceptable variation in these scenarios, as stated by Walsh et al. (2023).

There is no doubt that APVR in the OR is a handy tool for learning in the healthcare environment and will most likely improve performance and increase patient safety in the OR—the best example of the use of this technology is aviation, where the existence of the black box where all the information that occurred during the flight is recorded, collected, and analyzed, exists since the 60s (Helmreich et al., 1999) and has made this industry one of the safest in the world. However, we must not stop thinking and asking ourselves how much technology can replace our decision-making processes, together with the ethical consequences implied there.

In light of these challenges, we propose several best-practice recommendations for the governance of APVR systems. First, consent procedures should explicitly include information about video and audio recording, specifying their scope and intended uses. Second, data retention should be limited to defined time windows (e.g., 30–90 days) unless longer storage is explicitly justified by clinical, research, or legal needs. Third, access to recordings should be based on role-specific permissions, with patients and staff informed about who can view or use the material. Finally, staff should receive clear and standardized notifications about the presence of recording systems to mitigate the observer effect while fostering transparency and trust.

The impact of AI and ML on the surgical decision-making process

According to Ngiam and Khor (2019), "machine learning is a type of artificial intelligence (AI) that encompasses algorithmic methods that enable machines to solve problems without specific computer programming." It never tires, never loses information, and the speed of analysis can be incredibly fast. The growing popularity of AI across many different industries has attracted venture capital investment of up to \$5 billion in 2016 alone (Bellini et al., 2019).

Most current ML publications report outcomes of preclinical studies and are associated with essential methodological pitfalls: (1) Selection bias: only high-quality images are used to train and validate the algorithms; (2) Overfitting: always using the same data set does not allow accurate prediction on new images; and (3) Lack of independent external validation. However, Computer-Aided Detection (CAD) using ML can be used to make crucial decisions for patients (Arribas et al., 2021). On these topics, legal

and ethical developments are needed. For example, bias may be reduced in surgical decision-making using a well-designed AI system, as Lazcoz and de Miguel (2025) suggest. We don't need more data, indeed (Lazcoz and de Miguel, 2025): we need better strategies, ethical criteria, and laws to select the best data, in order to avoid bias (1) and overfitting (2) and allow independent external validation (3).

The application of ML in surgery can be in the preoperative, intraoperative, and postoperative phases (Sakamoto et al., 2022). In each one, the main objectives of machine learning applications that have been proven to be effective are diagnosis, prognosis prediction, surgical risk stratification; skill level classification, identification of anatomy and intraoperative surgical phase, identification of instruments and surgical gestures; predictions of complications and prognosis prediction (Sakamoto et al., 2022).

For example, AI and ML have been widely used for the endoscopic detection of malignant lesions. It is known that many times (e.g., in gastric cancer), the diagnosis of malignancy is late because the lesions of the gastric mucosa—when they are incipient or small—are difficult to diagnose via the human eye and require a lot of experience and expertise to do so (de Groof et al., 2021). For example, there may be a situation where the endoscopic diagnosis, supported by CAD, suggests a malignant disease in a small lesion; yet, biopsies of the lesion are inconclusive because the sample was too small. In the eyes of expert endoscopists, the lesion does not appear malignant: will we subject the patient to highly invasive surgery, or should we ignore the machine suggestion? There is no doubt that we will increasingly face situations where a computational algorithm supports the decision of a specific treatment, but we cannot be sure of its correctness. Indeed, the case for IBM's Watson for Oncology supercomputer shows that even ML technologies are fallible (McDougall, 2019), and their use in the decision-making process may be quite dubious. Even though Watson does not make any decisions, it makes recommendations based on hypotheses and evidence (Luxton, 2019), and its recommendations may induce the surgeon to make certain decisions, given that their data may look reliable and the physician him/herself is not able to understand the processes at stake (Smith et al., 2024). Indeed, if the physician follows the algorithm, is he/she responsible for that decision (and the possible mistake) or not? The response depends on our idea of responsibility and the paradigm of technology we employ—as we will see in the last section.

Another example is the real-time analysis of laparoscopic videos and automated identification of anatomy in the intraoperative setting. Here, computer vision uses mathematical techniques to analyze visual images as quantifiable features that can be used within a data set to identify statistically meaningful events (Hashimoto et al., 2018). In these situations, a machine can warn the surgeon of an adverse event that occurred during the surgery that could potentially have consequences for the patient. However, we should consider that in many surgeries, small events occur, and the surgeon may believe they are irrelevant or not possible to correct. In these cases, what would happen if a machine registers that an adverse event occurred, no matter how small, and the surgeon did not correct it? Will the surgeon have to explain and justify ignoring the machine's suggestion? Similarly, ML has the potential to predict pathological features in CT and MRI for liver and pancreatic lesions (Hamm et al., 2019; Liu et al., 2021; Luo et al., 2020). Also, ML has been proven to be more effective than TNM classification in predicting overall survival and the need for adjuvant chemotherapy in colorectal and gastric cancer patients (Jiang et al., 2021; Peng et al., 2016).

