Received: 22 February 2018 | Revised: 18 October 2018 | Accepted: 26 October 2018
DOI: 10.1002/rcs.1968

REVIEW ARTICLE

WILEY The International Journal of Medical Robotics and Computer Assisted Surgery

Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery

Shane O'Sullivan[1] | Nathalie Nevejans[2] | Colin Allen[3] | Andrew Blyth[4] | Simon Leonard[5] | Ugo Pagallo[6] | Katharina Holzinger[7] | Andreas Holzinger[8] | Mohammed Imran Sajid[9] | Hutan Ashrafian[10]

[1] Department of Pathology, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil
[2] Research Center in Law, Ethics and Procedures, Faculty of Law of Douai, University of Artois, France
[3] Department of History and Philosophy of Science, University of Pittsburgh, Pennsylvania
[4] Department of Computing and Mathematics, Faculty of Computing, Engineering and Science, University of South Wales, UK
[5] Department of Computer Science, Johns Hopkins University, Baltimore, Maryland
[6] Department of Jurisprudence, University of Turin, Italy
[7] Secure Business Austria, SBA Research gGmbH, Vienna, Austria
[8] Holzinger Group, HCI-KDD, Institute for Medical Informatics/Statistics. Medical University of Graz, Austria
[9] Department of Upper GI Surgery, Wirral University Teaching Hospital, UK
[10] Department of Surgery and Cancer and Institute of Global Health Innovation Imperial College London, UK

Correspondence
Dr Shane O'Sullivan, Department of Pathology, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Av. Enéas Carvalho de Aguiar, 155, São Paulo/SP, Brazil Cep 05403-000.
Email: doctorshaneosullivan@gmail.com

Funding information
CNPQ (Brazilian National Council for Scientific and Technological Development)

Abstract

Background: This paper aims to move the debate forward regarding the potential for artificial intelligence (AI) and autonomous robotic surgery with a particular focus on ethics, regulation and legal aspects (such as civil law, international law, tort law, liability, medical malpractice, privacy and product/device legislation, among other aspects).

Methods: We conducted an intensive literature search on current or emerging AI and autonomous technologies (eg, vehicles), military and medical technologies (eg, surgical robots), relevant frameworks and standards, cyber security/safety- and legal-systems worldwide. We provide a discussion on unique challenges for robotic surgery faced by proposals made for AI more generally (eg, Explainable AI) and machine learning more specifically (eg, black box), as well as recommendations for developing and improving relevant frameworks or standards.

Conclusion: We classify responsibility into the following: (1) Accountability; (2) Liability; and (3) Culpability. All three aspects were addressed when discussing responsibility for AI and autonomous surgical robots, be these civil or military patients (however, these aspects may require revision in cases where robots become citizens). The component which produces the least clarity is Culpability, since it is unthinkable in the current state of technology. We envision that in the near future a surgical robot can learn and perform routine operative tasks that can then be supervised by a human surgeon. This represents a surgical parallel to autonomously driven vehicles. Here a human remains in the 'driving seat' as a 'doctor-in-the-loop' thereby safeguarding patients undergoing operations that are supported by surgical machines with autonomous capabilities.

1 | INTRODUCTION

In this paper, we attempt to provide the building blocks of frameworks for development of standards in ethics, and in legal and regulatory compliance of artificial intelligence (AI) and autonomous robotic surgery. Acceptance of such robots requires a debate on ethics and trust. We approach this debate in the context of both open surgical procedures (eg, robotic laparotomy) and less invasive types of endoscopic surgical procedures (eg, robotic bronchoscopy). In addition to other use cases, we discuss AI and autonomous robotics for endoscopy to prevent infection or contamination, enhance precision, and provide automatic image analysis for navigation, evaluation, or diagnosis. This paper takes input from the existing body of literature and generates new classification, organisation, and analysis of it. We address many intriguing issues concerning the socio-legal implications of "robotic-assisted surgery" and "autonomous robotic surgery." We also examine the specifics of safety and security in modern day robotic surgery. We question 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? Indeed, with all new emerging technologies, there can be many gaps in the law, as laws struggle to stay ahead of these developing technological advancements. This paper summarized the current literature on robotic surgery, and we considered the fundamental principles of self-governing future autonomous robotic systems, ranging from their practical use to the regulatory and legal principles required for their incorporation into health care practice.

2 | SEEKING APPROVAL FOR USING AUTONOMOUS SURGICAL ROBOTS

In February 2018, the United States (US) Department of Defense closed its funding deadline for the Foundational Research for Autonomous, Unmanned, and Robotics Development of Medical Technologies (FORwARD) Award.[1] This recent call for proposals showed that there is clearly an interest in autonomous robotic surgery. We take into account that the level of this "interest" is merely relative, and that a research program of a few million dollars does not necessarily mean the US military considers it of equal importance to other programs. Nevertheless, the US military seems to be, at minimum, curious about exploring this new area.

Successfully developing state-of-the-art reliable autonomous surgical robots is one challenge. Seeking acceptance and approval for their use is an entirely different issue. A most obvious concern relates to the consequences surrounding patient death or disability resulting directly from a surgical error. This problem arises in other areas, eg, in autonomous driving.[2] One immediate solution might be the same as in autonomous driving: to put the human-in-the-loop and only assist the human in routine tasks, without giving full autonomy to the machine. In other words, "*let the human do what the humans can do best, and let the machines do what's more suited to the machines*",—we see the combination as being the most powerful.[3]

For Autonomous Driving, there already exists a standard. The Society of Automotive Engineers (SAE) has established the standard SAE J3016, which defines different levels of automation: Level 0 is full control by the driver, just warnings provided by the machine; level 1 requires hands-on, driver must be ready to retake full control at any moment; level 2 means hands-off, but the driver must monitor the vehicle at all times and must be prepared to immediately intervene in case of errors; level 3 means eyes-off, the driver can safely turn her/his attention elsewhere, but the driver may take over at any time; level 4 means mind-off, driver may even sleep or leave the driver's seat; and finally, level 5 is fully automatic, no human intervention required, eg, a driverless robotic taxi.[4] It is obvious that in levels 0 to 2 the human takes full responsibility; however, to date there are no solutions for the other levels. Therefore, at least in Europe, only levels 0 to 2 are currently legally allowed and European car manufacturers name it simply "driver assistance."

Today, similar attempts to define standards for automated surgical systems in the medical domain are missing, but this will become an important issue for medical product law. Currently however, time is not a major pressing factor as even the autonomous surgery equivalent of level 2 autonomous vehicles, in which a machine is capable of performing the operation entirely without human assistance, is many years away. A further issue is to determine how acceptable autonomous robotic surgery would be for patients in the future.[5] This leads one to pose the question of how law and malpractice suits will impact autonomous surgical robots. For computer error, there are many legal and ethical issues to overcome before we can establish a legal and regulatory framework that works effectively. It would seem somewhat easier to overcome these issues in military scenarios. Many of the earlier ideas, regarding robotics in health care, originated from past war experiences.

Attempts to save wounded soldiers can be extremely taxing on resources and medical care. In hostile environments, it would be preferable to deploy robotic surgeons in the proximity of wounded soldiers, whereby the wounded soldier can be loaded into a vehicle that has robotic surgical equipment on board, which can be operated remotely by a surgeon from the relative safety of a nearby Mobile Advanced Surgical Hospital.[6] The RAVEN[7] and Trauma Pod[8] prototype systems are examples of the (US) military interest in developing surgical robotics systems that can be deployed in combat zones and controlled remotely by surgeons. This triage saves having to deploy doctors, who have had years of costly training, into areas where they run the risk of being wounded, killed, or taken captive. It is safe to assume that military scenarios are less ethically and legally stringent than other kinds of hostile environments due to the life-and-death nature of military conflict. Although a hostile environment need not imply a military setting, the deployment of automated surgery in public health care or general hospitals involves considerations making it more problematic to devise adequate legal and regulatory frameworks.

