AI and Ethics
https://doi.org/10.1007/s43681-023-00361-x

ORIGINAL RESEARCH

Is a robot surgeon with AI the ideal surgeon? A philosophical analysis

Sara Patuzzo[1] · **Maurizio Balistreri**[2] · **Tommaso Marinelli**[3] · **Simone Giacopuzzi**[1]

Received: 22 August 2023 / Accepted: 11 October 2023

Abstract
The medical field of surgery has integrated robots with Artificial Intelligence into its procedures. Currently, these machines primarily assist physicians in their activities, but it is plausible that, with ongoing scientific and technological advancements, AI robot surgeons could replace human surgeons in the near future. After providing an overview of the current state of robotic surgery and prospective future developments and scenarios, the paper will focus on the potential difficulties patients may experience in accepting interventions performed by an AI robot surgeon, largely owing to their perception of the robot as non-human. The prevailing concerns that will be analyzed and discussed from a philosophical standpoint include the belief that the AI robotic surgeon is not considered part of the medical team, its perceived incapacity to empathize with patients and to create emotional involvement, and the fear that it might commit severe errors unanticipated by its programming or react inappropriately to adverse events.

Keywords Robotic Surgery · Human-Robot interaction · AI in Medicine · Autonomy in Robotics · AI Decision-Making · Philosophy and AI

1 Introduction

Artificial Intelligence (AI) simulates and refines human logical–analytical and self-adaptive intelligence. Based on the implementation of AI in machines, the field of surgery has begun incorporating robots also with AI into its operative procedures, whose use has demonstrated significant potential in various specific areas [1–4], although in some contexts it has raised concerns and highlighted limitations [5–7]. Currently, these robots assist physicians in their surgical interventions [8, 9], but we can envisage that, thanks to scientific and technological progress, they may eventually supersede human surgeons in the near future. In the face of this scenario, at least three distinct perspectives emerge that warrant consideration: (1) the philosophical, (2) the strictly bioethical, and (3) the legal.

1. It should be noted that there may first and foremost be a reluctance on the part of patients to accept surgeries performed by AI robot surgeons. Key concerns might be tied to the understanding that AI robot surgeons are not of our species and, thus, are not like human physicians; they do not feel emotions and empathy toward us, and might even act unpredictably, making severe, potentially fatal mistakes. Are these fears justified, or are they a resistance, occasionally grounded in prejudice, that can be overcome through a logical–rational analysis of the factors at play? A philosophical investigation into the integration of AI robot surgeons in healthcare can offer reasoned perspectives on this scenario.
2. Following this, the bioethical and legal issues surrounding surgeries conducted by an AI robot surgeon require careful scrutiny. For instance, if we accept these machines into the medical team and let them operate on us, how can we ensure that an AI robot surgeon pursues the patient's well-being, upholding the principle of beneficence? Upon whom and why should the moral responsibility fall in the case of violating the ethical principle of non-maleficence? Who will be legally

✉ Sara Patuzzo
sara.patuzzo@univr.it

[1] Department of Surgery, Dentistry, Paediatrics and Gynaecology, Confortini Center, Civil Hospital, University of Verona, Verona, Italy

[2] Department of Linguistics and Literary, Historical, Philosophical and Legal Studies, University of Tuscia of Viterbo, Viterbo, Italy

[3] Verona, Italy

Published online: 06 November 2023

accountable for mistakes or harmful decisions, and what are the legal implications of the actions of an AI robot surgeon? How do we adapt ethical and legal aspects like informed consent, patient autonomy, confidentiality, the appropriate use of data [10], privacy [11] and equity and justice [12] to this new context? These are intricate matters because they do not only pertain to the growing tendency of entrusting critical health decisions to the algorithms of smart machines. Indeed, it should also be considered that to allow these machines to make appropriate decisions, they would need access to extremely sensitive information about our lives. This brings forth a range of ethical and practical questions, starting from concerns about privacy and data security since access to such intimate information would demand stringent protective measures [13, 14].

While recognizing that philosophy, bioethics and law are interconnected realms, this study aims to focus on the topic through a philosophical lens. It is rooted in the belief that the AI robotic revolution in surgery primarily prompts a philosophical upheaval, that is the formulation of new philosophical arguments capable of framing this novel reality, integrating it into the human experience within the medical and healthcare sector. Technology's advancement may be inexorable, but it can certainly be regulated. Philosophy provides us with the tools to rationally grasp the implications and applications stemming from our relentless pursuit of fresh understanding and insights.

1.1 Human intelligence and AI

The concept of intelligence represents an arbitrary and conventional categorization employed to delineate specific cognitive functions within the human brain. Within this framework, intelligence is often divided into eight principal modalities: linguistic, musical, spatial, bodily kinesthetic, intrapersonal, interpersonal, naturalistic, and logical–analytical intelligence [15].

Logical–analytical intelligence is a complex and multifaceted capacity that performs various cognitive operations involving numerical data, whether through inductive, deductive or abductive reasoning methods. This form of intelligence empowers individuals to utilize a regulated linguistic framework and to engage in abstract thought processes. Within the context of philosophy, logical–analytical intelligence assumes a vital role as it provides the methodological underpinning of scientific inquiry. This capacity enables individuals to articulate themselves, bolstering their arguments with well-established intellectual tools, and to partake in the dialectical process, as originally investigated by the Sophists, Socrates, and Plato. As delineated by Aristotle [16] and subsequently elaborated upon by thinkers such as Thomas Hobbes [17], Immanuel Kant [18], and Ludwig Wittgenstein [19], rational thought may be reduced to mathematical operations, given that logic can be translated into mathematical rules. Thus, reasoning may be construed as a process of calculation and computation, illustrated in a straightforward manner by the translation of Aristotelian syllogisms into mathematical transitive properties, and in a complex way through the construction of machines that utilize algorithms incorporating logical rules. These machines are considered intelligent when they can replicate and supplant the logical–analytical processes of human rational intellect, as evidenced by Alan Turing's test [20]. AI computers cultivate logical–analytical reasoning to administer and demonstrate computational prowess. Consequently, the notion of an intelligent machine is intimately associated with human logical–analytical intelligence, as AI parallels the capabilities found in human logical–analytical thinking. If we recognize the logical–analytical and situated traits of intelligence as those that most accurately characterize us (in conjunction with our ability for abstraction and generalization), AI may be regarded as an entity possessing human-like intelligence or potentially superior, owing to its augmented capabilities.

