Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-024-02806-y OPEN

Civil liability for the actions of autonomous AI in healthcare: an invitation to further contemplation

Ahmed Eldakak [1,2]✉, Abdulla Alremeithi[1], Emad Dahiyat [1], Moatasem El-Gheriani[1,2], Hassan Mohamed[1] & Mohammad Ibrahim Abdulrahim Abdulla[1]

There are already a number of autonomous robots that play a significant role in improving the quality of healthcare in different areas ranging from basic health diagnosis to complex surgeries. However, using robots and machine learning applications in the healthcare context poses concerns over liability for patient injury. This paper will thus attempt to investigate the potential legal problems that might arise if AI technology evolves or is commonly used in clinical practice. It also examines whether the traditional doctrines of liability can adequately address the liability for the injuries stemming from acts of autonomous robots. As such, this paper adopted both descriptive and analytical methodologies to explore the main focus of the study. while the descriptive methodology was used to spot light on various theories of liability, the analytical methodology was used to critically examine the main theories that have been advanced to deal with autonomous robots and predict the necessity of legal reform. Throughout this paper, the authors insist on the importance of distinguishing between robots in light of their degree of autonomy and then drafting liability rules depending on whether the action was done autonomously by an unattended robot or whether it was done automatically by an attended robot. Finally, the paper concludes with the proposal of a series of factors to be considered for the future regulation of AI Robots in the healthcare context.

[1] Private Law Department, College of Law, United Arab Emirates University, Al Ain, United Arab Emirates. [2] Faculty of Law, Alexandria University, Alexandria, Egypt. ✉email: a.eldakak@uaeu.ac.ae

Introduction

Artificial intelligence and machine learning are dramatically changing the healthcare industry. Healthcare providers are currently benefiting from their assistance. However, it is predicted that in the future, the amount of contribution by AI-enabled medical robots and other devices will increase. It is worthy of note that no robot nor any artificial intelligence, whether within or outside the medical context, has managed to reach the state of full autonomy. However, in this article, we are attempting to foresee the legal challenges that would arise in the event of realizing full autonomy and theorize some solutions to those challenges.

As such, this study would be divided into three main sections. The first section briefly discusses the concept of AI and its current application in healthcare, while the second section identifies the actors who are potential perpetrators of AI-related harm and explores how the liability rules would apply to AI systems. Finally, the third section explores the possible legal approaches to address the difficulties raised by AI-related harm: (i) ascribing personhood to autonomous AI; (ii) collective evidence and other procedural devices to ease the burden of proof; and (iii) the adoption of specialized compensation funds.

The development and integration of AI innovation in healthcare

AI can be defined as a computer program designed to initiate specific tasks like, or better than, humans without additional review or involvement by a human user (Griffin, 2021). In fact, the advanced generations of AI systems are capable of independent action rather than merely following instructions. They further exhibit high levels of autonomy, intelligence, and mobility, according to which their behaviors are not always completely predicted by their users. At that stage, they would no longer be mimicking human behavior or thinking, but they would be able to autonomously handle more complex tasks and become true "cognitive allies". This degree of sophistication led many scholars across the world to reconsider legal remedies that must be applied to robot torts (Guerra et al., 2021).

Presently, AI is being used in the robots industry. A robot refers to a machine managed by software capable of operating it to perform certain tasks. George Devol invented the first robot in 1954 called 'unimate'. Since that time, the use of robots has emerged in different fields, including healthcare. Through machine learning, AI has enabled robots to become more adaptable in their work and capable of solving problems more efficiently.