3 | SAFETY-CRITICAL SYSTEMS AND CYBER SECURITY

Cyber security is another enormously increasing area of general concern that applies to surgical robots. Evidently, robotic surgery can be hacked, as was shown recently by Bonaci et al (2017).[9] This recent hacking exploit is interesting as it concerns seizing control of a teleoperated surgical robot, rather than dealing with issues potentially arising more speculatively for higher levels of robot autonomy. The greater the autonomy of the robot, the more avenues that exist for attack, and the questions of responsibility become rather different in such cases (See next section on Responsibility). Consider that someone could attempt to attack an autonomous robot, for example, by compromising the sensory input to such a system, in order to enable its AI to make a wrong/harmful/lethal decision. In contrast to seizing control of a tele-operated robot, where the attacker is in direct control of what the robot does, the more futuristic type of exploit works through deception of the robot's algorithms. In this case, there may be an additional onus on the software designers to certify that their algorithms are not vulnerable to perceptual tricks that would be readily detected by human perceivers.

In effect, autonomous surgical robots are safety-critical systems. To ensure reliability and obviate failure, the development process requires testing and certification. During operation, we require a live and complete record of all tasks and/or events, so that when unintended events occur, the record can be reviewed by a human operator (eg, medically certified surgeon).

The recent malware attacks on the UK National Health Service affected many medical devices.[10] A wide variety of the affected machines were embedded systems. One may argue that no software operating system can ever be considered secure and error-free. In many cases, software has not been updated since systems were deployed. Therefore, we question how to effectively perform patch management for robotic surgery, and we argue for raising awareness of such threats and considering defensive strategies at large.[11]

To experience any kind of cyber-attack during robotic surgery would be entirely unacceptable. Steps should be taken to prevent the possibility that software would be altered during surgery. Such a policy would also prevent the downloading of patches from the software developer during surgery, and the initiation of new surgery should be suspended until all patches have been applied.

4 | COMMON CRITERIA

For robotic surgery on living patients, Common Criteria security certification should be a requirement, similar to ISO/IEC 15408 for the automotive industry.[12] We must ensure that the robot can work within a safe and secure ecosystem. Reliable patch management could be vital to overcome vulnerabilities that may emerge—this would ensure capabilities of the device at all times. If no patch is available, and a known or an unknown vulnerability exists, that can be exploited, then we need more than simply patch management.

Surgical robots lack clear standards that define the meanings of safety, accuracy, and specific procedures to evaluate them, unlike their industrial counterparts for which standards such as ISO 9283 exist. A similar situation is also present in other emerging technologies such as autonomous vehicles. Standards such as IEC 60601 provide guidelines to define the degrees of autonomy of medical equipment, but its intent is aimed to assess the risk of a device rather than manage it. Paraskevopoulos et al (2010) presented five different metrics to assess the accuracy of a surgical navigation system.[13] We question how to effectively test autonomous robot surgeons running on software. One can encounter many difficulties in the specifics of robot testing. To date, most of the successful autonomous robotic surgeries have only involved animal subjects. For example, one pig study showed human surgeons were outperformed[14] by the Smart tissue anastomosis robot (STAR)—an autonomous system co-developed by Leonard et al (2014)[15] (See next section on Performance of autonomous robotic surgical systems and tort law). STAR uses several Food and Drug Administration (FDA) approved parts and dyes. However, the developers never pursued FDA clearance. Experimentation with robotic surgery on humans presents a very different proposition than work on animals or surgery training phantoms (eg, plastic human models). There are several other systems aimed at autonomous surgeries (or part thereof) in different areas (orthopaedic, radiotherapy, etc.). There also exist several robots, not designed for autonomy, which are adapted for autonomous tasks (Raven, DVRK).[16]

5 | LEGAL AND REGULATORY FRAMEWORKS

Although there is sound legal reasoning to support the use of autonomous surgical robots in hostile military environments, they are nevertheless a potential target for liability claims (See next section on Liability). In threat assessment of a military setting, there exist very strong rules of engagement and frameworks governing actions or accountability (See next section on Accountability). In hostile environments, one could assume these frameworks function to govern the deployment of surgical robots. They may provide the best option for surgical operations in regions where no human could safely reach (compared also to Marchant et al[17]). Although such deployments may be permissible despite lessened capability and ability than might be required in other contexts, state-of-the-art safety-critical systems are a requirement. When unintended events occur, one can clearly argue that the state-of-the-art was used at the time of development, and to an extent in military scenarios mistakes are highly likely anyway. This becomes problematic, however, when extending this emergency service in the context of civilian casualties—civilian deaths during autonomous robotic surgery could result in prosecution for manslaughter. We speculate that this is where the legally binding scenario might initiate. Under certain circumstances, soldiers can be legally obliged to provide medical assistance to civilian casualties. At that point, it becomes irrelevant whether they are a soldier or a civilian trapped (for example) in a minefield. Regardless of whether it was “remote robotic-assisted tele-surgery” or “autonomous robotic surgery,” any unintended event mid-surgery could lead to an interesting debate about legality, or about grey areas in the law. In their defence against legal action, human operators have the potential to argue that the developer did not deliver adequate software[18,19] or provided false statements regarding its proper testing.

6 | HUMAN-IN-THE-LOOP AND MEDICAL MISTAKES

In a military scenario, remote robot-assisted tele-surgery may experience transmission loss. Medical systems and radio frequencies (eg, ISM Band) under International Law are protected. However, this approach would not necessarily protect the system from an asymmetric threat, such as rogue non-military groups jamming or disrupting the signal. Thus, the option to activate autonomous surgery would be most advantageous. For disrupted tele-surgeries, autonomous support features can be activated to continue surgery and save the patient's life.[8] Emergency activation of autonomous features could occur during power-cuts, transmission-loss, (eg, Space station), jamming (eg, battlefield), or even a cyber-attack.

Switching to fully autonomous mode could be problematic if there is a computer error. Up until the activation point for autonomous operation, a human operator is responsible for the robot's actions. In autonomous mode, we question who is responsible for the actions of the robot since robots cannot understand blame, sanctions, accountability, liability, or culpability. Human operators can be discharged from military service, or even found guilty of manslaughter and sent to prison. So, when autonomous robots are deployed, where does the responsibility chain begin and end? Today, in respect of Culpability, we would expect the responsibility chain to continue to be incumbent on a human operator (See next section on Culpability). Makary and Daniel (2016) suggest that medical mistakes may be the third leading cause of death in the US.[20] When unintended events occur, human operators can stand trial because they are truly responsible for their actions. The question of the legal consequences of the damage caused by robots is paramount.

7 | JUSTIFYING THE NEED FOR AI AND AUTONOMOUS ROBOTIC SURGERY

There is ample legal reasoning to support the use of autonomous robots for surgery in remote or hostile environments, such as remote research stations, battlefields, or long-duration space flights. However, under high-risk conditions that may lead to contamination—machine learning (ML) enhanced robotic surgical procedures could be a better choice for surgical staff as well. AI surgery has potential to safeguard medical staff from infection and provide better diagnosis. For instance, we envision that current AI approaches, particularly ML, can enhance certain endoscopic surgical procedures (eg, robotic bronchoscopy) by augmenting the medical professional. The potential value in terms of performer safety is clear. ML enhanced robotic bronchoscopy can be tested in animal models. Currently, there are robotic systems for navigational endoscopic procedures,[21] but current methods still require initial access to the patient's body by medical personnel, and there exists the possibility for contamination of the atmosphere surrounding the patient. In 2018, Auris Surgical Robotics had its new bronchoscope approved by the FDA.[22] However, this bronchoscope (known as the Monarch Endoscopy Platform) does not include any autonomous capacities. There is a lot of room for further research and improvement with these types of new technologies. Tele-operated or autonomous robotic endoscopic procedures in combination with ML image analysis would be preferable. During the outbreak of the severe acute respiratory syndrome between 2002 and 2003,[23] health care workers used tracheal procedures to intubate the patient. Suctions, intubations, and resuscitations were some of the ways by which respiratory physicians became contaminated and infected by severe acute respiratory syndrome whilst also using nebulizers and high flow rates of oxygen.[24] In addition to disease spreading, the patient's breathing can also lead to the contamination of equipment. This is a prime example, where ML-enhanced autonomous robotic or robotic-assisted endoscopic procedures would be particularly desirable and beneficial for patients with high-risk pathogens. In this instance, autonomous features could enhance a tele-operated procedure (with the surgeon outside the operating theatre).