Although the idea that ML could enable surgical robots to completely evade human control remains hypothetical, ML is already being applied in surgical robots to give these systems

different degrees of autonomy. For example, the TSolution One orthopedic robot can generate patient-specific surgical plans and autonomously execute tasks such as bone milling, with the surgeon observing rather than actively manipulating the instruments (Lee et al., 2024). Similarly, the Smart Tissue Autonomous Robot (STAR) has demonstrated fully autonomous functioning by performing an intestinal anastomosis in preclinical models (Saeidi et al., 2022). Integrating ML into surgical robots promises benefits like enhanced precision, reduced variability in surgical quality, and the potential for more patient-specific surgical planning (Lee et al., 2024). On the other side, increasing robotic autonomy introduces new risks, as control shifts from human operators to algorithms; there is concern about technical unpredictability in novel situations, and ambiguity in accountability if an autonomous decision leads to error.

Medical staff should thus evolve to interpret decision-support tools and offer wisdom to patients and caregivers, ensuring effective and safe integration of intelligent and machine learning decision process tools for both patients and caregivers. In any case, many concerns (as well as practical problems) remain: due both to the difficulty in understanding the ever-changing functionality of artificially intelligent systems and the impossibility of interrogating their reasoning, in most cases, the clinician in charge is not able to assess the information he/she receive (Smith et al., 2024).

The robotic-assisted surgery (RAS)

In 1985, robots were first used to assist surgeons with Computerized Tomography (CT) guided biopsies (Leddy et al., 2010). In the last few years, progress has been impressive, and robots are becoming more and more autonomous. Nowadays, robots can be classified by their characterization as active, semiactive, or passive (as mentioned above). In this section, we will discuss the implications of passive robots, which are the approved robots used on patients regularly.

“Da Vinci” is the most commonly used RAS: under the surgeon’s control, the robot can be used to cut, suture, grasp, and dissect (Meadows, 2002), providing an excellent tridimensional view, more ergonomic for the surgeon, and in some cases improving surgical outcomes. One of the most substantial reasons the FDA panel gave for approving “da Vinci” has been its future potential. However, this kind of technology raises several ethical questions that we must address.

First, the surgeon-patient relationship, particularly in the long-distance telesurgery where the surgeon is out of the control loop: is it a “real relationship”? This problem concerns telemedicine in general, not only “da Vinci.”

The second concern regards the responsibilities involved. Robotic malfunctions, though rare, can occur and may necessitate a change in the planned surgical procedure; mechanical failure or malfunction can rarely even cause patient injury (Pai et al., 2023). So, who is responsible if one of the robotic arms fails and harms the patient: the surgeon, the company, the staff who prepared the robot for the surgery, or the patient who assumed the risk? There are various stakeholders involved, and balancing this legally and attributing responsibility is a whole challenge. There is quite a consensus that the adoption of robotic surgery does not exonerate the surgeon from their legal accountability. Courts of law have traditionally seen the robot as a tool for assisting the surgeon, but still expect the surgeon to be able to use their discretion on the proposed actions and provide human criteria (Pai et al., 2023).

Finally, the third problem concerns external events and agents: who is responsible if the robot is under a cyber-attack (O’Sullivan et al., 2019), and at a crucial moment of the surgery, breaks an important blood vessel, and the patient dies? Judicial systems

have limited experience in assigning liability for errors made by intelligent machines and differentiating between human and machine errors.

Because of the reduced literature and previous verdicts on this matter, an approximation can be made from another area: autonomously driven cars. In legal proceedings on cases involving crashes of self-driven cars, the culpability varies on a case-by-case basis (Pai et al., 2023). They usually consider the level of autonomy of the car and, in most cases, when it is a fully autonomous car, the culpability is mainly attributed to the manufacturer or the legal authorities that issued the license to drive the car (Pai et al., 2023). Nevertheless, we think the case of robotic surgery is a little different because the surgeon is expected to have control and criteria when indicating a robotic surgery or using a robotic platform to operate (Pai et al., 2023).

Despite the possibility of adverse events occurring during RAS due to multiple causes (e.g., Alemzadeh et al., 2016), it is of paramount importance to have emergency protocols for these situations. This is the responsibility of the treating surgeon and OR personnel. As an example, some institutions have developed emergency undocking protocols for specific occurrences, such as life-threatening intraoperative bleeding, anaphylaxis, or cardiac events (Pai et al., 2023).

There is no doubt that robotics represents an impressive advance in surgery. There are still several ethical and legal gaps that must be addressed by experts and authorities worldwide to offer a clear ethical framework on this topic (Clanahan and Awad, 2023). Indeed, RAS introduces specific error types distinct from those in traditional surgery (e.g., Iacovazzo et al., 2023; Chabot et al., 2024), primarily due to its reliance on complex technology and system integration, as we mentioned above. A systematic review on robotic spine surgery identified three main error types: registration errors (60% of failed screws), skiving errors (26.8%), and interference errors (19.5%) (Gautam et al., 2025). Registration errors arise from mismatches between preoperative imaging and intraoperative anatomy; skiving errors result from instrument deviation on bone surfaces; interference errors involve unexpected interactions with soft tissues (Gautam et al., 2025). These kinds of errors are rare in conventional surgery, where tactile feedback and direct visualization are key. Some meta-analyses (e.g., Farivar et al., 2023; Klock et al., 2023; Negrut et al., 2024; Ogihara et al., 2024; Zhang et al., 2024) comparing robotic and conventional laparoscopic surgeries found that RAS may cause more risks than traditional surgeries. Finally, Paul and Pandya (2025) emphasized that although robotic platforms can reduce surgeon fatigue, they introduce new technical failure points, highlighting the importance of specialized training and system familiarity.