For instance, surgical robots enable surgeons to perform cardiovascular surgery without splitting breastbones by making small incisions between the ribs (Hodge, 2022). This represents an amazing breakthrough in such a kind of surgery. In fact, there are several popular surgical robots, including the famous Da Vinci robot.[1] It was estimated that the Da Vinci robot had been used in 1.5 million surgeries in 2021 alone (Xue & Liu, 2022). However, there have been two documented death cases resulting from the use of the da Vinci Robot despite the fact that it is not fully autonomous and it still requires constant input from surgeons. The first related to botched kidney surgery and was settled out of court (Allain, 2013), while the second took place in November 2015 at the Freeman Hospital when the Da Vinci robot damaged part of Mr Pettitt's heart (Dyer, 2018).

Moreover, robots are being used in rehabilitation. For example, patients recovering from disabilities caused by strokes frequently use rehabilitation robots for an extended period after the stroke (Habuza et al., 2021). Robots can guide patients to do the exercises that a therapist would do (Clark et al., 2019). Also, robots are being used in the disinfection of surgical rooms and medical centers in general. Disinfection robots use sensors to identify the surfaces and targets to disinfect (Fan et al., 2021). AI enables robots to determine the critical areas, the disinfectant dose, and the disinfection times. It has been argued that robotic disinfection is faster and more efficient than manual disinfection, which depends on the human conscience, training level, and skills of the worker (Fan et al., 2021). Another example is the Cyberknife robot, which autonomously uses radiation to treat tumors, including cancerous ones. This robot has a tracking system that is capable of tracking the tumor, which normally moves during respiration (Goyal et al., 2010). This ensures that radiation beams are targeted at the exact tumor position instead of where it was located moments ago, reducing damage to surrounding areas.

Applicable civil liability theories to AI-related harm

The use of AI-enabled medical robots in healthcare settings poses a host of legal and ethical questions, especially regarding the attribution of liability for the injuries and deaths caused by such AI-driven devices. After reviewing the current literature and relevant legislation on this matter, the following models of civil liability have been identified as potentially applicable to AI-related harm.

Strict liability. According to §§ 519–524A(1) of Restatement (Second) of Torts (1977), strict liability, also known as absolute liability, is normally applied in relation to unreasonably hazardous products and activities. According to this doctrine, liability exists whenever damage occurs without having to prove fault on the part of the defendant. This implies that liability could be attributed to the hospital using AI-based technology, regardless of whether the medical device's operations were authorized, intended, or controllable, and also regardless of whether the manufacturer exercised the utmost care in manufacturing, marketing, and selling the AI-enabled medical device.

This doctrine of liability may deter healthcare providers from using AI-enabled medical devices, including robots, and disincentivize companies from producing self-learning systems. While the adoption of this doctrine seems convincing in relation to hazardous products that can cause major bodily injury or death, the same cannot be said about medical devices that are primarily used to enhance the quality of treatment and reduce deaths and injuries (Dahiyat, 2011).

Negligence (fault-based) liability. Given that negligence is the typical cause of action in medical malpractice cases, it has been suggested as a possible basis to address the harm caused by AI-enabled medical and surgical devices (Chan, 2021). Furthermore, it has been asserted that liability based on negligence will become an even more viable option to handle harm brought on by AI-enabled devices (Rachum-Twaig, 2020), as the AI processes in those devices would develop to the point where the device can be fully autonomous and self-thinking, and capable of making decisions independently based on the information they gather.

According to § 3 of the Restatement (Third) of Torts: Liability for Physical and Emotional Harm (2010), the negligence doctrine establishes a framework for attributing liability to the person who fails to exercise due diligence that a reasonable person would have exercised under the same circumstances. Negligence in the medical field in the US law is called medical malpractice, and it is considered a form of professional negligence. Medical malpractice occurs when a physician or health care provider fails to provide a patient with the proper care, fails to take suitable action, or administers subpar care that results in damage.[2] For instance,

doctors are deemed to be negligent and not providing the appropriate care to their patients if they diagnose a patient's condition incorrectly, administer the incorrect medication or dosage to treat a medical condition, or make surgical mistakes like performing surgery in the incorrect location.