Programmed autonomous procedures for ENT could potentially reduce irritation and improve comfort among patients (especially when analgesic sprays are not used). Overall, the compelling benefits are mutual to both the patient and the doctor. For patients, certain endoscopic procedures used in ENT surgeries could be considered less life-threatening compared with open surgical procedures (eg, robotic laparotomy), and therefore a more suitable starting point for robotic surgery on human patients. Furthermore, ML-enhanced ENT surgery can safely support diagnosis from image-analysis (eg, images obtained in bronchoscopy or even flexible nasal endoscopy). ML techniques enhance the identification of lesions or features of disease that can potentially be overlooked by a physician. However, this brings us back to whether responsibility and final decisions should rest on humans. The ability to use ML-enhanced imaging is desirable when performing a biopsy—beyond classic approaches of ML.[25] Furthermore, ML-enhanced bronchoscopy provides many more benefits. Certainly, in the case of respiratory syndromes, there are applications which foster automatic approaches.[26,27]

ML facilitates automatic learning from data, the extraction of knowledge, prediction, and decision-making, without any human intervention. Automatic (aML) approaches show impressive success. For example, deep learning, applied for automatic classification of skin lesions and trained with large amounts of data sets, has shown similar performance to human dermatologists.[28] Statistical learning approaches can discover patterns in arbitrarily high-dimensional spaces, which is far beyond human capability.[29] Consequently, “big data” is not only beneficial for aML, but it is necessary to instruct the algorithms by many samples.[30] However, in pathology, we are often confronted with bias in data sets, and spurious patterns in data derived from affluent/racially/gendered, non-diverse samples. Here, we are confronted with the problem of having a small amount of data sets; therefore, our best learning algorithms suffer from unsatisfactory training samples, hence may not generalize well. Here, interactive ML (iML) is supportive: Humans can learn from very few examples and are able to understand the context; thus, a human-in-the-loop, particularly a doctor-in-the-loop, or even many doctors-in-the-loop (crowdsourcing)[31] can help to reduce complexity.[32]

A much more compelling argument for iML is that current approaches lack explicit declarative knowledge, therefore lacking transparency. The “black box” ML systems already used in many non-surgical applications defy easy causal analysis. In the medical domain, such lack of transparency does not foster trust and acceptance of ML among physicians.[33] Rising legal and privacy aspects, eg, with the new European General Data Protection Regulations, will make such approaches more difficult to use in the future. This is the case most frequently because automatic methods are often unable to explain why a decision has been made.[32] The most genuine and beneficial strength arises from combining both human and computer intervention into the decision-making process and consequently use of their joint efforts to achieve the best of both aspects overall.[34]

As humans, we have limitations that robots and machines do not have. Although, humans endeavour to do their best physically, at times, their surrounding environment and who they interact with can impact on them unknowingly. The consequences are detrimental to both the patient and the physician. These factors do not impact on a robot or machine; however, one can expect that robots usually need maintenance and can also experience breakdowns due to various reasons. Moreover, the FDA's Manufacturer and User Facility Device Experience database is filled with an alarmingly high number of medical robot failures.

So far, our justification has focused a lot on the benefit of AI enhancements and autonomous robotics for endoscopic procedures and ENT surgeries. However, any surgery with high infection risk or high likelihood of cross-contamination would benefit from robotic-assisted surgery and autonomous robotic surgery. Open surgical procedures (eg, robotic laparotomy) are considered more invasive and present more surgical trauma of exposure as they require more extensive incisions to gain tissue access. On the other hand, applying AI enhancements and autonomous robotics for minimally invasive procedures such as laparoscopic, thoracoscopic (both considered "key hole" techniques), or endoscopic procedures could be realised now or in the near future. In some cases, certain types of endoscopic procedures do not require external incisions into organs or cavities of the body (such as in Natural Orifice Transluminal Endoscopic Surgery). Moreover, due to technology advancements, robotic endoscopies have been increasingly developed and accepted.[21] We recognise that robotic-assisted endoscopy with remote control (operated by human surgeons) can also address the issue of infection or contamination. However, while autonomy may not be strictly necessary, it can provide an additional level of distance for valuable medical staff.

Other future realistic use cases are more applicable in emergency, military, or space exploration scenarios. These could include autonomous orthopaedic reduction procedures for minimally invasive or traditional "open" surgery. Wireless robots carrying (or installed with) portable digital radiography equipment could dissect tissue or expose fractured bone fragments and manipulate bone fragments without surgically exposing the bones or joints—be they fractured or dislocated. This is a procedure that can be accomplished by means of a short duration anaesthetic or nerve-blocking sedative. Fragments and/or bones can be repositioned robotically. This is achieved using casts, traction plates, or screws. Likewise, the use of implants which can be positioned either internally or externally.

8 | RESPONSIBILITY

We classify responsibility into the following: (1) Accountability; (2) Liability; and (3) Culpability. All three aspects are addressed when discussing responsibility for AI and autonomous surgical robots, be these civil or military patients.

I. Accountability

By accountability, we mean the capacity of a system to give an explanation for its actions. In one sense, accountability of a robot is simple to achieve since all it has to provide is a record of its input(s), internal state, and output(s). When an autonomous surgical robot makes a decision, humans need to know the decision was made for two reasons. On the one hand, the humans want to understand the decision and improve the system. Indeed, there may be technical problems that remain unsolved amongst other difficulties. In this regard, Wachter et al (2017)[35] believe that systems can make unfair and discriminatory decisions, develop bias, or have unexpectedly dangerous behaviour. On the other hand, humans may also want to understand why damage has occurred. There is therefore a strong link between accountability and liability (see following section on Liability).

With respect to accountability, the surgical robot may be responsible for recording its actions by using a "flight recorder" (similar to that used in aircraft) to track all aspects of telemetry and performance. Here the safety engineers' meaning of "flight recorder" is a recording device that is intended to provide data that will assist in failure analysis - often described in layman's terms as a "black box". However, this is very different from the ML sense of "black box" as something that is hard to analyse. Cooper et al (2013)[36] suggest the use of a black box to identify dysfunctions. Decker (2014)[37] proposes the same solution in a different context, as he tries to keep track of the modifications of the robotics system related to the learning algorithm of the robot. He also believes that *"the black box can make sure that the robotic system is something that is always understandable"* (Decker, 2014: 84).[37] This may be overly optimistic, however, because while having a read-out of the system's internal states may be necessary for reconstructing and understanding the causes of its behaviour, it may not be sufficient when the system is at a level of complexity which defies easy analysis. In other words, even with a black box recorder, the system's workings may still be a black box in the sense of having opaque inner workings. For example in "deep learning" with a "black-box" approach, the user can feed-in data as the input, then press/click a command button to obtain a result as the output. However, the user cannot retrace how this result has been obtained.

A system for recording the system's input(s), internal state, and output(s) can be implemented when designing and building the system. According to Wachter et al (2017),[35] emphasis on the design and the manufacture of the system tends towards an accountability requirement which necessarily bypasses current debates on the transparency and the explicability of AI and ML systems. Indeed, there are many cases in which transparency is not available as a solution, eg, trade secrets and the protection of other intellectual property rights. Today, American and European policies are divergent regarding accountability of AI.[38] The US is moving towards ethical design, education, and self-regulation, while the European Union is focusing on individual rights by adopting a regulatory approach in the Regulation of 27 April 2016 on the personal data.[35,39] It is for this reason that Wachter et al (2017)[35] state that *"Regulatory standards need to be developed to set system- and context-dependent accountability requirements based on potential bias and discriminatory decision-making and risks to safety, fairness, and privacy"*.

II. Liability

According to the current legal rules, resulting from legal issues that have arisen in the automotive area, a robot cannot be liable for its actions. Yet, this is not an immutable "ontological" truth: the current approach may one day be changed for pragmatic reasons. Sheriff (2015)[40] recommends a variation of Ugo Pagallo's "digital peculium"[41] liability scheme for "hard cases" where fully autonomous robots make decisions absent appropriate linkage to the original programmer and thus fall outside the scope of pre-programmed uncertainty. Situating Pagallo's "hard cases" in the larger abstractions of ethical and legal theory laid out by H.L.A. Hart and Ronald Dworkin, Sheriff (2015) concludes by considering whether determination of a right answer, or conclusive indetermination of any answer, exists for applications of legal liability to ever-increasing robotic autonomy.[40] As the scenario of full autonomy does not yet apply to robotic surgery, we distinguish between the application of current legal rules and the ongoing debate on whether such rules should be amended in the future, so as to focus on current legal rules. We do this either (1) because this is the appropriate level of abstraction for the present paper, or (2) because we deem new legal rules unnecessary in the field of AI and autonomous robotic surgery.