Emerging technologies are our environment

The “ancient” (and out-of-date) paradigm used to interpret technologies claimed that they are exclusively useful means (or devices) to achieve human purposes (Valera, 2020). This is not true anymore (Valera, 2022). The examples we put forward—as well as robotics in general—show that such an interpretation is at least deficient, as it fails to grasp the specificity of those devices. In this sense, a new and fresh paradigm is necessary. Among the current philosophical paradigms introduced to interpret emerging technologies, post-phenomenology (Ihde, 1990) seems to be more than adequate, as it stresses that technologies are not just the “means” (or tools) we use but environments we interact with. From this, the idea of “technological mediation” (Verbeek, 2016) as a possible interpretation of human experience emerges: we interact in technological environments more than act *with* technological tools (Valera, 2020).

Indeed, besides the “passive” role of mediating between humans and the world, technologies “actively” structure unprecedented interactions with the world.

The radical change, here, mainly concerns the interpretation of technologies as passive tools (like a scalpel) or as inter-active devices and environments (like “da Vinci” and “OR Black Box”) (Gunkel 2020). Indeed, Chang et al. state: “The average physician is even more ‘plugged-in’ to the modern technological ecosystem, given the use of electronic medical records, decision support tools, and imaging software” (Chang et al., 2020). If technologies are our environment (Valera, 2022), then, a new kind of consideration is necessary. Retaking Jonas’s (1984, 1) reflections, we may argue that since “with certain developments of our powers the nature of human actions has changed” and “ethics is concerned with actions,” we need new ethical considerations, concerned with inter-actions with devices more than with actions on tools. This ethical shift would be impossible without the “ontological shift” concerning these emerging devices. The pivotal point of this ethical assessment is reframing the surgeon’s responsibility—which is now hybrid—due to the reduced range of his/her actions. The question: “Who is responsible for this procedure?” acquires a new meaning now. To put it in another way, how do we fill, thus, the “responsibility gap” (Matthias, 2004) opened by these technologies?

On the one hand, considering robots responsible for a possible mistake would be incorrect, as ML processes, as far as we know at present, do not admit consciousness, which is the *conditio sine qua non* of responsible behaviors. On the other, it seems to be equally incorrect to place/lay all responsibility (or liability and culpability) on the medical team since the range of the action developed by the robot clearly exceeds the human domain. Hence, we should analogically argue that semi-autonomous robots work like living beings, unconsciously moved by internal dynamics and processes (e.g., ML or AI). In this regard, there wouldn’t be room for full responsibility: since responsibility deals with power, conscious actions, and their predictable consequences, robots would have to be considered not responsible. Thus, in most cases, robotic processes are to be thought of as spaces of non-responsible behaviors (beyond the programmer’s responsibility, which must be causally demonstrated), just like natural occurrences. In this sense, and going back to our initial considerations, we may state that “hybrid responsibility” (Matthias, 2004, 94; Gunkel, 2020) mainly means that human responsibility has certain limits and constraints.

Taddeo and Floridi (2018, 751) called it “distributed responsibility” as a consequence of a new form of “distributed agency”: “The effects of decisions or actions based on AI are often the result of countless interactions among many actors, including designers, developers, users, software, and hardware.” This emerging form of responsibility would be, thus, the result of an action developed by different agents and stakeholders simultaneously (Fosch-Villaronga et al., 2021): we cannot identify a direct cause (and agent) for a particular action. Indeed, there are multiple causes and agents (or stakeholders): there is insufficient track to causally link an event to a particular agent.

In the context of the operating room, this notion of hybrid/distributed responsibility becomes particularly salient. For instance, when a surgical robot autonomously adjusts its movements in response to real-time sensor data, responsibility cannot be entirely ascribed to the surgeon, as the robot’s internal processes function beyond direct human oversight. Similarly, when an AI system provides intraoperative recommendations derived from complex algorithms, the medical team cannot be held solely accountable for the outcome, given that the result arises from the interplay of software, hardware, and pre-established parameters shaped by multiple actors. These

scenarios illustrate that responsibility is inherently distributed across a network of human agents and non-human devices, highlighting the limits of individual human accountability in technologically mediated environments.

Finally, more radical considerations concern totally autonomous robotic surgeons enhanced via AI and microrobots, integrally replacing human surgeons: human responsibility may totally disappear there. Anyway, assessing these actions is beyond the scope of the present paper, given the current state of the art in robotic surgery (Attanasio et al., 2021, 653). What we are interested in to date is “to build trust in these technologies by their human counterparts,” “as the role of AI in surgery becomes more prominent with the emergence of autonomous and intelligent robots” (Capelli et al., 2023, 113). This last point allows us to disentangle a little bit the use of the concept of autonomy in robotics linked to AI.

What kind of autonomy?

The most relevant ethical problems concern robotic autonomy (Yang et al., 2017), insofar as they entail more-than-human responsibilities, as we previously highlighted. In this regard, it is worth recalling the six-level classification of robotic autonomy (Fosch-Villaronga et al., 2021; 2023; Opfermann and Krieger, 2023) elaborated by Yang et al. (2017). It ranges from level 0 (no autonomy) to level 5 (full autonomy –i.e., no human needed). A paper by Attanasio et al. (2021) provides some examples of practical applications of these levels² – i.e., 0. No autonomy; (1) Robot assistance; (2) Task autonomy; (3) Conditional autonomy; (4) High autonomy; (5) Full autonomy. In this sense, robot autonomy is inversely proportional to operator interference: if level 0 gathers “tele-operated robots or prosthetic devices that respond to and follow the user’s command” (Yang et al., 2017), in level 5 we are referring to robotic surgeons “that can perform an entire surgery,” which is “currently in the realm of science fiction” (Yang et al., 2017).³ The most used and affordable robots in surgery nowadays range from level 0 to 2, that is to say, assisted surgery (and autonomy). Regarding the more extended and high level of autonomy to date available—i.e., level 2—it seems that such devices are not really autonomous, as the action is developed by the doctor with the help of a device—in this regard, the robot may be considered a tool.