Courts have a tradition of considering human users as the locus of liability even if he was unaware of the operations of his program.[3] This might be because AI-based systems do not have legal personality, and hence they cannot be subject to the same laws as persons and cannot be held liable for damages. They are merely viewed as tools for doctors to provide better and more advanced healthcare services to patients. However, with the advancements that we currently witness in AI technologies, we may no longer hide the possibility that one day fully autonomous medical robots will surpass human skills and work on their own without the need for human supervision. At that point, AI-enabled medical devices must not be viewed as mere tools, and their legal status should be carefully re-evaluated.

Applying medical malpractice elements to AI-enabled medical devices. To establish a medical malpractice claim, the plaintiff must prove that the defendant owed him/her the duty of care, that this duty has been breached, and lastly, that this breach caused damages to the plaintiff (Raskin, 2018). The first element of a medical malpractice claim is the duty of care. This refers to a healthcare professional's commitment to delivering treatment to a patient that satisfies medical community standards. The physician-patient relationship establishes this responsibility of care. To prove the existence of a duty of care in a medical malpractice case, the plaintiff must show that a doctor positively acts in the plaintiff's case by looking at, diagnosing, treating, or consenting to do as such (Bush, 2010).

Moreover, the duty of care can be established between the patient and the physician even if they did not meet in person. For instance, in *Hand v Tavera* (1993), Mr. Hand, who was a member of Humana HMO, went to the emergency in a hospital approved by HMO, complaining of a 3-day headache. The emergency doctor examined and reviewed his medical history and decided to admit him to the hospital and called Dr Tavera, who was responsible for approving patients' admission to the hospital, but he refused to admit Mr. Hand to the hospital and directed the emergency doctor to treat him as an out-patient and was sent home. Soon after Mr. Hand reached home, he suffered a heart attack. Mr. Hand brought a lawsuit against Dr Tavera, but the latter declined any responsibility, claiming that there was no physician-patient relationship between them because he did not see Mr. Hand in person. The court found that the physician-patient relationship existed in this case and stated that 'when a patient who has enrolled in a prepaid medical plan goes to a hospital emergency room and the plan's designated doctor is consulted, the physician-patient relationship exists and the doctor owes the patient a duty of care'.

The duty of care element cannot be established with the current medical and surgical devices available because they cannot work on patients on their own, and they still need to be operated by a physician or surgeon, who will decide on the action to take to treat the patient (Allain, 2013). However, we think that the duty of care can be established on future fully autonomous medical robots because they would be capable of having direct interaction with patients without depending on human intervention or guidance. Also, it is easier for patients to prove the physician-patient relationship on medical devices, as it can be simply traced by the electronic medical records (EMR) that are generated by the medical system itself.

The second element is a breach of duty. This happens when a healthcare provider fails to deliver treatment that complies with accepted medical standards. To establish a breach of duty of care, a plaintiff must prove that the healthcare professional's actions or omissions fell below the medical community's recognized level of care, or what is referred to as the 'standard of care'. In the US, it is a question of fact for the jury to decide whether a doctor breached the duty of care, by providing improper care to the patient after hearing an expert testimony from a physician (Kessler, 2011). As we mentioned earlier in this paper, fully autonomous medical robots and other devices can make mistakes and breach the duty of care. However, sophisticated AI-driven medical devices should be held to a higher standard of care than human physicians because they surpass the human brain in the amount of data they can store and how fast they can make decisions based on this data. Also, they are more accurate than human physicians in diagnoses of diseases and can perform very accurate maneuvers during surgeries that human surgeons cannot match. Therefore, in deciding whether or not AI-enabled medical devices have breached the duty of care owed, they should be judged by the level of care that other comparable sophisticated medical devices would provide in similar circumstances and not be compared to the level of care of the average physician or surgeon in similar circumstances. If we applied the current rules of deciding the breach of duty of care in medical malpractice cases on AI-driven medical devices, which typically require expert testimony from a human doctor, the jury will decide in favor of medical robots as they make fewer mistakes than human doctors would make. Thus we need a new model of the standard of care in medical malpractice cases filed against AI-enabled medical devices.