Under current law, the robot, even if autonomous, may not be held liable for its actions or its inactions in case of damage. It cannot therefore be ordered to compensate the victim. In this case, the damage caused to the patient by a surgical robot is imputed, either (1) to the manufacturer (if the robot has a manufacturing defect), or (2) to the operator (if the use of the robot is implicated or if it has made a medical error), or (3) to the person responsible for performing the maintenance or the robot's adjustments if the damage results from its failure. In regards to (1), one could also raise the question as to what happens if the manufacturer does not exist anymore (ie, went out of business). This question may be difficult to address judiciously due to the different conditions and assumptions that are likely to be relevant in each case. However, in situations where the operator has reduced control over the system, the operator may be less responsible or not at all liable for damage (Matthias, 2004).[41] In a war context, this scenario can be illustrated with the harm caused to a patient (soldier or civilian), for example, by a remotely operated surgical robot in a hostile environment (ie, when the robot has become less visible to the operator, such as during missile fire or an enemy attack).

However, the question arises whether the development of the complexity of robotic systems, methods of learning and reasoning, and especially autonomy, does not make it more difficult to determine liability. The operation of an autonomous robot, capable of learning new things for itself and adapting to its environment, could lead to unpredictable behaviour. It could therefore cause damage that would not be the result of its programming. For Matthias (2004),[42] the traditional rules of liability law would then be unsuitable for unpredictable machines because the human would no longer really have control of the actions of the robot. This same reasoning can be found in the European Parliament Resolution of 16 February 2017 on the rules of civil law in robotics.[43] For the Resolution, the legal framework currently in place is insufficient to cover the damage caused by robots with adaptive, learning, and autonomous capabilities, since they lead to some unpredictability in their behaviour (Paragraph AI). Paragraph AF states "*in the scenario where a robot can take autonomous decisions, the traditional rules will not suffice to give rise to legal liability for damage caused by a robot, since they would not make it possible to identify the party responsible for providing compensation and to require that party to make good the damage it has caused*".

This stance is not convincing for several reasons. First of all, in order for the behaviour of the robot to be unpredictable, a real break with the initial programming is required. This means that it is not enough that the situation is abnormal, because even the anomaly can be predictable (eg, malfunction of a sensor, incorrect programming of the robot, software error, hardware failure, etc.). For example, a personal-assistant robot has abnormal behaviour if it delivers a drug dose 100 times above the recommended dose but remains predictable insofar as this behaviour is related to its initial tasks (if it announced that it would rather be a poet than a medic, then that would be unpredictable!). Actually nowadays, it is difficult to find examples of really unpredictable robot behaviour in this sense. This is due to fact that when a robot learns new things on its own, it does so within the framework specified by the designers, for example, who provide it with objectives. In this discussion, a lot depends on the level of analysis. For example, it was entirely predictable that AlphaGo would produce moves that would help it to defeat its human opponents, even though, as widely reported in the press coverage (eg, Metz 2016),[44] some specific moves were very surprising to experts watching the games. Similarly, each robot is physically designed for certain tasks. Even if it has to learn new and unpredictable things, its dedicated architecture hinders the realization of acts incompatible with its material structure. Finally, if the behaviour of the robot is unpredictable to the point of causing damage, it must be deduced that the machine does not provide sufficient safety for humans, and perhaps, that the law may require it to be removed from the market. Moreover, in this case, the unpredictability of the behaviour even testifies to the existence of a defect of the dangerous robot for the human. This lack of security can then be sanctioned by the courts.

Court sanctions are likely to drive the development of the insurance market for autonomous surgical robots. The way that humans in modern societies deal with liability is via insurance (such as for a doctor when he/she practices medicine). In the current state of the law, as the robot is not liable, it is up to humans to incur liability by design, manufacture, or use of the robot to take insurance. The solution is identical in surgical robotics.

It is important to point out that the European Union could move towards a different path than insurance to solve the problem of responsibility for the damage of autonomous robots. The European Resolution of 16 February 2017 believes that a possible way would be to provide robots with a legal personhood to make them responsible for their damage so they can compensate victims. This text calls on the Commission, when carrying out an impact assessment of its future legislative instrument, to consider the implications of "*creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personhood to cases where robots make autonomous decisions or otherwise interact with third parties independently*" (Paragraph 59, f). There exist differences between legal systems in various countries (and parts of the world), and the European Resolution of Feb 16, 2017 may be the starting framework for robotics regulation, but it is not legally binding for nations such as the US or those in Asia. Furthermore, it is not immediately clear that it will become adapted as a de-facto industry standard.

Recourse to the legal personhood for robots is problematic for several reasons. First of all, in the Report of COMEST on Robotics ethics of UNESCO of 14 September 2017, Paragraph 201 explains about robots, "*it is highly counterintuitive to call them "persons" as long as they do not possess some additional qualities typically associated with* *human persons, such as freedom of will, intentionality, selfconsciousness, moral agency or a sense of personal identity*".[45] Secondly, even if the attribution of legal personhood derives from the model of the legal person and not from the natural person, the solution is untimely because a natural person always exists behind the legal person. This is not the case for the autonomous robot. Moreover, European Economic and Social Committee on AI of 31 May 2017 considers "*the comparison with the limited liability of companies is misplaced, because in that case a natural person is always ultimately responsible*" (Paragraph 3.33). Finally, even if the aim of the legal personhood of the robot is to set up a specific legal regime, for example in the form of a trust, the most serious risk is that of disempowering the manufacturer in the event of damage caused by the machine.

Insurance is managed through compliance and auditing of actions (see previous section on Accountability). Thus, we can say that accountability is a prerequisite for liability. The use of a black box recorder in surgical robotics to ensure the accountability of the robot would also play a role in legal disputes (Decker, 2014: 84).[37] Other authors also believe that the black box is an effective way of solving liability problems (Palmerini et al, 2014: 89).[46] Compliance implies best practice, and this would in turn apply to the development and maintenance of the system.

III. Culpability

Culpability relates to punishment, and a computer/robot cannot be punished as it has no concept of civil liberties. Culpability comes from the Latin concept of fault (culpa). The notion of culpability is paradigmatically linked to the notions of free will and conscience, although it is a matter of philosophical discussion whether freewill in an absolute metaphysical is an essential requirement for the attribution of culpable guilt. Regardless of the philosophical subtleties concerning the metaphysics of free will, it is common understanding that a person is guilty of a punishable act if either (1) the act is intentional and freely chosen, or (2) the person has not taken the trouble to avoid the act. It is the court that determines guilt. In this respect, a robot or AI program cannot, in any way, be found guilty for its acts, because it has neither free will nor conscience nor freedom in any sense currently recognized by courts. These concepts are foreign to it.

Culpability for autonomous robots will attach to the people who manufacture, distribute, own, and operate them, and the legal and regulatory framework under which they operate. In the event of an autonomous robot's infraction, only a human could be considered criminally responsible. Thus, a surgical robot that causes the death of the patient as a result of an error cannot be guilty of this offense. It will be necessary to determine which humans are culpable. If the robot is tele-operated, it is easy to think of the surgeon himself, unless the death is due to a malfunction of the machine (manufacturer). Signal loss, during tele-surgery,[47] raises many questions as to whether autonomous surgery should even be activated at all. Of course, it can be argued that an already open surgery should continue (eg, robotic laparotomy). However, arguing for autonomy as a "last resort" or "only option" runs the risk of having to address more questions. For example, if AI can take control, then what is the need for a human-in-the-loop, other than to provide responsibility? If the human is not necessary for the surgery itself, but is needed to provide responsibility, then perhaps there can be no justification for removing the human from the control loop.

9 | ETHICAL CONCERNS

It is important to cover the ethical domain and the need of informed consent or the respect of the individual's (patient or even the physician as operator) autonomy, elaborated in relation to robotic applied technology. The March 2018 special issue of the IEEE Technology and Society Magazine focused on "Social Implications of Robotics and AI,"[48] which illustrates the increasing interest and relevance of ethical questions, generally in AI, and specifically in AI for robotics. Research is urgently needed because the progress of AI for robotics is expected to mature—but the relevant remaining unanswered question is when?