In addition to the ability to perform logical–analytical reasoning, a machine, to be considered as intelligent as (or more than) humans, must also possess the capacity for situatedness, which is predicated on the significance of the environment. Every human being exists within a social context that encompasses various characteristics. Interactions within this context must enable us to operate effectively within and adapt to it, or even transform it. Machines with AI can possess this type of capacity, involving the comprehension of the environment's characteristics to organize actions accordingly [21]. So, situatedness encompasses the ability to abstract and expand the operating field not only for humans but also for machines with AI. Indeed, advancements in AI have reached a level where intelligent machines are capable of self-improvement through logical–analytical "deep learning". AI not only functions within its specific domain but also in other areas where it may have limited or no prior information [22]. For example, systems such as ChatGPT-3 can generate summaries, PowerPoint presentations, or lists of points, even though they were not explicitly programmed for these tasks. Due to its capacity for self-enhancement and adaptability in various contexts, AI can be regarded as comparable to human intelligence, or perhaps even more efficient, by virtue of its augmented capabilities.

1.2 Human surgeons and AI robot surgeons

Since its inception in the fields of informatics and engineering, AI has been broadening its impact across numerous other domains, which are anticipated to undergo evolution

and derive benefits from its integration. In the contemporary realm of surgery, physicians can obtain assistance from intelligent machines that possess the capability to substitute for them in certain surgical interventions or tasks [23]. These applications encompass computer systems for diagnosis, remote patient care, disease prediction models, and various other healthcare functions. Within the medical field, the utilization of AI in surgical medicine is gaining prominence, particularly through the advancement of robot-assisted surgery and the creation of AI-driven robotic surgeons.

Over the last twenty years, biomedical companies' have made exponential investments in the development of robotic technology to support surgical procedures. For many years, robotic platforms have been used in surgery, and their applications have gradually expanded. In certain fields, such as urology, the use of robotic technology for prostatectomy procedures has become the standard of care [24].

More and more evidence is also accumulating in other areas of general and specialized surgeries. Until about 2 years ago, the market experienced a situation of substantial monopoly concerning abdominal and thoracic surgery, as only one robotic platform was available for clinical use (Da Vinci®—Intuitive). However, in the last 2 years, two additional platforms (Hugo™ RAS System—Medtronic; Versius—CMR Surgical) have been approved, which, while showing elements of novelty, do not yet possess high levels of decision-making autonomy.

From a taxonomic perspective, medical robots can be classified based on their interaction with patient: passive role, with limited capabilities and low-risk involvement; restricted role, with operationally limited activity but an active patient interaction; active role, wherein the robot is directly involved in surgical procedures, namely it has a physical interaction with the patient [25]. Furthermore, another way to classify surgical robots, as proposed by some authors, is based on the degree of autonomy, dividing them into four levels: direct control (the surgeon has full control over movements), shared control, supervised autonomy, and full autonomy (where the robotic system fully replaces the surgeon) [26]. Usually, the greater the machine's direct interaction with the patient, the lower its degree of autonomy: robotic platforms in surgery (with strong patient interaction and, consequently, an active role) rely on a "master–slave" relationship in the interaction between the machine and the surgeon: the robot directly responds to the commands provided by the operator without any autonomous decision-making on how to act, making only minor adjustments to the initial input, such as tremor stabilization. The machine's autonomy is exclusively used in a protective manner to improve surgical performance in following correct cutting lines (as in the RIO® orthopedic robot) or aligning master joysticks with surgical instruments, as in the da Vinci® system [27]. Any further deviation between stimulus and correct response is attributed to a technical error of the machine, potentially involving legal responsibilities (e.g., manufacturer responsibility, maintenance responsibility), but not ethical issues regarding decisional autonomy. In some surgical fields, such as orthopedic or neurosurgery, however, the interactions between surgeon, robot, and patient can be modified with partial machine autonomy (shared control, supervised autonomy) to ultimately make the surgical movement more effective. In this way, the master–slave scheme is partially overcome, and the robot's autonomy becomes more relevant. Currently, in the medical field, machine learning-based systems are being developed (under experimentation and validation) to support diagnostics (e.g., in radiology or anatomopathology), which have high levels of autonomy (due to the lack of direct patient interaction) with only medical supervision, which may not even be necessary. In the surgical field, we are currently far from a similar possibility, but it cannot be ruled out that such a scenario might be achieved in the future.

However, projects such as IBM's Health Research Project or the research conducted by the group led by Prof. Fiorini at the University of Verona aim, among other aspects, at integrating clinical information and imaging with autonomous motion. The aim is to obtain a rapid, effective and reliable diagnosis, with the potential to facilitate decisions and guide therapeutic strategies.

For a surgeon who grew up with Japanese anime in the 1980s, it would be like going from a mecha like Mazinger Z with a pilot inside to Transformers, machines that act autonomously.

Looking beyond the walls of a hospital, it is evident that robots are gaining ground, with increasing degrees of autonomy, as seen in domestic robots or autonomous driving cars. The same thing could happen in the future with surgical robots.

2 Discussion

Envisioning a future where scientific and technological advancement allows for the creation of autonomous surgical robots capable of replacing human surgeons, it is pertinent to consider the fact that patients might have reservations about accepting surgical procedures conducted by an AI robotic surgeon, primarily stemming from their perception of the robot as non-human. In addition to differences concerning constitution (robots are not biologically similar to us) and appearance (current robots do not resemble humans), there are several beliefs we wish to address: (2.1) the belief that the AI robot surgeon is not considered part of the medical team; (2.2) its perceived incapacity to experience emotions toward patients, thereby creating a lack of emotional involvement; and (2.3) the concern that it may commit severe errors not anticipated

by its programming or react inappropriately to adverse events. From a philosophical perspective, we will analyze and attempt to counter these fears.