This classification should push us to consider the very meaning of the word “autonomous” and the possible difference with the term “automatic”: what are we claiming when we say that the robot is autonomous?⁴ A philosophical problem emerges here since the word autonomy usually refers to conscious and free human actions. Indeed, the problem we must face is not lexical—i.e., choosing one word or another to describe a certain behavior—but ontological: stating that a behavior is either autonomous, automatic, or independent, is totally different. For example, if we take the following sentence: “Generally, by the word ‘autonomy’ one means that the robot operates on its own so as to perform a specific task” (Moustris et al., 2011, 377)... are we referring to autonomy or independence? Another definition of machine autonomy could be: “The ability of a computer to follow a complex algorithm in response to environmental inputs, independently of real-time human input” (Etzioni and Etzioni, 2016, 149; Formosa, 2021). Once again, it seems that the autonomy these authors are referring to could be better defined as independence—i.e., the capacity to carry out a process without human intervention.

On the contrary, “in the philosophical literature, however, one finds rather more emphasis on the reasons *why* one is acting (i.e., the goals one has chosen to pursue) than on *how* the goals

are achieved. Auto-nomos, being or setting a law to oneself, indicates the importance of self-regulation or self-government. Autonomy is deeply connected to the capacity to act on one's own behalf and make one's own choices, instead of following goals set by other agents" (Haselager, 2005, 519). In this sense—and we agree with this interpretation—autonomy is something more than independence: if the latter merely refers to the human non-interference in some processes, the former requires, moreover, the possibility to have a *know-how* and a *know-why*. More briefly, independence is a pre-condition for autonomy, and autonomy is not included in independence. A different case is that of the concept of "automatic." Indeed, the focus of this concept is on the process itself and its predictability, more than on the agent involved in that process. Automatic systems operate based on pre-defined instructions and perform repetitive tasks without adapting to changing conditions (e.g., an infusion pump that delivers a constant medication dose once programmed). In contrast, autonomous systems can adapt their actions based on real-time data and environmental changes. In this sense, Chiodo (2022) correctly distinguishes autonomy from automation (Chiodo, 2022). This latter points to the dynamism of the process realized, as Attanasio et al. (2021, 652) clearly explain: an "important clarification is the difference between automatic and autonomous behaviors. Automatic behaviors are completely predictable, as they follow well-established theories, either deterministic or probabilistic. Although there are variations of behaviors for an automatic system, these are due to small adaptations of the controller parameters to external conditions. If the variations are too large, an automatic system cannot adapt and consequently fails. An autonomous system, by contrast, is able to make large adaptations to a change in external conditions by planning its tasks. The planning function requires wider domain knowledge and the use of cognitive tools, such as ontologies or logical rules that do not exist within an automatic system." Following this distinction and the previous classification given by Yang et al. (2017), it seems that "the more autonomous medical robots are, the less human oversight is" (Fosch-Villaronga et al., 2021): an increase in robot autonomy seems to imply a decrease in human responsibility and vice versa, while an increase or decrease in automaticity would not change the degree of human responsibility at all.⁵ Nonetheless, rather than implying a simple reduction of human responsibility as robot autonomy increases, it is more accurate to speak of a redistribution of responsibility across multiple layers of accountability. These layers may include the surgeon (who indicates and supervises the procedure), the institution (which ensures training, protocols, and infrastructure), the manufacturer (who guarantees technical reliability and safety), the developers (who design and maintain the algorithms), and the data suppliers (whose datasets influence system performance), just to mention a few. Ethical analysis should therefore focus on how accountability is allocated and shared within this network, rather than searching for individual responsibilities.

Going back to the aforementioned statements on the difference between autonomous and automatic behaviors and the emphasis on the reason for acting, it is safe to argue that "robots may be operating independently—even 'freely'" choosing how to act in order to achieve goals—but the goals they are trying to achieve are still set by human programmers" (Haselager, 2005, 519), or better, without sufficient reasons to take one or another decision.⁶ Once again, it seems that "different conceptions of autonomy" emerge here, due to the different emphasis on "the capacity for independent (unsupervised) action versus the freedom to choose goals" (Haselager, 2005, 528).

Conclusions

This paper has defended three interconnected claims. First, OR technologies should be treated not only as tools but as environments that shape human action and relationships. Building on this, the second claim holds that effective responses require reconceptualizing responsibility in the OR as hybrid or distributed across surgeons, institutions, manufacturers, developers, and data suppliers. These conceptual shifts support a third claim: the possibility of designing governance mechanisms, including oversight requirements, adapted consent models, and emergency safeguards, to address the ethical and legal challenges posed by these technologies. Articulating these claims extends the foundational work of Floridi and Verbeek and translates the concept of technological environments into actionable ethical and regulatory guidance. We have shown how specific cases (robotic-assisted surgery, APVR, AI decision-support, and experimental autonomous systems) can be mapped to autonomy levels, ethical/legal risks, and governance measures. This operational mapping helps clarify how responsibility and accountability can be redistributed rather than diminished as technologies evolve.