Due to the fact that the military is fostering AI research generally (eg, DARPA), and particularly investigating AI for robotic surgery (eg, FORwARD), some previous work on ethics originates from that application domain, eg, Coeckelbergh (2011).[49] Coeckelbergh discusses a range of ethical questions in the military context generally, while Rosen et al (2011)[50] present a good general overview on the field of robotic surgery with some ethical considerations, and Lin et al (2011)[51] cover the standard work on ethical questions.

The rapid developments of robotic technologies in the last decades have naturally fostered the use of robotic devices for medicine and health, eg, for surgery, diagnosis, rehabilitation, prosthetics, and beyond. All of these applications give rise to ethical considerations, which are dramatically accelerated by the increasing fear of the ways that AI can harm human physical and mental integrity, and most of all, reduce human autonomy.[52] The latter needs a broader discussion. A general worry is that people will trust these machines to make decisions that are actually beyond the machines' capabilities, and this reliance on machines will effectively reduce human agency in ethically important domains. Steinert (2014)[53] provides a taxonomy of ethically relevant issues regarding robots in general, usefully dividing them into issues that arise from treating a robot as a sort of mechanical amplifying manipulator to achieve a specific goal vs treating the robot as a moral agent,[54,55] and thus perhaps imbuing the machine with more capabilities than it actually possesses. Analysis of the different kinds of cases leads to developing an ethical "impact factor" (such an impact factor could serve as a proxy measure for the broader field of robot ethics [roboethics],[56] which brings together ethical, social, and economic questions). Some researchers argue that the ethical impact of the new technologies will be quite low because they can be assimilated to existing paradigms for technological responsibility. For example, Datteri (2013)[57] addresses prospective and retrospective responsibility issues connected with medical robotics; it is suggested that extant conceptual and legal frameworks are sufficient to address and properly settle responsibility problems arising in connection with injuries caused by medical robots. Datteri (2013) emphasized that many issues of medical robots are nothing more than well-known general robotics engineering problems in disguise, which are routinely addressed by roboticists as part of their research and development activities.[57] Consequently, Datteri (2013) also emphasized that medical robotics does not necessarily raise novel ethical issues, as the same problems will occur in many other application domains as well.[52] Stahl and Coeckelbergh (2016)[58] argue that next to ethical analysis, classic technology risk assessment (a kind of reflection on the use of medical robots) is necessary, ie, "embedded" ethics in engineering projects from the very beginning.

In fact, ethical questions about surgical robotics vary according to the autonomy of the robot. Thus, if the robot is not autonomous, such as current surgical assistance robots, then there are fewer ethical problems. However, the ethical problems are still not absent. For example, it is essential to insist on training the surgeon on how to use the robotic technology, bearing in mind that each brand differs. Manufacturers of surgical robots often provide training for surgeons. There are also simulators, dual console systems, or Tele-mentoring (Canada). However, considering the number of cases in the US in which mishandling had caused harm to the patient,[59] current solutions may not be effective enough. It would be necessary to go further and set up real-world training in surgical robotics. It is also important to focus on patient information and consent in surgical robotics.[60] While robotic-assisted surgery has highly recognized benefits, the disadvantages in terms of increased duration of the procedure, aseptic problems, or lack of haptic feedback prove that the patient must be properly informed in order to consent to the surgical procedure. Whether in the US or Europe, the doctor must secure informed consent from the patient before carrying out the surgical procedure. The surgeon must therefore discuss all the possible risks, alongside the benefits, the alternative solutions, as well as the consequences of the operation. The physician who fails to fulfill this fundamental obligation may be held liable by the patient.

We enter into new uncharted territory when we switch the discussion from tele-operated ethics to autonomous robotic ethics. While it is important to encourage the designers of robots to adopt an ethical framework as part of their design process. It is also worth asking whether the AI/ML methods used to carry out complex tasks such as surgery can themselves be used to give robots the capacity for ethically appropriate responses to a variety of situations. An autonomous surgical robot in a battle zone may, for example, be required to triage among multiple wounded soldiers. Should such choices be made simply on the basis of likelihood of survival, or should the different capabilities of those who survive be considered? Hence, may it be necessary to attend to an injured helicopter pilot or a field medic, before attending to an ordinary unranked infantryman, even if the chances of successful surgery on the infantryman are slightly higher? Such decisions could be informed by what Wallach and Allen (2009)[55] call "top-down," ie, theory-driven, approaches to ethical decision making, or they could be shaped by "bottom-up," or data-driven processes that have been trained to emulate human decision makers. The authors argue that ultimately, hybrid top-down and bottom-up systems are needed. However, in the present era of deep learning and other advances in ML, bottom-up approaches seem to have the technological edge. The workings of deep neural networks and other advanced ML techniques are considerably more opaque than those of rule-driven, top-down systems. The technological landscape forces us to consider whether the benefits afforded by such systems outweigh the disadvantages for accountability, liability, and culpability that result. In other words, even if some capacity for ethical decision-making can be built into the robots themselves, the issues of human involvement in accountability, liability, and culpability remain the same.

10 | PERFORMANCE OF AUTONOMOUS ROBOTIC SURGICAL SYSTEMS AND TORT LAW

We must address questions, such as: should we evaluate AI and autonomous robotic surgery systems based on average performance? and, what if an autonomous system performs better in most cases, but could hurt a subset of patients that would be safer receiving conventional surgeon-led surgery? Currently, there is a robot-assisted surgery system that is enabling surgeons to undertake the most intricate of operations with very high dexterity. Known as the "da Vinci Robotic system," it allows these specialists to perform delicate manipulation in minimally invasive surgeries[61] (other similar systems include Mako, ROBODOC, CyberKnife, or Renaissance, and prominent robotic surgery companies include Intuitive Surgical, Smith & Nephew, Stryker, Mazor Robotics, Zimmer Biomet[62]). The term *surgical robot* refers to any electro-mechanical device which carries out surgical functions, including cutting tissue, cauterizing tissue, and suturing tissues, without direct mechanical feedback to a human operator. One could question whether robotic-assisted surgery devices such as the da Vinci[63] should be considered a robot or be considered a manipulator instead (ISO 8873). The da Vinci system is routinely used for operations such as renal surgery, urogenital surgery, heart surgery, colon surgery, and prostatectomies (even though many more procedures are currently under development). Nowadays, many assume there is approval for robotic-assisted surgery that requires human telepresence. However, this is not entirely accurate. In the US, and similarly in Europe and elsewhere, the da Vinci surgical robot is approved for use on living humans, but that robot is completely controlled by a highly skilled and specially trained surgeon. The surgeon sits in the same room where the robot operates, and at almost any point in time, if something goes wrong with the robotic procedure, the surgeon can revert to a traditional open procedure.

It is speculated that autonomous and semi-autonomous robotic methods will become accepted as a standard modality and thus revolutionize surgery.[64] These autonomous robots will form an essential element of cutting-edge technology. Enhanced ML capabilities can enable autonomous virtual robot surgeons and launch the next generation AI. It is worth investigating whether this could be comparable to human-level "Turing test" intelligence and consciousness.[65,66] Meanwhile, existing techniques will herald the first four generations (1^{st}—Stereotaxic; 2^{nd}—Endoscopic; 3^{rd}—Bioinspired; 4^{th}—Microbots) of surgical robots utilizing added autonomous (5^{th} generation) decision-making capability[64] even if being short of fully general AI.

In terms of autonomous robotic methods, we previously highlighted the Smart tissue anastomosis robot (known as STAR). This is the proof-of-concept for a vision-guided robotic surgical system.[15] The system's actuated laparoscopic suturing tool implemented image-based commands to perform specified tasks. It outperformed human surgeons in laparoscopic suturing, exhibiting higher accuracy, consistency, and faster speed.[15] They further demonstrated that STAR's in vivo supervised autonomous procedure is superior to surgery undertaken by expert surgeons.[14] Likewise, with robotic-assisted surgery techniques in ex vivo porcine tissues. Such positive results confirm the potential for autonomous robots. They are shown to improve the efficacy, consistency, functional outcome, and the feasibility of implementing surgical techniques.[14] However, the key question remains as to whether the robotic suture is better in terms of failure rates? If failure is defined as bursting pressure, then the answer is affirmative in this case.