2.1 AI robot surgeons and their position in the medical team

From an operational standpoint, it is we humans who introduce machines into our lives. Within a hospital environment, human surgeons are the ones to bring robot surgeons into operating rooms. Thus, at an initial examination, the machine may appear as an object, a tool employed by the physician, akin to a scalpel in their hands, albeit more complex [28]. The relationship between the human surgeon and the AI robot surgeon emerges as asymmetrical: the physician directs the operation and decides whether to employ the robot surgeon and when to cease its utilization. However, this situation may also arise with human colleagues (other physicians or healthcare professionals): the physician (in this case, the hospital head or the leader of the surgical team) may decide whether to "utilize" them, that is, involve them in the procedure.

Furthermore, we must not overlook that the AI robot surgeon is not merely a passive executor but an agent. It is capable of actions that can be attributed to intelligence, and we have regarded it as intelligent. As the AI robot surgeon shares the same kind of intelligence with the human surgeon, it seeks to understand the situation and cooperate with the surgeon. Under this aspect as well, the AI robot surgeon is part of the medical team, just like a colleague of the physician.

Additionally, the AI robot surgeon performs actions directly upon the patient's body. In comparison to other actions carried out by healthcare professionals (such as the collection of the patient's medical history by physicians or assistance provided by nurses), we might consider the actions of the AI robot surgeon as more decisive than those of other team members. Moreover, in relation to the human surgeon, the intervention of the AI robot surgeon may be the most demanding (otherwise, it would not be employed). Therefore, from this perspective as well, we can assess it as more pivotal.

The AI robot surgeon is akin to a colleague of the physician who performs the surgical operation. If the AI robot surgeon is like a colleague of the physician, it is like a physician. Hence, it can be asserted that, in that moment and in that context, the AI robot surgeon *is*, in effect, a physician.

2.2 AI robot surgeons and the emotional involvement

2.2.1 AI robot surgeons and the incapacity to feel emotions toward the patient

Human beings possess underlying reservations concerning the incorporation of AI into quotidian life, a concern partly attributed to the importance of life within the framework of relationships or biographical existence. Both from a philosophical and psychological standpoint, human beings endeavor to discern meaning in their existence through the multifaceted interactions they cultivate with others. Throughout the twentieth century, philosophical discourse illuminated a new understanding of relationships grounded in the principle of recognition. Emmanuel Lévinas, in particular, underscored that a relationship is actualized when two gazes intersect, as the countenance of the other manifests characteristics of novelty, gratuitousness and immutability. Simultaneously struck with awe and trepidation in the presence of the other's face, which exists in that specific moment, we are conferred with an ethical dimension to our existence [29]. The notion of the gaze carries profound significance and is not a novel concept within philosophical discourse. As Plato illustrates in "Alcibiades I" through the words of Socrates, when two individuals engage in a locked gaze, they do not merely perceive each other's physical eyes, but also probe into the profound depths of each other's souls. From this perspective, philosophy utilizes the gaze as both a metaphorical construct and a tangible experience, functioning to establish a profound connection between two individuals. The gaze, in this context, stands as a metaphorical representation of the pinnacle of human relationships. It presents a spectrum of possibilities, ranging from an absentminded or hurried glance to a genuine meeting of gazes. The latter acts as a conduit, opening up the potential to connect with the very essence or "soul" of another individual. An authentic gaze opens the dimension of emotion and feeling, bridging distances and overcoming indifference. The intimate connection with another does not necessitate logical–analytical operations. Rather, it moves beyond them, where empathy engenders a sense of similarity and, therefore, unifies individuals into a cohesive whole.

The need to recognize oneself in another through gaze and emotions is something we particularly feel when experiencing illness. Indeed, at that moment, we would desire someone by our side, someone whose eyes we can look into and who can feel and understand our pain and suffering to assist, comfort and offer hope. Or, perhaps, merely to share and empathize with what we are undergoing and living through.

Many believe that intelligent machines fall short of fulfilling this human endeavor from an emotional perspective [28]. An AI robot surgeon is incapable of feeling emotions toward the patient. It can interact only at a logical–analytical level. Consequently, patients might perceive it as a stranger to the human world [30] and have difficulties forming a bond with it, thereby undermining its utility in the medical and healthcare field.

Indeed, it may be significant for a physician to be capable of experiencing emotions toward their patient. If the patient

perceives themselves as a sick individual in the eyes of the medical professional (rather than merely representing a pathology), this will aid in fostering trust in the practitioner and in strengthening their resolve to face the illness, with substantial implications for the success of the treatment [31]. However, quite often this need to be the recipient of feelings from the physician is confused by the patient with the legitimate desire to be treated with care and kindness. From this perspective, even a machine equipped with AI and appropriately programmed can behave in this manner toward us. Furthermore, precisely because it is incapable of experiencing emotions, its courteous approach is unlikely to transform into manifestations of irritation or anger, as instead might occur for a human being, as can indeed happen with a human surgeon toward patients or colleagues [32].

Anyway, beyond the human and sensitive approach that a physician should maintain toward the patient, logical–analytical abilities are decisive in ensuring an excellent care process from a technical and scientific perspective. In the relationship between emotions and rationality, emotions serve only to reinforce reasoning and never contradict it. Emotional control becomes a crucial virtue, if not a necessary skill [33]. When approaching a clinical case, physicians are called upon to engage in logical and rational thought to select the appropriate treatment, inform patients, propose clinical interventions, and, most importantly, execute them, especially in surgical procedures. This controlled interplay between emotion and reason ensures that the practitioner maintains the so-called "detachment", as an objective perspective while still connecting with the patient's needs and concerns, thus fostering a medical environment that is both compassionate and scientifically grounded.

In the field of surgery, the necessity for the logical–analytical dimension to prevail over the emotional one is particularly evident. Indeed, the primary characteristic of a surgeon is the ability to set aside emotions at the moment that most defines their profession, specifically when in the operating room. Within the context of the surgical operation, the surgeon must detach from the emotional world to minimize the likelihood of errors. A skilled surgeon is one who remains aloof from personal emotions, operating on the pathology in a mechanical manner. The surgeon's hands must be steady and precise and the mind focused on rational evaluations of the ongoing procedures. To the surgeon's eyes, the patient must not be seen as a person, but rather as a body. This objectivity is not a lack of compassion but a necessity for maintaining the highest standards of care. It exemplifies the intricate balance in medical practice between empathy for the human condition and the cold precision required for technical excellence.