The third and fourth elements of medical malpractice, respectively, are the proximate cause and damages. To establish this element, the injured patient must prove that there is an adequate connection between the breach of the duty of care by the healthcare giver and the damages that are caused by such a breach (Sonny, 2009). Because proving causation in medical malpractice cases is a complicated issue, courts have adopted a 'relaxed causation' rule (Stein, 2012). According to this rule, even if an expert is unable to specify the precise degree to which the doctor's act or omission decreased the patient's chance of a better outcome or increased the injury, the jury may still find the testimony sufficient if it can be inferred that the doctor's actions decreased the patient's chance of a better outcome or increased the injury (Stein, 2012). Therefore, a patient's prior illness complicating the causation inquiry does not allow the doctors to avoid liability, and they may be deemed negligent in such situations if their conduct 'increased the patient's risk' of disease or harm (Frierson, 2019).

Establishing the causation element in medical malpractice cases against AI-enabled medical devices might be easier and less complicated because patients can retrieve the EMRs that are generated by the medical device and see exactly what actions the medical device took to treat the patients and what medical knowledge and cost/benefit analysis it used to base its medical decisions on. Also, EMR in intelligent medical devices may also solve the complicated issue of who is responsible for the harm that is caused by the medical device, is it a mistake in the programming of the device by the developer? Is it an erroneous decision taken solely by a self-learning and self-thinking device? Is it a defect in the device caused by the manufacturer? Did the patient provide the wrong information to the medical device? Or is it caused by one or more of the previously mentioned situations? These issues will be discussed further in the final section of this article, where we consider the implementation of a 'black box' in intelligent medical devices as a way to gather information about their actions which can be used in litigations. However, it is necessary before moving forward to recognize that mistakes in AI medical systems do not always reveal themselves

clearly and it is often difficult to detect their source or to prove their occurrence. This is especially true if one thinks of the fact that the robot's algorithm is, to some extent, unexplainable or unlikely to be well understood in all cases.

The development of standard of care in the era of AI. The standard of care that should be followed by healthcare providers to avoid being held liable for medical malpractice is not a fixed one, but it keeps on changing to catch up with the developments in the medical field regarding new medical methods that are being used by doctors to provide better and safer treatments to patients (Price et al., 2019). For example, some states in the US adopted clinical practice guidelines to help define the standard of care in healthcare, however, they were outdated shortly after being adopted and no longer can be used in malpractice cases due to the fast-paced advances in medical research and innovations (Cooke et al., 2017). With the numerous benefits that AI provides in the medical field, there is no doubt that it will affect the standard of care and be a critical part of it. Therefore, in the future, doctors who fail to use AI technologies in their diagnoses, treatments, and surgeries will be held liable for breaching the standard of care (Griffin, 2021).

Products liability. Product liability generally applies to manufacturers of finished products and manufacturers of component parts integrated into a finished product. It may also apply to, among others, importers, wholesalers, and retailers of the product. However, the question of liability in the case of using AI medical devices and systems is far from simple, as it is still unclear whether product liability law applies to algorithms and whether AI robot coincides with the hardware or with the software. The prevailing rule on what qualifies as a 'product' for the application of products liability law is to be found in Restatement (Third) of Torts: Products Liability (Restatement 3d) § 19, which strictly defines a product as 'tangible personal property distributed commercially for use or consumption.' Conversely, services and intangible information and ideas do not qualify as products for the purposes of product liability.[4] This strict view has been recently extended to AI-enabled software, according to *Rodgers v Christie* (2020).