For tort law, we raise this question: if robotic surgery eventually performs better than regular surgery, what effect will that have on medical malpractice standards? Do we then hold surgeons accountable for non-robotic surgery if it performs worse or do we hold them accountable for any kind of surgery they perform in general? This is an issue to be addressed and figured out for tele-operated surgical robots. Eventually, in the case of autonomous surgery, it raises another set of issues both in terms of (1) having to certify autonomous surgery, and secondly, (2) do we hold it to the same professional standards as human surgeons, or do we hold it to higher standards? There is a surprisingly small but growing literature on the issue of robot tort law. Hubbard (2014) argues that the existing legal system in the US is flexible enough to accommodate the technological developments, and he explicitly considers the case of da Vinci surgical robots.[67] Britten (2016) also attempts to address the subject in detail.[68] Arguably, if the average success rate for autonomous robotic surgery systems is higher than for human systems, this raises the bar for humans and robots alike. That is, if average robot performance is high, when something goes wrong in robotic surgery it is even more likely to demand scrutiny. The history of technology is full of such examples: today's car owners would no longer accept the level of unreliability of motor vehicles that was the norm before the last decades of the 20th Century; we have similarly raised our expectations for air travel. From a legal point of view, could the victim of an injury claim liability on the grounds that a surgical robot was *not* used? First of all, the answer depends on the country concerned. Indeed, in a poor country, it seems difficult to blame the doctor for not using a robot if the hospital does not have one.

Moreover, even in a rich country, not all hospitals are equipped with surgical robots because of their high cost. In this case, in the absence of a robot, and if a robot-assisted operation has greater benefits for the patient, it seems fundamental that the doctor informs the patient, or even refers the patient to a hospital that has it. It seems conceivable that a patient would blame the doctor for not providing this information before the operation. However, if the hospital has a surgical robot, the doctor has to give the patient the best possible care. If a robot-assisted operation has greater health benefits for the patient, then the doctor has to resort to robotic surgery, proving that advances in surgical robotics are having an impact on medical malpractice standards.

We focus on the legal issues relating to technology underlying *actual* surgical systems that are likely to be in use in the next 10 to 20 years, as well as the more futuristic concerns about systems that might one day be produced. We refer to "the learning algorithm of the robot"; however, there is a lack of evidence on learning algorithms used in autonomous surgical robots. Shademan et al (2016) refer to some abstracts in engineering conferences that propose enhancing motor control in a standard tele-operated surgical system (a da Vinci robot) by having it learn from surgeon's movements.[14] This implies that they learned from surgeons and coded their skills. The robot did not learn by itself nor by observing other robots.

In contrast, we suggest that "black box" recorder solutions to accountability are unlikely to be workable for sufficiently complex learning systems, although autonomous surgical systems such as STAR are not so complex as to preclude such an approach. Therefore, for STAR and others that work on similar principles, a black box may be acceptable. However, for other technologies that are not yet deployed for surgery, storing "snapshots" of the internal states of the system may produce only the ability to recreate them, but no adequate understanding of the exact causes.

Engineers tend to argue that this issue is irrelevant because such systems will outperform humans on average, so we can expect and tolerate some rate of unexplainable failures. This may be what lies behind their willingness to experiment, eg, with autonomous vehicles on open roads; engineers are already convinced that their vehicles are safer than human drivers, and the manufacturers are accustomed to absorbing the costs of product failures. Therefore, they can view this level of failure as something that lowers the cost of doing business, as ethically justifiable on general utilitarian grounds, and that does not require any special accountability. Perhaps, such a cost-benefit analysis is less acceptable for surgery because of the cultural and legal contexts within which medical practice occurs. The FDA Recommendations[59] for Patients and Health Care Providers about Robotically Assisted Surgery state that *"The FDA does not regulate the practice of medicine and therefore does not supervise or provide accreditation for physician training nor does it oversee training and education related to legally marketed medical devices. Instead, training development and implementation is the responsibility of the manufacturer, physicians, and health care facilities*." In June 2018, the FDA Software as a Medical Device Precertification Program[69] focussed on establishing processes for Software as a Medical Device technologies with AI and ML software functions. Additionally, the Program seeks clarification on elements or domains critical to evaluating the development of software functions using AI and ML algorithms.[69]

We recommend a unique proposition point: currently, most robots cannot (or do not) explain why they do something. Robots do not explain "what they think." This Explainable AI evolves from the iML with the human-in-the-loop. The component of explainability is in our opinion a huge benefit to the current robotics state-of-the-art and is novel on an international scale. This is urgently needed in the robotic field and can be used for education, but also for routine operations, since explainable robotics would raise trust among medical experts and patients undergoing operations that are supported by surgical machines with autonomous capabilities. It would be informative if one could specify what level of explainability for robot actions is needed to build trust among medical practitioners. For example, would showing the input features with higher weights in machines' decision-making process be sufficient? It seems that this suggestion would not work for sophisticated systems with many layers of processing as the weights on the input layer contribute only indirectly to the final decision. However, these are also somewhat empirical questions, since one cannot specify (a priori) what features contribute most effectively to human trust in robots. This area of robot explainability and trust is really only beginning to be explored. While there is some relevant literature on factors affecting human trust of robots,[70,71] there is a lot more essential work to be done.

11 | HEALTH DISPARITY AND REAL-WORLD APPLICATIONS OF AI AND AUTONOMOUS ROBOTIC SURGERY

We argued that the regulatory framework could be less strict when deploying AI and autonomous robotic surgery agents to hostile environments. However, this leads one to question whether it follows that the same holds true for deploying them to developing countries with fewer health care resources as well? When deploying an autonomous robotic surgical system abroad, be it built and/or approved in USA for example, this does not necessarily mean that it will be approved for its intended purpose outside USA. International legal and regulatory frameworks relate strongly to disaster or military use of autonomous robotic surgical devices, even in developing countries. Perhaps, the model used by humanitarian organizations can be used as a template to deploy surgical robots in humanitarian missions. The regulatory framework for medical volunteering varies greatly between jurisdictions. There are two different issues here. On the one hand, there is the permissibility of using prototype systems on populations for whom it is difficult to obtain informed consent. Here, it seems that the only ethically defensible position is to require the same safeguards for patients as would be required in the countries where this technology is being developed. On the other hand, there is the question of indemnification of those who are providing the services or operating the equipment. Some countries have regulations to indemnify claims for specific emergencies, such as natural disasters, or if the volunteers are citizens. Furthermore, some countries provide exemptions for actions "in good faith" while some require the volunteers to be registered and accredited.[72]

A report published by the International Federation of Red Cross concludes that "*jurisdictions adopt a rather piecemeal approach to the legal protection of volunteers in emergencies. Although there are some significant examples of how legislation protects volunteers, protection remains fairly inconsistent among countries. Some countries have legislation in place specifically applicable to either volunteering or disaster, whereas in others, volunteering in emergencies is covered by general laws.*"[73] Therefore, humanitarian organizations must adapt by training and certifying their volunteers and by defining the scope of each mission to comply with the heterogeneous legislations. The International Federation of Red Cross recommends that "*the most appropriate way to create an enabling environment would depend on the particular context of each country and a "one size fits all" approach will not work.*"[73]

In this light, it is very unlikely that the deployment of autonomous surgical robots would be any different. If surgical robots are deployed to support a humanitarian mission, they will likely have to be programmed to conform to the legislation and context under which they will operate. This significant consideration adds to the engineering hurdles of developing these systems: not only must these systems be programmed to execute difficult tasks, they must also be configurable to adapt to different jurisdictions.