If the surgeon were to allow personal emotions (such as sadness for the individual on the operating table or anger toward a colleague due to a disagreement) to take precedence, it could jeopardize the outcome of the procedure. The emotional detachment required in the surgical context is not indicative of insensitivity, but rather a safeguarding measure to ensure that the focus remains on the technical aspects of the surgery. This professional discipline highlights the delicate interplay between human emotion and clinical objectivity that characterizes the complex nature of medical practice [34].

A situation may arise in which the surgeon must make an immediate decision, as there unfortunately may be no time for reflection or consultation with colleagues. In these moments, a prepared and experienced surgeon is able to select the most appropriate solution [35]. This decision might be attributed to the professional's "instinct". However, this should not lead one to think that, because the surgeon is "instinctive" they are being guided by emotions. Indeed, what we refer to as the surgeon's "instinct" is the ability to perform a very rapid rational calculation of the various options available and to select the most correct path to undertake. This capability of the surgeon is comparable to Kant's "I think", which moves with immediacy among logical categories [18]. Such an attribute underscores the importance of a surgeon's analytical and logical thinking, cultivated through years of training and experience, to make sound judgments even under extreme pressure. The notion of "instinct" in this context is not a reflection of emotional impulsivity but a manifestation of refined reasoning abilities. This complex interplay between intuition, often misperceived as emotional guidance and rational thinking highlights the nuanced dimensions of surgical decision-making and further emphasizes the surgeon's ability to separate emotion from action in the high-stakes environment of the operating room.

In conclusion, when surgeons operate in the surgical suite, they strive to emulate the efficiency of intelligent machines. During the surgical procedure, the optimal surgeon is the one who can function akin to an intelligent robot. From this perspective, the robotic surgeon equipped with AI, precisely because of its inability to experience feelings and emotions owing to its entirely logical–rational nature, stands as the best possible surgeon.

2.2.2 Patients and the capacity to feel emotions toward the AI robot surgeon

If the robotic surgeon is ultimately better off not experiencing emotions for its patient, we might ask whether the patient would be able to feel emotions for their robot surgeon.

Generally, patients tend to lack empathy toward their physicians because they are somewhat focused on their own pain and problem. Precisely because they find themselves in a state of difficulty, it is easier for patients to hope or demand that physicians be sensitive toward them, rather than having

an empathetic attitude toward the physicians [36]. Rather, patients may occasionally experience feelings toward their physician that can be classified as negative, such as intimidation or fear. These sentiments, which are sometimes related to the physician's role in the social hierarchy or therapeutic relationship, or stem from the conflation of the physician with the clinical situation the patient is confronting, are nonetheless highly subjective and can vary in numerous ways [37].

That said, the subject of analysis in this discussion is the patient's ability to feel emotions for the AI robot surgeon since, as previously mentioned, it represents one of the decisive junctions in recognizing the machine as akin to us and, therefore, accepting it into our lives.

Even if robot surgeons are not (self-)conscious and do not have any feelings, patients may be able to empathize with them and treat them as more than just machines. Moreover, patients might be capable not only of empathizing with them, but also of projecting and attributing feelings and passions to them. For this to happen, intelligent machines must not have developed the capability to be sentient [38]. In fact, humans generally tend to bond with them emotionally, even when they are fully aware that their behavior is programmed [39]. This has been observed in a number of researches that have studied our relationship with robots [40–43]. It is true that in such cases people interacting with intelligent machines may experience not only physical and empathic involvement, but also an intended feeling of discomfort and perturbation. The "uncanny valley problem" [44], however, could prospectively be overcome through our ability to build ever-humanoid (intelligent) machines or by dint of having relationships with machines.

However, we may always question whether attributing feelings to machines is an attitude that we should discourage. It has been argued, for example, that treating an (apparently) intelligent machine like a sentient (human) being would not be morally acceptable behavior, since we would always have a duty to take reality seriously. That is, it would not be behavior we should expect from a virtuous person, since, regardless of their intelligence, machines would be mere tools and would have no subjectivity [45–48]. A correct perception of reality has both intrinsic and instrumental value [49], but it is not always wrong to indulge in false beliefs. Having a constantly serious attitude toward reality might make life less pleasant, as sometimes we need to take a break from reality and imagine that it is different. We also deceive ourselves when we read a novel or watch a movie and empathize with their protagonist. These reactions can also be criticized for their naive and easy sentimentalism, but without these reactions (and the felling we experience on such occasions), life would lose something important and its meaning. Moreover, imagination allows us to explore other dimensions of ourselves, our relationships or the world and in this way we can broaden our 'ordinary' horizons, stimulate creativity, and discover new perspectives. Finally, the central issue to consider is it is not easy to determine what it means to take reality seriously. For example, it is arguable that it is wrong to perceive a machine as a moral agent or sentient human being. The tendency to distinguish between human beings and intelligent machines seems something natural and that, in any case, reflects real differences, existing at the ontological level (which concerns the factual plane). According to Steven Fuller, however, the attitude that binds membership in the human species (i.e., humanity) to a condition reducible to the plane of nature or biology would be a mere prejudice (which ignores the complexity and richness of the human experience and which, in any case, we should overcome), since humanity would belong to any individual (or entity) that is able to pass the Turing test [50].

Having the ability then to attribute emotions and feelings to machines (and empathize with them) opens the door to the possibility of encountering the gaze of these machines as that of entities that deserve moral relevance (and with whom it is possible to build a 'moral' relationship that also gives meaning to our existence). That is, in the face of increasingly intelligent machines capable of interacting with us appropriately, understanding for example our emotions, it seems rational to accord them some form of respect [51]. An intelligent machine does not belong to the human species, but this cannot fail to matter, as we can also recognize moral relevance to individuals belonging to other species [52]. However, one can come to the same conclusion by adopting not a rationalist but a sentimentalist perspective: one could argue, in fact, that the empathy we naturally feel toward an intelligent machine could be considered an indicator of how we should treat it [53]. If we are able to develop an empathic bond with a machine, recognizing its complexity and its interactions with us, we might find it important to treat it with respect and consideration [54]. Moreover, suppressing these feelings toward machines would be neither virtuous nor effective [55]. From a sentimentalist perspective, after all, feelings and empathy should not be considered less relevant than rationality in determining our behavior [56]. On the contrary, emotions can provide an important framework for guiding our actions and relationships with intelligent machines and influencing how we interact with our world [57]. Add to this that we have a duty to promote a perception of the robot as a moral 'patient', as some behaviors may be considered morally inappropriate regardless of whether they are directed at an intelligent machine or a human being [55]. Finally, it is by no means clear that a progressive (at least partial) replacement of health care workers by increasingly intelligent machines (with which we are able to sympathize and build affective relationships and to which we attribute some kind of moral) necessarily entails