The central ethical task ahead is not only to anticipate the capabilities of autonomous systems but also to build trust in them through transparent oversight, fairness in design, and careful allocation of responsibilities. In this sense, surgical ethics must move beyond the physician–patient dyad and address the distributed networks of agency and accountability that now define technologically mediated medicine.

The introduction of emerging technologies and the development of artificial intelligence tools have introduced ethical challenges in the medical (and especially the surgical) field. As suggested by Taddeo and Floridi (2018), regulations (or moratoria) are not enough: ethical and legal frameworks need to be clarified regarding autonomy and liability when talking about these topics. Indeed, as suggested by Yang et al. (2017), "at the higher levels of autonomy (specifically Level 5 and possibly Level 4), the robot is not only a medical device but is also practicing medicine. The FDA regulates medical devices but not the practice of medicine, which is left to the medical societies." These possibilities should bring up new considerations about the agency of surgical robots and surgeons as well. For this reason, in this paper, we suggest some concepts and frameworks to address and reframe the current technological challenges in the field of surgery. In this regard, it is worth noticing that both the concepts of action, autonomy (or independence), and responsibility drastically change in the emerging environments above-mentioned: the first step to ethically assess these topics is to reframe the paradigm we are currently using in the field of technological surgery, to offer new hermeneutics of radically new situations, contexts, and concerns. This is the first challenge we must face in the new surgical era: address the emerging problems through new paradigms, concepts, and points of view. In this sense—and this is the main conclusion in the present paper, summarizing the last two sections—any legal and ethical consideration must stem from ontological considerations. We need new ontologies to define the emerging technologies we mentioned above: this need is necessarily prior to any ethical, political, or legal discussion regarding the "use" of such devices.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its Supplementary material. To write the paper—which is a theoretical reflection—we used only published papers and shared the DOIs of all articles

that had them available. We can share the originals of other articles upon request.

Received: 9 July 2024; Accepted: 18 December 2025;

Published online: 05 February 2026

Notes

- 1 These actions “realized” by surgical robots can be clustered in three main “functional-areas”: 1. acquisition and analysis of information; 2. division of surgical trajectories or plan of actions; 3. execution of the surgery (Fosch-Villaronga 2023, 566).
- 2 For reasons of space, we won’t present examples of surgical robots that exhibit different levels of autonomy, neither we will discuss the ethical implications of each level of autonomy. Examples of this may be found in the abovementioned paper by Attanasio et al. (2021).
- 3 Recent scoping reviews of the literature show that almost 75% of studies refer to Level 1, 23% to Level 2, and only 2% to Level 3 (Vasey et al., 2023, 898).
- 4 It is worth noticing that for many authors “the debate about autonomy of robots is a non-starter” (Haselager, 2005, 517).
- 5 Chiodo (2022) argues that an increase in robot autonomy would lead to an increase in human automation: we do not agree with this perspective, due to our characterization of autonomy.
- 6 This is precisely the aforementioned case for IBM’s Watson for Oncology supercomputer, which needs to be assisted by human beings during the decision-making process since it *actually doesn’t take any decision*.

References

Aboy M, Minssen T, Vayena E (2024) Navigating the EU AI Act: implications for regulated digital medical products. *NPJ Digit Med* 7(1):237. <https://doi.org/10.1038/s41746-024-01232-3>

Alemzadeh H, Raman J, Leveson N, Kalbarmczyk Z, Iyer RK (2016) Adverse events in robotic surgery: a retrospective study of 14 years of FDA data. *PLoS ONE* 11(4):e0151470. <https://doi.org/10.1371/journal.pone.0151470>

Andras I, Mazzone E, van Leeuwen FWB et al (2020) Artificial intelligence and robotics: a combination that is changing the operating room. *World J Urol* 38(10):2359–2366. <https://doi.org/10.1007/s00345-019-03037-6>

Arribas J, Antonelli G, Frazzoni L et al (2021) Standalone performance of artificial intelligence for upper GI neoplasia: a meta-analysis. *Gut* 70:1458–1468. <https://doi.org/10.1136/gutjnl-2020-321922>

Attanasio A, Scaglioni B, De Momi E, Fiorini P, Valdastri P (2021) Autonomy in surgical robotics. *Annu Rev Control Robot Auton Syst* 4(1):651–679. <https://doi.org/10.1146/annurev-control-062420-090543>

Bellini V, Guzzon M, Bigliardi B, Mordinini M, Filippelli S, Bignami E (2019) Artificial intelligence: a new tool in operating room management. Role of machine learning models in operating room optimization. *J Med Syst* 44(1):20. <https://doi.org/10.1007/s10916-019-1512-1>

Biasin E, Kamenjašević E, Ludvigsen KR (2024) Cybersecurity of AI medical devices: risks, legislation, and challenges. In: Solaiman B, Cohen IG (eds.) *Research handbook on health, AI and the Law*. Edward Elgar Publishing Ltd, Cheltenham, pp 57–74. <https://doi.org/10.4337/9781802205657.ch04>

Bui AH, Guerrier S, Feldman DL et al (2018) Is video observation as effective as live observation in improving teamwork in the operating room? *Surgery* 163(6):1191–1196. <https://doi.org/10.1016/j.surg.2018.01.019>

Capelli G, Verdi D, Frigerio I et al (2023) White paper: ethics and trustworthiness of artificial intelligence in clinical surgery. *AIS* 3(2):111–122. <https://doi.org/10.20517/ais.2023.04>