12 | CONCLUSION

We provided a discussion on responsibility by classifying it into: (1) Accountability; (2) Liability; and (3) Culpability. All three aspects were addressed when discussing responsibility for AI and autonomous surgical robots, be these civil or military patients (however, these aspects may require revision in cases where robots become citizens, although this does not seem required in the case of Sophia robot being granted citizenship of Saudi Arabia[74]). We provided specific suggestions for managing reliability for the first two categories, but discussion on the third component, Culpability, is less clear, since it assumes capabilities that are far beyond the current state of technology. For the accountability and the black box, we highlighted the current reflections. However, all of our discussion on liability corresponds to what is already taking place in surgical robotics. Indeed, for the future, we have the unique position in the world of the European Parliament on the status of robots. This would apply to the autonomous robot in surgical robotics. For various international regulatory frameworks, autonomous robotic surgery must undergo a reliable certification process for standardization. The legal and regulatory framework can be less strict when deployed in space or hostile environments. Even so, the military still have a duty of care towards their soldiers. Moreover, the military may not welcome the idea of being dependent on a machine that they cannot fully control. In addition, many patients may not be prepared to hand over control of their life and integrity to electronic circuits, arms and fingers made of stainless steel. Besides legal issues for less hostile environments, jobs are also at risk. Professional groups may also lobby against the use of AI and autonomous robotic surgical devices (as seen in the case of J&J's FDA approved automated sedation system; this was intended to replace anaesthesiologists[75]). Furthermore, it may potentially be too early to discuss whether robots can have psychiatric problems,[76] as currently that question does not apply to anything on the horizon for surgical robots.

In recent years, robotics has seen tremendous advances, so has AI and ML. AI and autonomous robots could assist in time-consuming microsurgery processes and in some orthopaedic operations that can already last for 12 to 16 hours. For emergency scenarios (such as in conflict zones and mankind or nature catastrophes), surgeons would utilise electronic simulations and alternative strategies while operating in situ through stereoscopic lenses. Consequently, what we envision in the near future is that a surgical robot can learn and perform routine operative tasks, which can then be supervised by a human surgeon. This represents a surgical parallel to autonomously driven vehicles. Here a human remains in the 'driving seat' as a 'doctor-in-the-loop' thereby safeguarding patients. Finally, we recommend a unique proposition point, that Explainable AI is urgently needed in the robotic field. It can be used for education, but also for routine operations, since explainable robotics would raise trust among medical experts and patients undergoing operations that are supported by surgical machines with autonomous capabilities.

ACKNOWLEDGEMENTS

We thank the organization CNPQ (Brazilian National Council for Scientific and Technological Development). This entity provided support that was invaluable to our research. The authors declare they have no actual or potential competing interests.

ORCID

Shane O'Sullivan http://orcid.org/0000-0003-4287-7380

REFERENCES

1. DoDFY17 Medical Simulation and Information Sciences FORwARD Award. http://cdmrp.army.mil/funding/pa/FY17-DMRDP-MSISRP-FOR.pdf. Accessed 10 October 2017.
2. Franke U, López AM, Imiya A, et al. Autonomous driving. In: López AM, Imiya A, Pajdla T, et al., eds. *Computer Vision in Vehicle Technology: Land, Sea & Air*. Chichester (UK): John Wiley; 2017:24-54 https://doi.org/10.1002/9781118868065.ch2
3. Holzinger A. Human-Computer Interaction and Knowledge Discovery (HCI-KDD): what is the benefit of bringing those two fields to work together? In: Cuzzocrea A, Kittl C, Simos DE, et al., eds. *Multidisciplinary Research and Practice for Information Systems, Springer Lecture Notes in Computer Science LNCS 8127*. Heidelberg, Berlin, New York: Springer; 2013:319-328. https://doi.org/10.1007/978-3-642-40511-2_22
4. Kircher K, Larsson A, Hultgren JA. Tactical driving behavior with different levels of automation. *IEEE Transactions on Intelligent Transportation Systems*. 2014;15(1):158-167. https://doi.org/10.1109/TITS.2013.2277725
5. Sharkey N, Sharkey A. Robotic surgery: on the cutting edge of ethics. *Computertomographie*. 2013;46(1):56-64. https://doi.org/10.1109/MC.2012.424
6. Lanfranco AR, Castellanos AE, Desai JP, Meyers WC. Robotic surgery: a current perspective. *Ann Surg*. 2004;239(1):14-21. https://doi.org/10.1097/01.sla.0000103020.19595.7d
7. Rosen J, Hannaford B. Doc at a distance. In: *IEEE Spectrumm*; 2006.
8. Garcia P, Rosen J, Kapoor C, et al. Trauma Pod: a semi-automated telerobotics surgical system. *Int J Med Robot*. 2009;5(9).
9. Bonaci T, Yan J, Herron J, et al. "Experimental analysis of cyber security attacks on teleoperated surgical robotics". *in preparation to the ACM Transaction on Cyber-Physical Systems* 2017.
10. Collier R. NHS ransomware attack spreads worldwide. *CMAJ*. 2017;189(22):E786-E787. https://doi.org/10.1503/cmaj.1095434
11. Amoroso EG. Cyber attacks: protecting national infrastructure. *Elsevier*. 2012.
12. Schmittner C, Ma Z. Towards a framework for alignment between automotive safety and security standards. International Conference on Computer Safety, Reliability and Security. Springer 2014:133–143.
13. Paraskevopoulos D, Unterberg A, Metzner R, Dreyhaupt J, Eggers G, Wirtz CR. Comparative study of application accuracy of two frameless neuronavigation systems: experimental error assessment quantifying registration methods and clinically influencing factors. *Neurosurg Rev*. 2010;34(2):217-228.
14. Shademan A, Decker R, Opfermann JD, et al. Supervised autonomous robotic soft tissue surgery. *Sci Transl Med*. 2016;8(337):337. https://doi.org/10.1126/scitranslmed.aad9398
15. Leonard S, Wu KL, Kim Y, et al. Smart tissue anastomosis robot (STAR): a vision-guided robotics system for laparoscopic suturing. *IEEE Trans on Biomedical Engineering*. 2014;61(4):1305-1317.
16. Opfermann JD, Leonard S, Decker RS, et al. Semi-autonomous electrosurgery for tumor resection using a multi-degree of freedom electrosurgical tool and visual servoing. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2017;2017:3653-3659. https://doi.org/10.1109/IROS.2017.8206210
17. Marchant GE, Allenby B, Arkin RC, et al. *International governance of autonomous military robots*. 2011.
18. Beard TR, Ford GS, Koutsky TM, et al. Tort liability for software developers: a law & economics perspective, 27 J. Marshall J. *Computer & Info* 2009:L.199.
19. Scott MD. Tort liability for vendors of insecure software: has the time finally come. *Maryland Law Review*. 2008;62(2). Available at SSRN: https://ssrn.com/abstract=1010069) Accessed 10 October 2017
20. Makary MA, Daniel M. Medical error—the third leading cause of death in the US. *BMJ* (Online. 2016;353:i2139. https://doi.org/10.1136/bmj.i2139
21. Li Z, Chiu PW. Robotic Endoscopy. *Visc Med*. 2018;34(1):45-51.
22. The Monarch Endoscopy Platform. Available at: https://www.accessdata.fda.gov/cdrh_docs/pdf17/K173760.pdf. Accessed 10 October 2017.
23. Hung LS. The SARS epidemic in Hong Kong: what lessons have we learned. *J R Soc Med*. 2003;96:374-378. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC539564/. Accessed 10 October 2017.
24. Tomlinson B, Cockram C. SARS: experience at Prince of Wales Hospital, Hong Kong. *Lancet*. 2003;361(9368):1486-1487.
25. Ye Q, Qin L, Forgues M, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nat Med*. 2003;9(4):416-423.
26. Mori K, Ota S, Deguchi D, et al. Automated anatomical labeling of bronchial branches extracted from CT datasets based on machine learning and combination optimization and its application to bronchoscope guidance. *Medical Image Computing and Computer-Assisted Intervention–MICCAI*. 2009;2009:707-714.
27. Kim SY, Diggans J, Pankratz D, et al. Classification of usual interstitial pneumonia in patients with interstitial lung disease: assessment of a machine learning approach using high-dimensional transcriptional data. *Lancet Respir Med*. 2015;3(6):473-482.
28. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.
29. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. https://doi.org/10.1038/nature14539
30. Sonnenburg S, Rätsch G, Schaefer C, et al. Large scale multiple kernel learning. *Journal of Machine Learning Research*. 2006;7(7):1531-1565.
31. Holzinger A. Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics*. 2016;3(2):119-131.
32. Holzinger A, Plass M, Holzinger K, et al. A glass-box interactive machine learning approach for solving NP-hard problems with the human-in-the-loop. *arXiv:1708.01104* 2017.
33. Holzinger K, Mak K, Kieseberg P, et al. Can we trust machine learning results? *Artificial Intelligence in Safety Critical News*. 2018;112(1):42-43.
34. Holzinger A. Introduction to machine learning and knowledge extraction (MAKE). *Machine Learning and Knowledge Extraction*. 2017;1(1):1-20. https://doi.org/10.3390/make1010001
35. Wachter S, Mittelstadt B, Floridi L. Transparent, explainable, and accountable AI for robotics. *Science Robotics*. 2017;2(6):eaan6080. https://doi.org/10.1126/scirobotics.aan6080
36. Cooper MA, Ibrahim A, Lyu H, et al. Underreporting of robotic surgery complications. *J Healthc Qual*. 2013.