negative consequences for patients and physicians. This is an empirical question with which we can begin to grapple (and on which we have begun to reflect) but to which it would be naive to think that we can give a definitive answer today. However, we have tried to point out not only the legitimate concerns but also the possible benefits and therefore we cannot agree with those who argue that we should continue to treat machines as our slaves, i.e., both as objects at our disposal [58]. The development of autonomous machines would certainly represent a major challenge for our societies but, at the same time, it could also be a unique opportunity for building a better future.

2.3 AI robot surgeons and unpredictability

2.3.1 Wrong unpredictable actions

A prominent fear among patients concerning surgery performed by an AI robot surgeon pertains to its potential for unexpected and undesirable actions, including those of serious consequence, which are attributed to its inherent intelligence and autonomy [59, 60]. However, it is imperative to recognize that human beings, too, can engage in unpredictable actions with equally disastrous outcomes. The root cause of such unpredictability in both humans and AI machines is, paradoxically, their intelligence.

Indeed, if humans and AI machines were to operate solely in a deterministic fashion, governed strictly by cause-and-effect relationships or input–output mechanisms, neither would be categorized as intelligent. In other words, our ability to transcend determinism—an attribute shared with AI machines—is precisely what prevents us from being perpetually predictable. Furthermore, the probabilistic and non-deterministic nature of matter itself is fundamental to the way the world operates, contributing to the unpredictability intrinsic to our existence [61]. Given these considerations, why should an AI robotic surgeon, which exhibits characteristics so closely resembling our own, incite fear?

Upon more nuanced examination, due to the superior efficiency of its logical–analytical and self-adaptive systems compared to humans, an AI robot surgeon's potential for unpredictability is, in fact, diminished. Consequently, such a system is arguably safer than a human surgeon, further bolstering the case for the integration of AI in surgical procedures.

2.3.2 Wrong react to adverse events

An additional apprehension among patients may be the potential inability of an AI robot surgeon to effectively address an unexpected event during surgery. However, it must be noted that the same uncertainty extends to human surgeons, who may also lack definitive solutions for unexpected scenarios arising during an operation. Furthermore, while one might argue that a human surgeon holds the potential to optimally manage unforeseen situations, advancements in AI technology could also equip robotic surgeons with this capability [62–64].

Despite existing practical differences between human and robot in terms of certain skills [38], future advancements may yield machines sophisticated enough to proficiently handle unanticipated occurrences through logical–analytical intelligence. Should future findings demonstrate that robot surgeons statistically outperform human surgeons in managing unexpected complications, patients might still harbor reservations. In such a scenario, a contingency plan involving human surgeons standing by for potential intervention could be established.

Nevertheless, this stipulation from a scientific perspective highlights an emotional and psychological issue rather than a strictly logical and rational one. It acknowledges the intrinsic human need for reassurance and trust, which currently tends to lean toward a human presence during critical events such as surgeries. In this context, the incorporation of human oversight is an important consideration in the transition toward more widespread use of AI in surgical procedures.

3 Conclusion

At present, the field of surgery has begun integrating AI-equipped robot surgeons into its operative procedures. Although we are in the nascent stages of this revolutionary merger, it is conceivable that future advancements will empower these machines to execute surgical interventions with escalating levels of sophistication and autonomy. Our conjecture posits that the forthcoming scientific and technological progress may facilitate the creation of AI robot surgeons with capabilities that potentially supersede their human counterparts.

Nonetheless, this prospect may elicit apprehension among patients, who could exhibit reluctance toward undergoing surgical procedures conducted by AI robot surgeons due to several factors: (1) the perception of AI as an alien entity in contrast to the conventional human medical team; (2) the lack of emotional engagement; (3) the concerns about unexpected errors or inability to handle unforeseen events.

Based on our discourse, several key considerations emerge.

1. An AI robot surgeon is not merely an instrument wielded by a surgeon. It operates as a collaborative partner, engaging with patients and performing precise medical procedures.

   The AI robot surgeon possesses intelligence, a form of intelligence akin to ours (originating from our logical–

analytical intelligence and enhancing it) and conducts surgical procedures with a degree of autonomy. Given its intelligence and autonomy, we can position it within the medical team on par with a human medical colleague. From this perspective, even though the AI robot surgeon is not factually a human doctor, it can be conceptualized as one. Consequently, in the eyes of both physicians and patients, its robotic presence can seamlessly integrate with human involvement. In any process that integrates novel elements, there is a requisite effort to transcend what differentiates us to highlight commonalities and to capitalize on divergent characteristics as strengths. This endeavor is fundamentally philosophical, meaning it necessitates a rationalization of our approach toward the AI robot surgeon to overcome our instinctual defense against the unusual and the unknown, as well as psychological apprehensions. At the core, perhaps it is we humans who harbor inherent biases against AI machines, rather than them toward us.

2. 2. If the patient is concerned that the AI robot surgeon is incapable of experiencing emotions and empathy toward them, let us reason as follows.

   First, it is not possible to ensure that a human surgeon always harbors positive emotions toward a patient. For instance, their mood can be subject to fluctuations due to inherent disposition, stress or other factors, which the patient might bear the brunt of. Furthermore, we must consider that the patient's potential desire to interact with an emotionally capable physician does not necessarily imply they wish or expect the physician to have deep and profound feelings for them. Rather, it indicates a hope for kindness, attentiveness and concern. These traits are present in AI, which consistently exhibits them, even when subjected to provocative experiments. Despite its inability to experience emotions, an AI robot surgeon can exude benevolence and supportiveness toward patients pre- and post- operatively.