Chabot S, Calleja-Agius J, Horeman T (2024) A comparison of clinical outcomes of robot-assisted and conventional laparoscopic surgery. *Surg Tech Dev* 13:22–57. <https://doi.org/10.3390/std13010003>

Chang M, Canseco JA, Nicholson KJ, Patel N, Vaccaro AR (2020) The role of machine learning in spine surgery: the future is now. *Front Surg* 7:54. <https://doi.org/10.3389/fsurg.2020.00054>

Chiodo S (2022) Human autonomy, technological automation (and reverse). *AI and Soc* 37:39–48. <https://doi.org/10.1007/s00146-021-01149-5>

Clanahan JM, Awad MM (2023) How does robotic-assisted surgery change OR safety culture? *AMA J Ethics* 25(8):E615–E623. <https://doi.org/10.1001/amaethics.2023.615>

Datteri E, Tamburini G (2009) Ethical reflections on health care robotics. In: Capurro R (ed) *Ethics and Robotics*. IOS Press, Amsterdam, pp 35–48

de Groot J, Antonelli G, Dinis-Ribeiro MJ, Bergman JJ (2021) Meta-analyses of machine learning in endoscopy: stacking apples and oranges. *Gastrointest Endosc* 93(5):1016–1018. <https://doi.org/10.1016/j.gie.2021.01.010>

Ebers M (2024) AI robotics in healthcare between the EU medical device regulation and the artificial intelligence act gaps and inconsistencies in the protection of patients and care recipients. *Oslo Law Rev* 11(1):1–12. <https://doi.org/10.18261/olr.11.1.2>

Elliott A (2021) The Routledge Social Science Handbook of AI. Routledge; New York and London

Etzioni A, Etzioni O (2016) AI assisted ethics. *Ethics Inf Technol* 18(2):149–156. <https://doi.org/10.1007/s10676-016-9400-6>

Ewing DR, Pigazzi A, Wang Y, Ballantyne GH (2004) Robots in the operating room: the history. *Semin Laparosc Surg* 11(2):63–71. <https://doi.org/10.1177/107155170410110020>

Farivar D, Kim TT, Sy CA et al (2023) Complications have not improved with newer generation robots. *Global Spine J* 15(2):898–905. <https://doi.org/10.1177/21925682231216081>

Ficuciello F, Tamburini G, Arezzo A, Villani L, Siciliano B (2019) Autonomy in surgical robots and its meaningful human control. *Paladyn J Behav Robot* 10(1):30–43. <https://doi.org/10.1515/pjbr-2019-0002>

Floridi L (2015) The onlife manifesto. *Being Human in a Hyperconnected Era*. Springer, Cham

Formosa P (2021) Robot autonomy vs. Human autonomy: social robots, artificial intelligence (AI), and the nature of autonomy. *Mind Mach* 31:595–616. <https://doi.org/10.1007/s11023-021-09579-2>

Fosch-Villaronga E, Khanna P, Drukarch H, Custers B (2023) The role of humans in surgery automation. Exploring the influence of automation on human–robot interaction and responsibility in surgery innovation. *Int J Soc Rob* 15:563–580. <https://doi.org/10.1007/s12369-022-00875-0>

Fosch-Villaronga E, Khanna P, Drukarch H, Custers BH (2021) A human in the loop in surgery automation. *Nat Mach Intell* 3(5):368–369. <https://doi.org/10.1038/s42256-021-00349-4>

Gabrielli M, Valera L, Barrientos M (2021) Audio and panoramic video recording in the operating room: legal and ethical perspectives. *J Med Ethics* 47:798–802. <https://doi.org/10.1136/medethics-2019-106056>

Gallant JN, Brelsford K, Sharma S, Grantcharov T, Langerman A (2022) Patient perceptions of audio and video recording in the operating room. *Ann Surg* 276(6):e1057–e1063. <https://doi.org/10.1097/SLA.0000000000004759>

Gautam D, Vivekanandan S, Mazur MD (2025) Robotic spine surgery: systematic review of common error types and best practices. *Oper Neurosurg* 28(3):295–302. <https://doi.org/10.1227/ons.0000000000001293>

Goldberg K (2023) Supervised autonomy in robot-assisted surgery and telesurgery. *Science* 381(6654):143

Goldenberg MG, Jung J, Grantcharov TP (2017) Using data to enhance performance and improve quality and safety in surgery. *JAMA Surg* 152(10):972–973. <https://doi.org/10.1001/jamasurg.2017.2888>

Gunkel DJ (2020) Mind the gap: responsible robotics and the problem of responsibility. *Ethics Inf Technol* 22:307–320. <https://doi.org/10.1007/s10676-017-9428-2>

Hamm CA, Wang CJ, Savic LJ et al (2019) Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol* 29(7):3338–3347. <https://doi.org/10.1007/s00330-019-06205-9>

Haselager WFG (2005) Robotics, philosophy and the problems of autonomy. *PRAGMATCOGN* 13(3):515–532. <https://doi.org/10.1075/pc.13.3.07has>

Hashimoto DA, Rosman G, Rus D, Meireles OR (2018) Artificial intelligence in surgery: promises and perils. *Ann Surg* 268(1):70–76. <https://doi.org/10.1097/SLA.0000000000002693>

Helmreich RL, Merritt AC, Wilhelm JA (1999) The evolution of crew resource management training in commercial aviation. *Int J Aviat Psychol* 9(1):19–32. https://doi.org/10.1207/s15327108ijap0901_2