An alternative approach is for courts to seek guidance from the US Food and Drug Administration (FDA) in relation to the use of AI-enabled medical devices and systems. The FDA does not govern the practice of medicine (i.e., service), but it does regulate medical devices,[5] and if self-thinking healthcare systems are classified as such in a particular case, product liability, at least the preliminary question of classification, is not outside the realm of possibility in the future. In acknowledgement of the integration of AI-based technologies in the healthcare sector, the FDA, in January 2021, published an action plan to modernize its paradigm of medical regulation to cover adaptive AI technologies.[6]

Even if we were to assume that AI-driven medical systems would qualify as 'products' it would be difficult to determine what category of 'defect' applies to AI-related harm. Restatement 3d recognizes three product defectiveness types: manufacturing defects, design defects, and inadequate instructions and warnings. There is a tentative consensus amongst commentators that problems concerning the design and coding of the AI algorithm should be treated as a design defect as it will normally impair the entire product line (Hodge, 2022; Duplechin, 2018). However, if this proposition is accepted, then the plaintiff would face the daunting task of establishing the presence of a design defect. This is because the principal test is whether an adequate alternative design would have reduced the expected risks of harm posed by the product at a reasonable cost and, if so, whether the omission of the alternative design by the seller or a predecessor in the distributive chain rendered the product not reasonably safe. In short, the test of proving the availability of a technologically feasible and safer alternative AI-enabled medical device would prove too burdensome, highly contentious, and uncertain and subject the plaintiff to exorbitant legal and expert fees.

Besides the difficulty of satisfying the 'reasonable alternative design' test, the product liability theory may hardly be applied when the AI-enabled medical devices are modified after their manufacturing or programming via open-source software. It will be difficult here to adopt the product liability principles and conclude that the original product originally sold caused the damage in question. What complicates the matter further is that AI medical devices might operate autonomously according to their own experience and then commit errors independently of those who produce, sell, or use them. Holding manufacturers responsible is unfair, as long as no one of them has done anything that specifically caused harm nor could have prevented or foreseen it. Manufacturers shall only be responsible for the defects that may be attributed to the design process or for inadequate instructions when there is a foreseeable risk of harm posed by the product. Bearing their unpredictable nature in mind, AI-related risks cannot be covered by product/design defects or duty of warning and instruction doctrines.

Vicarious liability. The doctrine of vicarious liability attributes an indirect and secondary liability for an employee's misconduct on his employer if such conduct was committed during the course of employment. Similarly, healthcare providers using AI-driven medical devices should be held responsible for any damages caused by autonomous systems during deployment (Glavanicová & Pascucci, 2022). The rationale for this doctrine of liability is to distribute costs for any injury caused by AI among hospitals or other relevant parties to fully compensate a patient.

However, the application of this doctrine in the area of AI and robotic technology has been subject to criticism since it ignores the fact that AI-driven systems, especially intelligent robots, lack legal personality or capacity and have no independent financial status to make reparations for damages it causes. On the other hand, it is extremely difficult to fully trace connections between deployment or environmental inputs and changes in their algorithms and approaches. Highly autonomous medical robots and systems would not be subject to the full control of human or corporate deployers, and hence their outputs are not necessarily completely determined by inputs received.

There is a lot of confusion out there about whether or not an AI robot stands fully under the hospital's control and whether it can be considered an "agent", "physician", or "employee" of a hospital in order to attribute responsibility to the "principal", in this case, a hospital (Schweikart, 2021). However, there is no reason why such intelligent robots and devices could not someday be granted some kind of subjectivity or a minimum level of legal capacity. The arrival of that day will depend on what technology can do to improve the reliability, sophistication, and autonomy of AI-enabled devices in the medical field. Once we reach that day, the most logical option might be to acknowledge some kind of legal personality to such robots so that we can treat them in the same manner as the law treats human employees, with some exceptions where the algorithmic processes impose particular requirements (Dahiyat, 2011).