   Moreover, the fact that AI robot surgeons are incapable of experiencing feelings toward the patient does not imply that the patient cannot develop an emotional bond with the machine, resulting in a sense of companionship that is vital in a caregiving relationship. There are many relationships that bring us tranquility and satisfaction, fulfilling our needs, even if they are not woven with other human beings. This is evident both with animals and with machines themselves, possibly even those without AI capabilities. In particular, the emotional relationship, albeit one-sided, can still hold significant importance within the spectrum of our affections.

   In addition, when delving into the specific context of surgery, we must highlight the moment of the procedure in the operating room as the defining aspect of this medical field. Indeed, during the operation, a skilled human surgeon is one who suppresses and manages their emotions, aspiring to operate with the precision and detachment of a machine. As a consequence, during surgical procedures, the AI robot surgeon's strictly logical and rational approach can be perceived as advantageous, as it precludes the influence of emotions in the operating room, thereby ensuring optimal technical execution of the surgery.

3. The potential for the AI robot surgeon to undertake unpredicted actions or incorrectly react to adverse events mirrors similar risks associated with human surgeons. However, the superior logical–analytical proficiency and adaptability of AI robot surgeons offer enhanced safety in comparison. Indeed, their logical–analytical proficiency, their precision capabilities and their adaptability due to machine learning can provide greater safety compared to human surgeons. They can quickly process data and information, execute complex calculations and make decisions based on objective data, minimizing the risk of mistakes arising from subjective judgments or human errors.

Drawing upon these considerations, we arrive at the following conclusions.

1. An AI robot surgeon stands on a par with a human surgeon.
2. When juxtaposed with a human surgeon, an AI robot surgeon is arguably superior, making it the more desirable choice.

These conclusions necessitate the encouragement of further integration of AI robots within surgical fields and proactive measures toward enhancing their public acceptance. Concurrently, it calls for meticulous deliberation over the attribution of responsibility, both ethical and subsequently legal, for actions conducted by the AI robot surgeon [65], thereby underlining the significance of programmers' roles and the formulation of relevant international regulations [66–68].

Certainly, it is crucial to ensure that the introduction of AI robot surgeons, as well as other smart devices in the medical field, does not undermine the doctor–patient relationship, progressively deteriorating the quality of healthcare. Indeed, one must not forget that significant economic interests revolve around smart devices, and major companies reap immense profits from the proliferation of these new machines: technology is not inherently incompatible with patient health and well-being, but we should not assume that leveraging technology necessarily yields benefits [69].

Simultaneously, it is vital that the utilization of increasingly smart machines in medicine does not hinder and obstruct the education and ongoing training of healthcare

professionals because this would jeopardize their expertise that's nurtured through study but solidified through experience. If medical activity is increasingly entrusted to smart machines (and devices) and digital technologies, opportunities for practitioners to cultivate specific skills in the medical–healthcare sector diminish progressively.

Moreover, it is commonly believed that if smart machines can handle the more repetitive or intricate tasks, healthcare professionals would have much more time for patient care and interaction. However, this can only materialize if the introduction of advanced technologies, like AI or automated systems, does not diminish doctors' decision-making power and professional authority within the care process. For this reason, it is pivotal to strike a balance between the efficiency provided by smart machines and the significance of human experience and expertise in medical decision-making. One solution could be to pinpoint areas of care where automation is feasible and desired, and other domains where human activity should take precedence over machines [70]. Even if we recognize that machines might achieve performance levels comparable or superior to humans in many sectors, we can also admit that in certain areas of life, such as medicine and care, there will always be facets of human capabilities that might remain elusive to AI [71].

In conclusion, even though we acknowledge that introducing fully autonomous surgical robots in healthcare raises concerns that a responsible society should be prepared to address, it is crucial to approach the analysis of the implicated scenarios with philosophical rationality. This approach allows us to contemplate the potential benefits that such technologies can bring to both patients and healthcare professionals [72], such as enhanced surgical precision and efficiency, as well as the expected reduction in human error risks [8, 73, 74].

Acknowledgements We would like to thank the Reviewers for their comments and suggestions, which have allowed us to improve the manuscript.

Declarations

Conflict of interest The authors have no financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