37. Decker M. Responsible innovation for adaptive robots. In: Battaglia F, Mukerji N, Nida-Rümelin J, eds. *Rethinking Responsibility in Science and Technology*, *Pisa*. Plus University Press; 2014:65-86.
38. Pagallo U. Algo-rhythms and the beat of the legal drum. *Philos Technol*. 2017. https://doi.org/10.1007/s13347-017-0277-z
39. Regulation (EU) 79 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), JOEU 4.5.2016, L 119/1–88.2016.
40. Sheriff KD. *Defining Autonomy in the Context of Tort Liability: Is Machine Learning Indicative of Robotic Responsibility?* 2015. https://doi.org/10.2139/ssrn.2735945
41. Pagallo U. The Laws of Robots—Crimes, Contracts, and Torts. Law, Governance and Technology Series 10, *Springer* 2013:I-XXV;1–200.
42. Matthias A. The responsibility gap: ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*. 2004;6(3):175-183.
43. European Parliament, n. Accessed 10 October 2017.
44. Googles AI viewed move no human understand. 2016. Available from: http://www.wired.com/2016/03/googles-ai-viewed-move-no-human-understand/. Accessed 10 October 2017.
45. World Commission on the Ethics of Scientific Knowledge and Technology. *Report of COMEST on robotics ethics* 2017. Available from: http://unesdoc.unesco.org/images/0025/002539/253952E.pdf. Accessed 10 October 2017.
46. Palmerini E et al. Regulating emerging robotic technologies in europe: robotics facing law and ethics. In: *D6.2 Guidelines on Regulating Robotics*; 2014. Available from: www.robolaw.eu. Accessed 10 October 2017.
47. Dumpert J, Lehman AC, Wood NA, et al. "Semi-autonomous surgical tasks using a miniature in vivo surgical robot". IEEE International Conference on Engineering in Medicine and Biology Society (EMBC). 2009.
48. Michael K, Bowman D, Jones ML, et al. Special issue on social implications of robotics and AI. IEEE Technology and Societyy Magazine. 2018.
49. Coeckelbergh M. From killer machines to doctrines and swarms, or why ethics of military robotics is not (necessarily) about robots. *Philosophy & Technology*. 2011;24(3):269-278.
50. Rosen J, Hannaford B, Satava RM. Surgical robotics: systems applications and visions. *Springer Science & Business Media*. 2011.
51. Lin P, Abney K, Bekey GA. Robot ethics: the ethical and social implications of robotics. *MIT Press*. 2011.
52. Datteri E, Tamburrini G. Ethical reflections on health care robotics. *Ethics and Robotics*. 2009;35-48.
53. Steinert S. The five robots—a taxonomy for roboethics. *International Journal of Social Robotics*. 2014;6(2):249-260. https://doi.org/10.1007/s12369-013-0221-z
54. Allen C, Varner G, Zinser J. A prologomena to any future artificial moral agent. *Journal of Theoretical Artificial Intelligence*. 2000;12(3):251-261.
55. Wallach W, Allen C. *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press; 2009.
56. Veruggio G. The birth of roboethics. In: Doe J, ed. *ICRA*. *IEEE International Conference on Robotics and Automation*, *Workshop on Roboethics*; 2005:1-4.
57. Datteri E. Predicting the long-term effects of human-robot interaction: a reflection on responsibility in medical robotics. *Sci Eng Ethics*. 2013;19(1):139-160.
58. Stahl BC, Coeckelbergh M. Ethics of healthcare robotics: towards responsible research and innovation. *Robotics and Autonomous Systems*. 2016;86:152-161.
59. Computer Assisted Surgical Systems. Available from: http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/SurgeryandLifeSupport/ComputerAssistedSurgicalSystems/default.htm#3. Accessed 10 October 2017.
60. Nevejans N. Traité de droit et d'éthique de la robotique civile (Treatise of Law and Ethics of Civil Robotics). *LEH Edition (in French)*. 2017;646-651.
61. Ashrafian H, Clancy O, Grover V, et al. The evolution of robotic surgery: surgical and anaesthetic aspects. *Br J Anaesth*. 2017;119(Issue suppl_1):72-i84.
62. Robotic Surgery Companies Medical Technology. Available from: http://www.investors.com/news/technology/robotic-surgery-companies-medical-technology/. Accessed 10 October 2017.
63. Toh JWT, Kim SH. Port positioning and docking for single-stage totally robotic dissection for rectal cancer surgery with the Si and Xi Da Vinci surgical system. *J Robotic Surg*. 2017.
64. Moustris GP, Hiridis SC, Deliparaschos KM, Konstantinidis KM. Evolution of autonomous and semi-autonomous robotic surgical systems: a review of the literature. *Int J Med Robotics Comput Assist Surg*. 2011;7(4):375-392. https://doi.org/10.1002/rcs.408
65. Ashrafian H, Darzi A, Athanasiou T. A novel modification of the Turing test for artificial intelligence and robotics in healthcare. *Int J Med Robot*. 2015;11(1):38-43.
66. Ashrafian H. Intelligent robots must uphold human rights. *Nature*. 2015;519(7544):391.
67. Hubbard FP. "Sophisticated robots": balancing liability, regulation, and innovation. *Florida Law Review*. 2016;66(5):1803-1872.
68. Britton D. Autonomous surgical robots. *Life Sciences Innovation: Law*. 2016;321. Available from: http://ssrn.com/abstract=2977943. Accessed 10 October 2017.
69. Software Precertification Program: Working Model—Version 0.2. 2018 Available from: http://www.fda.gov/downloads/MedicalDevices/DigitalHealth/DigitalHealthPreCertProgram/UCM611103.pdf. Accessed 10 October 2017.
70. de Graaf MMA, Malle BF. *How People Explain Action (and AIS should too)*. Arlington, VA, USA: AAAI Fall Symposium on "AI-HRI"; 2017. Available from: http://www.researchgate.net/profile/Maartje_De_Graaf/publication/320930548_How_people_explain_action_and_AIS_should_too/links/5a03106c0f7e9b3d401f6fc3/How-people-explain-action-and-AIS-should-too.pdf. Accessed 10 October 2017.
71. Ullman D, Malle BF. Human-robot trust: just a button press away. Proceedings of the Companion of the ACM/IEEE International Conference on Human-Robot Interaction. New York. 2017:309–310. Available from: http://dl.acm.org/citation.cfm?id=3029798.3038423. Accessed 10 October 2017.
72. Sharif E. Medical liability in humanitarian missions. *The Journal of Humanitarian Assistance*. 2014.
73. International Federation of Red Cross and Red Crescent Societies. *The Legal Framework for Volunteering in Emergencies*. Geneva; 2011. Available from: http://www.ifrc.org/PageFiles/125640/legalframework.pdf. Accessed 10 October 2017.
74. Pagallo U. Vital, Sophia, and Co. The Quest for the Legal Personhood of Robots. *Information*. 2018; 9: 230. https://doi.org/10.3390/info9090230
75. Pugliano J. *The Robots are Coming: A Human's Survival Guide to Profiting in the Age of Automation*. Berkeley, CA, United States: Ulysses Press; 2017.
76. Ashrafian H. Can artificial intelligences suffer from mental illness? *A Philosophical Matter to Consider Sci Eng Ethics*. 2017 https://doi.org/10.1007/s11948-016-9783-0;23(2):403-412.

How to cite this article: O'Sullivan S, Nevejans N, Allen C, et al. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *Int J Med Robotics Comput Assist Surg*. 2019;15:e1968. https://doi.org/10.1002/rcs.1968