Observations on possible solutions

In the discussion above, we have addressed the possible application of the different theories of liability to injuries attributed to the use of AI in healthcare. Each raises its own set of doctrinal

and practical questions regarding its applicability in that scenario. In addition to those theories, and in an attempt to solve or circumvent the difficulties they may face, several proposals have been advanced. We focus here on three of those proposals, mainly the proposal to grant an AI-driven medical robot a legal personality, the proposal to include a 'black box' to collect sufficient data regarding the operations of such robot for evidentiary purposes, and the reliance on no-fault insurance. We discuss each proposal and its advantages and disadvantages.

Granting the medical robot legal personality. This proposal has been thoroughly discussed by many scholars,[7] although no final judgment can be made until AI robots reach a level of sophistication and reliability at which it becomes desirable to treat them as legal persons who are capable of taking the initiative rather than waiting to be told what to do. According to that proposal, the medical robot will be assigned an independent legal personality and would be subject to liability for its actions, to some extent, just as a natural person would, and hence it will not be endowed with an unlimited power to bind its owners or operators.[8]

There are advantages to such a proposal. The use of AI has raised the dilemma of holding the different players responsible for actions that are only attributable to the AI's own ability to learn from its environment and to act on its own initiative. It is this ability that makes the use of AI most useful in medical diagnoses and surgical operations, but it is also the one that raises moral concerns about holding physicians responsible when they are not in control of the diagnosis or operation. Additionally, holding the medical device liable shields owners and operators from liability or limits their liability. It also facilitates judicial actions and collections of compensation by targeting the medical device itself.

From a practical point of view, however, the proposal seems less advantageous on different levels. First, a medical robot, as a legal person, still needs to have funds of its own if it is to bear the financial burden of liability, such funds will still need to be provided by someone. Second, hospitals may still be held accountable for the autonomous actions of their medical robots and systems on vicarious liability basis. Third, physicians may also still face liability based on negligence if they fail to supervise or take necessary actions when a medical device malfunctions. Fourth, the veil of the personality of the robot may still be pierced, just like the veil of a corporation could be pierced under certain circumstances.

Two more factors should have a decisive effect on assessing the proposal. The first concerns incentives (Guerra et al., 2022). Liability rules are certainly meant to compensate the injured person. For this, we have rules regarding compensatory damages, restitution, and so on. However a large body of tort law focuses on creating the right incentive for encouraging people to avoid and reduce risky behavior. Law and economic analysis focus on calculating who can better avoid the risk in a more efficient way. This whole incentive issue may be lost if we simply attributed the actions to the robot and let it bear the financial consequences. This needs to be considered, especially looking into the extent AI will actually be able to calculate such risk and respond correctly to the liability incentive.

The second is evidence of causality. Although granting the medical robot or system personality will make it easier to sue the robot itself, the effect on litigation is not readily clear. To hold the AI-driven robot liable, we still have to attribute the harm to the robot's own actions. Incorrect diagnoses or mistakes during operations causing patient injury may still be caused by failures on the part of the hospital, physicians, and staff, or they can be attributed to problems with manufacturing, programming, or the data fed to the system. Indeed, it is not clear whether the presence of this additional legal personality will make litigations in these cases easier or more difficult, with one more person to litigate and defend, assuming that we will give the medical robot or system the right to be duly represented by a lawyer.

This proposal has therefore caused a wide debate, especially because robots were essentially built by humans as artifacts that lack the cognitive ability to intend or express intent. Such argument against considering robots as intentional legal persons has its roots in Searle's theories which treat thinking and other relevant cognitive abilities as non-physical processes arising only from the soul and not from any material process (Searle, 1980). This line of thinking, however, can be subject to much criticism as there is still no decisive evidence that intentionality is exclusively linked to an organism's soul or that the brain process cannot be replicated or at least explained by the computational model. Therefore, to establish whether an AI robot truly possesses cognitive states, one needs to consider whether the concerned robot is operating "rationally" to accomplish its objectives. If so, then we can attribute a cognitive state to this robot, and we can say that it is intentional or that it should be considered so. Despite the initial allure of this scenario, the reality is that we have not reached that point yet. Thus, it might currently be useful to explore other approaches addressing the issue of how to deal with unintended outcomes of AI robots.