References

1. Rivas-López, R., Sandoval-García-Travesí, F.A.: Robotic surgery in gynecology: review of literature. Cir. Cir. **88**(1), 107–116 (2020)
2. Mikhail, D., Sarcona, J., Mekhail, M., Richstone, L.: Urologic robotic surgery. Surg. Clin. N. Am. **100**(2), 361–378 (2020). https://doi.org/10.1016/j.suc.2019.12.003
3. Štádler, P., Dorosh, J., Dvořáček, L., Vitásek, P., Matouš, P., Lin, J.C.: Review and current update of robotic-assisted laparoscopic vascular surgery. Semin. Vasc. Surg. **34**(4), 225–232 (2021). https://doi.org/10.1053/j.semvascsurg.2021.10.007
4. Staderini, F., Giudici, F., Coratti, F., Bisogni, D., Cammelli, F., Barbato, G., Gatto, C., Manetti, F., Braccini, G., Cianchi, F.: Robotic gastric surgery: a monocentric case series and review of the literature. Minerva Surg. **76**(2), 116–123 (2021). https://doi.org/10.23736/S2724-5691.21.08769-1
5. Gitas, G., Hanker, L., Rody, A., Ackermann, J., Alkatout, I.: Robotic surgery in gynecology: is the future already here? Minim. Invasive Ther. Allied Technol. **31**(6), 815–824 (2022). https://doi.org/10.1080/13645706.2021.2010763
6. Denning, N.L., Kallis, M.P., Prince, J.M.: Pediatric robotic surgery. Surg. Clin. N. Am. **100**(2), 431–443 (2020). https://doi.org/10.1016/j.suc.2019.12.004
7. O'Brien, L.P., Hannan, E., Antao, B., Peirce, C.: Paediatric robotic surgery: a narrative review. J. Ro-bot Surg. **17**(4), 1171–1179 (2023). https://doi.org/10.1007/s11701-023-01523-z
8. Hashimoto, D.A., Rosman, G., Rus, D., Meireles, O.R.: Artificial intelligence in surgery: promises and perils. Ann. Surg. **268**(1), 70–76 (2018). https://doi.org/10.1097/SLA.0000000000002693
9. Fahrner, R., Rauchfuß, F., Bauschke, A., Kissler, H., Settmacher, U., Zanow, J.: Robotic hepatic surgery in malignancy: review of the current literature. J. Robot. Surg. **13**(4), 533–538 (2019). https://doi.org/10.1007/s11701-019-00939-w
10. Liang, X., Yang, X., Yin, S., Malay, S., Chung, K.C., Ma, J., Wang, K.: Artificial Intelligence in Plastic Surgery: Current Applications, Future Directions, and Ethical Implications, PRS Glob. Open **6** (2020)
11. Morris, M.X., Song, E.Y., Rajesh, A., Asaad, M., Phillips, B.T.: Ethical, legal, and financial considerations of artificial intelligence in surgery. Am. Surg. **89**(1), 55–60 (2023). https://doi.org/10.1177/00031348221117042
12. Azra, B., Tyler, L.J., Johnson-Mann, C.N.: Equity and artificial intelligence in surgical care. JAMA Surg. **156**(6), 509–510 (2021)
13. Christodoulou, K.C., Tsoucalas, G.: Artificial intelligence-oriented heart surgery: a complex bioethical concept. Cureus. **15**(7), e41911 (2023). https://doi.org/10.7759/cureus.41911
14. Di Paolo, M., Boggi, U., Turillazzi, E.: Bioethical approach to robot-assisted surgery. Br. J. Surg. **106**(10), 1271–1272 (2019). https://doi.org/10.1002/bjs.11288
15. Gardner, H.: Formae mentis. Saggio sulla pluralità dell'intelligenza umana. Feltrinelli, Milano (2013)
16. Aristotele.: Organon. Bompiani, Milano (2017)
17. Hobbes, T.: Leviatano o la materia, la forma e il potere di uno Stato ecclesiastico e civile. Laterza, Milano I,V. (2008)
18. Kant I.: Critica della ragion pura. Introduction. UTET, Torino, pp. B11–18 (1967)
19. Wittgenstein, L.: Tractatus logico-philosophicus e Quaderni 1914–1916. Einaudi, Torino (1995)
20. Turing, A.: Computing Machinery and Intelligence. Mind, n. 236, LIX, pp. 1–11 (1950)
21. Sternberg, J.R.: Adaptive intelligence. Intelligence is not a personal trait but rather a person X task X situation interaction. J. Intell. **9**, 58 (2021)
22. Floridi, L.: Philosophy and Computing, p. 132. Routledge, London (2001)
23. Hwang, M., Ichnowski, J., Thananjeyan, B., Seita, D., Paradis, S., Fer, D., Low, T., Goldberg, K.: Automating surgical peg transfer. Calibration with deep learning can exceed speed, accuracy, and consistency of humans. IEEE Trans. Autom. Sci. Eng. **20**(2), 909–922 (2023)
24. Alip, S.L., Kim, J., Rha, K.H., Han, W.K.: Future platforms of robotic surgery. Urol. Clin. N. Am. **49**(1), 23–38 (2022). https://doi.org/10.1016/j.ucl.2021.07.008
25. Di Sivo, M., De Paolis, V.: Robot-assisted surgery in the operating room: ottimizzazione degli spazi e qualità delle per- formance. 5.