Iacobazzo C, Buonanno P, Massaro M et al (2023) Robot-assisted versus laparoscopic gastrointestinal surgery: a systematic review and metanalysis of intra- and post-operative complications. *J Pers Med* 13:1297. <https://doi.org/10.3390/jpm13091297>

Ihde D (1990) Technology and the lifeworld: from garden to Earth. Indiana University Press, Bloomington

Jiang Y, Jin C, Yu H et al (2021) Development and validation of a deep learning ct signature to predict survival and chemotherapy benefit in gastric cancer: a multicenter, retrospective study. *Ann Surg* 274(6):e1153–e1161. <https://doi.org/10.1097/SLA.0000000000003778>

Jonas H (1979) Toward a philosophy of technology. *Hastings Cent Rep* 9(1):34–43

Jonas H (1984) The imperative of responsibility: In: *Search of an ethics for the technological age*. University of Chicago Press, London and Chicago

Jung JJ, Jüni P, Lebovic G, Grantcharov T (2020) First-year analysis of the operating room black box study. *Ann Surg* 271(1):122–127. <https://doi.org/10.1097/SLA.0000000000002863>

Klock JA, Bremer K, Niu F, Walters RW, Nandipati KC (2023) Robotic-assisted bariatric surgery is associated with increased postoperative complications

compared to laparoscopic: a nationwide readmissions database study. *Obes Surg* 33(7):2186–2193. <https://doi.org/10.1007/s11695-023-06657-5>

Lazcoz G, de Miguel I (2025) Is more data always better? On alternative policies to mitigate bias in Artificial Intelligence health systems. *Bioethics* 39(5):414–424. <https://doi.org/10.1111/bioe.13398>

Leddy L, Lendvay T, Stava R (2010) Robotic surgery: applications and cost effectiveness. *Open Access Surg* 3:99–107

Lee A, Baker TS, Bederson JB, Rapoport BI (2024) Levels of autonomy in FDA-cleared surgical robots: a systematic review. *NPJ Digit Med* 7(1):103. <https://doi.org/10.1038/s41746-024-01102-y>

Liu X, Khalvati F, Namdar K et al (2021) Can machine learning radiomics provide pre-operative differentiation of combined hepatocellular cholangiocarcinoma from hepatocellular carcinoma and cholangiocarcinoma to inform optimal treatment planning? *Eur Radiol* 31(1):244–255. <https://doi.org/10.1007/s00330-020-02711-7>

Ludvigsen KR, Nagaraja S (2022) Dissecting liabilities in adversarial surgical robot failures: a national (Danish) and EU law perspective. *Comput Law Secur Rev* 44: 105656. <https://doi.org/10.1016/j.clsr.2022.105656>

Luo Y, Chen X, Chen J et al (2020) Preoperative prediction of pancreatic neuroendocrine neoplasms grading based on enhanced computed tomography imaging: validation of deep learning with a convolutional neural network. *Neuroendocrinology* 110(5):338–350. <https://doi.org/10.1159/000503291>

Luxton DD (2019) Should Watson be consulted for a second opinion? *AMA J Ethics* 21(2):E131–E137. <https://doi.org/10.1001/ajamethics.2019.131>

Ma R, Vanstrum EB, Lee R, Chen J, Hung AJ (2020) Machine learning in the optimization of robotics in the operative field. *Curr Opin Urol* 30(6):808–816. <https://doi.org/10.1097/MOU.0000000000000816>

Mahler T (2024) Smart robotics in the EU legal framework: the role of the machinery regulation. *Oslo Law Rev* 11(1):1–18. <https://doi.org/10.18261/olr.11.1.5>

Marcus HJ, Ramirez PT, Khan DZ et al (2024) The IDEAL framework for surgical robotics: development, comparative evaluation and long-term monitoring. *Nat Med* 30:61–75. <https://doi.org/10.1038/s41591-023-02732-7>

Matthias A (2004) The responsibility gap: ascribing responsibility for the actions of learning automata. *Ethics Inf Technol* 6(3):175–183. <https://doi.org/10.1007/s10676-004-3422-1>

McCulloch P, Marcus HJ, Hanrahan JG (2024) The IDEAL framework and its applications: balancing innovation with safety. *Bull R Coll Surg Engl* 106(8):492–494. <https://doi.org/10.1308/rcsbull.2024.158>

McDougall RJ (2019) Computer knows best? The need for value-flexibility in medical AI. *J Med Ethics* 45(3):156–160. <https://doi.org/10.1136/medethics-2018-105118>

Meadows M (2002) Robots lend a helping hand to surgeons. *FDA Consum* 36(3):10–15

Møller KE, Sørensen JL, Topperzer MK et al (2023) Implementation of an innovative technology called the OR black box: a feasibility study. *Surg Innov* 30(1):64–72. <https://doi.org/10.1177/15533506221106258>

Moustris GP, Hiridis SC, Deliparaschos KM, Konstantinidis KM (2011) Evolution of autonomous and semi-autonomous robotic surgical systems: a review of the literature. *Int J Med Robot* 7(4):375–392. <https://doi.org/10.1002/rcs.408>

Negrut RL, Cote A, Caus VA, Maghri AM (2024) Systematic review and meta-analysis of laparoscopic versus robotic-assisted surgery for colon cancer: efficacy, safety, and outcomes – a focus on studies from 2020–2024. *Cancers* 16(8):1552. <https://doi.org/10.3390/cancers16081552>