Collecting electronic evidence. No matter how we look at liability cases, the issue of evidence is always present, especially regarding causality. At an early stage of the discussion of medical robot or system responsibility, there was a proposal to require a black box to be present in all AI-enabled robotic operations or disease diagnoses. The idea, similar to the black box of an airplane, aims to collect sufficient evidence about how the operation or diagnosis was conducted, which can then be used in court. Once this black box is opened and scientists are able to detect what exactly occurred at the decision-making point, this may play a greater role in establishing evidence of either causality between the action and the injury, or the degree of negligence or fault, if any. Even if no black box exists, the reality of the presence of EMRs greatly facilitates accessing data about patients and operations, which provides a plethora of data in case of any liability litigation.

However, the use of this data raises concerns regarding the privacy of patients' personal data. Such concerns revolve around the issue of who has the right to access the medical data and whether the patient's bringing a liability case suffices as an implied consent for accessing such data for the purpose of defending the different actors. On the other hand, AI systems require vast amounts of sensitive personal data, and if this data falls into the wrong hands, it can cause considerable damage. This may also pose potential issues, particularly with regard to identifying the one who has the final say on when, how, and to what extent that data may be used or processed. The processing of medical data must be lawful, transparent, and limited to the purpose(s) for which such data was collected (Regulation (EU) 2016/679 (GDPR), art. 5). Appropriate safeguards must also be in place to maintain privacy and patient data. It should be noted, however that data protection legislation safeguards the personality of data subjects, not their property. This implies that any relevant use or processing of data would not necessarily be conditional upon a patient's verification or confirmation. What is needed is just to provide patients with an opportunity to know how and why their data is being collected. Patients should also be given direct access to their data and be aware of the privacy policy

surrounding the use, storage, and transfer of biometric data obtained from AI technology.[9]

A different issue that has not been sufficiently addressed is the effect of medical recording on the actions of different players. Supposedly, the recording does not affect the robots' or systems' actions, but it may affect the decisions taken by attending physicians, surgeons, and staff. Is it even possible that the AI can manipulate data and recordings, or that a hacker or malignant virus can still affect the robot's or system's actions while at the same time deleting all of its traces? Those threats will require further research and technical assessment and will certainly raise legal challenges.

US state legislatures and courts could seek guidance from the latest European Commission's proposal for AI Liability Directive, which attempts to minimize the need for evidence altogether and especially to avoid the black box effect. According to Article 3 of the proposed directive, the claimant 'must present facts and evidence sufficient to support the plausibility of a claim for damages', this will trigger a duty on the part of the defendant to 'disclose evidence at its disposal ordered by a court', and if the defendant fails to do so, the article then creates a rebuttable 'presumption of non-compliance with those duties of care which that evidence was intended to prove.'

Article 4 of the same proposal then goes further to establish a presumption of causal connection 'between the fault of the defendant and the output produced by the AI system or the failure of the AI system to produce an output under certain conditions, (1) proof or presumption of fault, (2) the reasonable likelihood that such 'fault has influenced the output produced by the AI system or the failure of the AI system to produce an output, and (3) the claimant's demonstration that 'the output produced by the AI system or the failure of the AI system to produce an output gave rise to the damage.'

No-fault insurance. Another approach that aims to mitigate and overcome the liability issue, is to separate compensation of patients from establishing the liability. This requires the health care system or the medical insurance to indemnify the patients regardless of the establishment of liability. The insurer will then have the option to litigate to establish who is actually liable and pursue their rights to recover. Because such a system does not require the establishment of fault, it makes it easier and faster to obtain compensation, and it also removes the incentive to cover for error, making it easier for the patients and their care providers to discuss, explain, and reconsider.