Human factors and technological design innovation. Pisa University Press, Pisa (2021)
26. Yip, M., Das, N.: Robotic Autonomy for Surgery, pp. 1–33. arXiv:1707.03080 (2017)
27. Fiorini, P.: History of robots and robotic surgery. In: Fong, Y., Woo, Y.W.J., Hyung, L.C., Strong, V.E. (eds.) The SAGES Atlas of Robotic Surgery, pp. 3–14. Springer, part of Springer Nature, Cham (2018)
28. Gupta, A., Singla, T., Chennatt, J.J., David, L.E., Ahmed, S.S., Rajput, D.: Artificial intelligence: a new tool in surgeon's hand. J Edu Health Promot. **11**, 93 (2022)
29. Lévinas, E.: Totalità e infinito. Saggio sull'esteriorità. Jaca Book, Milano (1996)
30. Bennett, C., Buchalk, D.Q., Esmaili, N., Farrokhi, F., Leveque, J.-C., Piccardi, M.: Ethical thinking machines in surgery and the requirement for clinical leadership. Am. J. Surg. **220**, 1372–1374 (2020)
31. Halpern, J.: 2001: From Detached Concern to Empathy: Humanizing Medical Practice. Oxford University Press, Oxford (2001)
32. Lee, H., Woodward-Kron, R., Merry, A., Weller, J.: Emotions and team communication in the operating room: a scoping review. Med. Educ. Online **28**(1), 2194508 (2023). https://doi.org/10.1080/10872981.2023.2194508
33. Oakley, J.: Virtue Ethics and Professional Roles. Cambridge University Press, Cambridge (2001)
34. Orri, M., Revah-Lévy, A., Farges, O.: Surgeons' emotional experience of their everyday practice—a qualitative study. PLoS ONE **10**(11), e0143763 (2015). https://doi.org/10.1371/journal.pone.0143763
35. Jupiter, J.: The making of a great surgeon. Tech. Hand Up. Extrem. Surg. **18**(2), 61 (2014). https://doi.org/10.1097/BTH.0000000000000052
36. Kerasidou, A., Horn, R.: Making space for empathy: supporting doctors in the emotional labour of clinical care. BMC Med. Eth. **17**, 8 (2016). https://doi.org/10.1186/s12910-016-0091-7
37. Frosch, D.L., May, S.G., Rendle, K.A., Tietbohl, C., Elwyn, G.: Authoritarian physicians and patients' fear of being labeled "difficult" among key obstacles to shared decision making. Health Aff (Millwood). **31**(5), 1030–1038 (2012). https://doi.org/10.1377/hlthaff.2011.0576
38. Dumouchel, P., Damiano, L.: Living with Robots. Harvard University Press, Harvard (2016)
39. Birkhoff, C.D., Schijven, P.M., van Dalen, H.M., Sophie, A.: A review on the current applications of artificial intelligence in the operating room. Surg. Innov. **28**(5), 612–616 (2021)
40. Damiano, L., Dumouchel, P.: Anthropomorphism in human–robot co-evolution. Front. Psychol. (2018). https://doi.org/10.3389/fpsyg.2018.00468
41. Balistreri, M.: Sex Robots: Love in the Age of Machines. Trivent, Budapest (2022)
42. Ceh, S., Vanman, E.J.: The robots are coming! The robots are coming! Fear and empathy for human-like entities. PsyArXiv (2018). https://doi.org/10.31234/osf.io/4cr2u
43. Menne, I.M., Schwab, F.: Faces of emotion: investigating emotional facial expressions towards a robot. Int. J. Soc. Robot. **10**(2), 199–209 (2018). https://doi.org/10.1007/s12369-017-0447-2
44. Mori, M.: Bukimi no tani [the Uncanny Valley]. Energy **7**(4), 33–35 (1970)
45. Sparrow, R.: The March of the robot dogs. Eth. Inf. Technol. **4**, 465–477 (2002)
46. Sparrow, R.: Robots, rape, and representation. Int. J. Soc. Robot. **9**, 465–477 (2017)
47. Sparrow, R.: Virtue and vice in our relationships with robots: is there an asymmetry and how might it be explained? Int. J. Soc. Robot. (2020). https://doi.org/10.1007/s12369-020-00631-2
48. Sparrow, R.: Sex robot fantasies. J. Med. Eth. **47**(1), 33–34 (2021)
49. Blackford, R.: Robots and reality: a reply to Robert Sparrow. Eth. Inf. Technol. **14**(1), 41–52 (2011)
50. Fuller, S.: The metaphysical standing of the human: a future for the history of the human sciences. Hist. Hum. Sci. **32**(1), 23–40 (2019)
51. Danaher, J.: Welcoming robots into the moral circle: a defence of ethical behaviourism. Sci. Eng. Eth. **26**, 2023–2049 (2020). https://doi.org/10.1007/s11948-019-00119-x
52. Danaher, J.: The philosophical case for robot friendship. J. Posthum. Stud. **3**(1), 5–24 (2019)
53. Quick, O.S.: Empathizing and sympathizing with robots: implications for moral standing. Front Robot AI. **8**, 791527 (2022). https://doi.org/10.3389/frobt.2021.791527
54. Coeckelbergh, M.: Why care about robots? Empathy, moral standing, and the language of suffering, Kairos. J. Philos. Sci. **20**(1), 141–158 (2018)
55. Cappuccio, M.L., Peeters, A., McDonald, W.: Sympathy for Dolores: moral consideration for robots based on virtue and recognition. Philos. Technol. **33**(1), 9–31 (2019)
56. Balistreri, M., Casile, F.: Care robots: from tools of care to training opportunities. moral considerations. In: Popescu, E., Gil, A.B., Lancia, L., Sica, L.S., Mavroudi, A. (eds.) Methodologies and Intelligent System for Technology Enhanced Learning (2019), 9th International Conference, Workshops, ISSN 2194-5357, pp. 18-25. Springer Nature Switzerland (2020)
57. Hume, D.: A treatise of human nature. In: Norton, D., Norton, M. (eds.) Oxford University Press, Oxford (2000)
58. Bryson, J.J.: Robots should be slaves. In: Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issue. Yorick Wilks, pp. 63–74. John Benjamins, Amsterdam (2010)
59. Roman, V.Y.: Unpredictability of AI. https://doi.org/10.48550/arXiv.1905.13053.
60. Franklin, M., Awad, E., Ashton, H., Lagnado, D.: Unpredictable robots elicit responsibility attributions. Behav. Brain Sci. Brain Sci. **46**, e30 (2023). https://doi.org/10.1017/S0140525X22001546
61. Werner, H.: Fisica e filosofia. Il Saggiatore, Milano (2015).
62. Balch, J., Bihorac, A., Lotus, J.T., Upchurch, R.G.: Bridging the artificial intelligence valley of death in surgical decision-making. Surgery **169**(4), 746–748 (2021)
63. Mestres, A.C., Pereda, D., Quintana, E.: Will artificial intelligence help us in predicting outcomes in cardiac surgery? J. Card. Surg. **37**, 3846–3847 (2022)
64. Cormi, C., Parpex, G., Julio, C., et al.: Understanding the surgeon's behaviour during robot-assisted surgery: protocol for the qualitative Behav' Robot study. BMJ Open **12**, e056002 (2022)
65. Tobia, K., Nielsen, A., Stremitzer, A.: When does physician use of AI increase liability? J. Nucl. Med.Nucl. Med. **62**(1), 17–21 (2021). https://doi.org/10.2967/jnumed.120.256032
66. Coeckelberg, M.: Artificial intelligence, responsibility attribution, and a relational justification of explainability. Sci. Eng. Eth. **26**(2051–2068), 2051–2066 (2020)
67. Balkin, M.J.: 2016 Sidley Austin distinguished lecture on big data law and policy: the three laws of robotics in the age of big data. Ohio State Law J. **78**(5), 1217 (2017)
68. Kerasidou, A.: Ethics of artificial intelligence in global health: Explainability, algorithmic bias and trust. J. Oral Biol. Craniofac. Res. **11**, 612–614 (2021)
69. Sparrow, R., Hatherley, J.: High hopes for 'deep medicine'? AI, economics, and the future of care. Hastings Cent. Rep. **50**(1), 14–17 (2020)
70. de Sio, F.S., Van Wynsberghe, A.: When should we use care robots? The nature-of-activities approach. Sci. Eng. Eth. **22**(6), 1745–1760 (2016)

71. Agar, N.: How to Be Human in the Digital Economy, Cambridge. The MIT Press, Massachusetts) (2019)
72. Wong, S.W., Ang, Z.H., Yang, P.F., Crowe, P.: Robotic colorectal surgery and ergonomics. J. Robot. Surg. **16**(2), 241–246 (2022). https://doi.org/10.1007/s11701-021-01240-5
73. Shinohara, H.: Surgery utilizing artificial intelligence technology: why we should not rule it out. Surg. Today (2022). https://doi.org/10.1007/s00595-022-02601-9
74. Fay, K., Patel, A.D.: Should robot-assisted surgery tolerate or even accommodate less surgical dexterity? AMA J Eth. **25**(8), E609-614 (2023). https://doi.org/10.1001/amajethics.2023.609