Ngiam KY, Khor IW (2019) Big data and machine learning algorithms for healthcare delivery. *Lancet Oncol* 20(5):e262–e273. [https://doi.org/10.1016/S1470-2045\(19\)30149-4](https://doi.org/10.1016/S1470-2045(19)30149-4)

O'Sullivan S, Nevejans N, Allen C et al (2019) Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *Int J Med Robot* 15(1):e1968. <https://doi.org/10.1002/rcs.1968>

Ogihara A, Omata M, Shidei H et al (2024) Intraoperative robotic surgical system-related problems in robot-assisted thoracoscopic surgery. *Gen Thorac Cardiovasc Surg* 72(9):593–598. <https://doi.org/10.1007/s11748-024-02013-1>

Opfermann JD, Krieger A (2023) Bringing highly autonomous surgical robotics to the clinic. *Science* 381(6654):144–145

Pai SN, Jeyaraman M, Jeyaraman N, Nallakumarasamy A, Yadav S (2023) In the hands of a robot, from the operating room to the courtroom: the medicolegal considerations of robotic surgery. *Cureus* 15(8):e43634. <https://doi.org/10.7759/cureus.43634>

Paul RA, Pandya A (2025) A system for surgeon fatigue monitoring in robotic surgery. *Robotics* 14(4):40. <https://doi.org/10.3390/robotics14040040>

Peng JH, Fang YJ, Li CX et al (2016) A scoring system based on artificial neural networks for predicting 10-year survival in stage II A colon cancer patients after radical surgery. *Oncotarget* 7(16):22939–22947. <https://doi.org/10.1863/oncotarget.8217>

Prigoff JG, Sherwin M, Divino CM (2016) Ethical recommendations for video recording in the operating room. *Ann Surg* 264(1):34–35. <https://doi.org/10.1097/SLA.0000000000001652>

Saeidi H, Opfermann JD, Kam M et al (2022) Autonomous robotic laparoscopic surgery for intestinal anastomosis. *Sci Robot* 7(62):eabj2908. <https://doi.org/10.1126/scirobotics.abj2908>

Sakamoto T, Goto T, Fujiogi M, Kawarai Lefor A (2022) Machine learning in gastrointestinal surgery. *Surg Today* 52(7):995–1007. <https://doi.org/10.1007/s00595-021-02380-9>

Smith H, Birchley G, Ives J (2024) Artificial intelligence in clinical decision-making: Rethinking personal moral responsibility. *Bioethics* 38(1):78–86. <https://doi.org/10.1111/bioe.13222>

Taddeo M, Floridi L (2018) How AI can be a force for good. *Science* 361:751–752. <https://doi.org/10.1126/science.aat5991>

Thia BC, Wong NJ, Sheth SJ (2019) Video recording in ophthalmic surgery. *Surv Ophthalmol* 64(4):570–578. <https://doi.org/10.1016/j.survophthal.2019.01.005>

Turnbull AM, Emsley ES (2014) Video recording of ophthalmic surgery –ethical and legal considerations. *Surv Ophthalmol* 59(5):553–558. <https://doi.org/10.1016/j.survophthal.2014.01.006>

Valera L (2020) New technologies. Rethinking ethics and the environment. In: Valera L, Castilla JC (eds.) *Global changes. Ethics, politics and the environment in the contemporary technological world*. Springer, Cham, pp 29–43

Valera L (2022) Espejos: filosofía y nuevas tecnologías. Herder, Barcelona

Vasey B, Lippert KAN, Khan DZ et al (2023) Intraoperative applications of artificial intelligence in robotic surgery: a scoping review of current development stages and levels of autonomy. *Ann Surg* 278(6):896–903. <https://doi.org/10.1097/SLA.00000000000005700>

Verbeek PP (2014) Some misunderstandings about the moral significance of technology. In: Kroes P, Verbeek PP (eds) *The moral status of technical artefacts*. Springer, Dordrecht pp 75–88

Verbeek PP (2016) Toward a theory of technological mediation. a program for postphenomenological research. Lexington Books, London

Walsh R, Kearns EC, Moynihan A et al (2023) Ethical perspectives on surgical video recording for patients, surgeons and society: systematic review. *BJS Open* 7(3). <https://doi.org/10.1093/bjsopen/zrad063>

Xiao Y, Schimpff S, Mackenzie C et al (2007) Video technology to advance safety in the operating room and perioperative environment. *Surg Innov* 14(1):52–61. <https://doi.org/10.1177/1553350607299777>

Yang GZ, Cambias J, Cleary K et al (2017) Medical robotics –regulatory, ethical, and legal considerations for increasing levels of autonomy. *Sci Robot* 2(4):eam8638. <https://doi.org/10.1126/scirobotics.aam8638>

Zhang Y, Dong B, Li G, Ye W (2024) Short-term outcomes of robotic vs. laparoscopic surgery for rectal cancer after neoadjuvant therapy: a meta-analysis. *Front Surg* 10: 1292031. <https://doi.org/10.3389/fsurg.2023.1292031>

Acknowledgements

We want to thank the Department of Digestive Surgery (Pontificia Universidad Católica de Chile) for its support and the grant Avanza UC n. AV25409.

Author contributions

All the authors contributed equally to the paper.

Competing interests

The authors declare no competing interests.

Ethical approval

Ethical approval was not required as the study did not involve human participants.

Informed consent

Informed consent was not required as the study did not involve human participants.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-025-06455-7>.

Correspondence and requests for materials should be addressed to Mauricio Gabrielli.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026