An example of such a system is New Zealand's Accident Compensation Act (2001), which was amended by the Injury Prevention, Rehabilitation, and Compensation Amendment Act (No2) (2005). The health care system compensates patients suffering injury resulting from *medical misadventures* with no need for litigation (art. 34). The patient suffering the accident may directly lodge his claim with the Accident Compensation Corporation (The Corporation).[10] The claim is either the cover of the injury suffered, certain entitlements enumerated by the law (art. 48, 69), or a combination of both. The proceedings under the act are the only recourse for the injured person, who may not bring any other proceedings to collect damages caused directly or indirectly by injury covered under the act (art. 317) (Wallis, 2013). The system's inclusion of "entitlement" goes beyond merely covering damages caused by an injury to provide social and professional rehabilitation.

A different example is the Swedish system. This system is more specific as it targets medical injuries through the Patient Injury Act. The funding of the coverage comes through private insurance (Hellborg, 2019), either through the employer who carries insurance with an insurance company or through direct insurance by the patient with the Patients Insurance Association. One important category of injuries, for present purposes, is 'injuries caused by faulty equipment or wrongful use of such equipment (Hellborg, 2019). Yet, the system does not solve all the problems, although the act does not require fault and is not concerned with contributory negligence by the injured, it still depends on the existence of causality between the action and the injury, as 'there must be a causal relationship between injury and care, and the injury must also not be a foreseeable consequence of the patient's basic illness (Hellborg, 2019). One final difference between the Swedish and New Zealander models is that the Swedish law does not preclude the patient from suing the parties responsible in court. The system's aim is to provide patients with speedy compensation. Those systems are already in place and do not have to be established specifically for AI, although the advent of AI may raise new questions for them.

Although the laws go a long way in guaranteeing coverage without the need to prove fault and causality, they have not been written with autonomous AI in mind and it may prove necessary to revise them as the need arises. For example, the New Zealand law, which is most liberal in its conditions of coverage, uses the criteria of "treatment". According to Article 32 of the Accident Compensation Act, it suffices that the injury is caused by treatment to constitute Treatment Injury. This dispensed with the previous requirements of the occurrence of medical misadventure, which depended either on the presence of negligence (medical error) or on the result being a rare or severe adverse event (medical mishap), and accordingly excludes any need to investigate any source for that effect (Wallis, 2013). Yet, the article goes on to require that the injury must be "not a necessary part, or ordinary consequence, of the treatment, taking into account all the circumstances of the treatment, including…ii. the clinical knowledge at the time of the treatment". Taking that AI is a self-learner, our conception of what constitutes "clinical knowledge" may need to expand to include the robot's knowledge. Another aspect lies in Article 25 of the same law. This article addresses the coverage of *Accidents* and excludes from insured accident occurrences that "constitute treatment given, i. in New Zealand, by or at the direction of a registered health professional or, ii. outside New Zealand, by or at the direction of a person who has qualifications that are the same as *or equivalent* to those of a registered health professional". Although the article is clear as to the treatment given "by or under the direction", it still requires that the health professional be a "person". The issue of autonomous AI may force us to consider who is a *person*, or whether at least we can consider a robot *equivalent* to a health professional without being a person.

Depending on the details of the system, different problems may arise. One of them is the topic of incentives and making sure that negligence does not go unpunished (Hellner, 2001). This is often covered through disciplinary actions by regulatory bodies of the medical professions. Another is the funding of the system and the increased burden of automatic compensation in the case of public funding. On the other hand, an insurance scheme might not present sufficient cover since the liability of AI robots could only be claimed exclusively on the patrimonial level or according to the sum of the insurance. Besides this very limited protection, the different policies of insurance could sometimes lead to uncertainty that might create various difficulties, and generate conflict between insureds and the insurance companies on what should be excluded and what should not. This is not the place to detail these problems, but the reference here does confirm the need for a deeper investigation to provide a more straightforward solution to the problems raised by the intervention of AI in medical treatment.

Law and economics and incentive-based solutions. As noted above in "Granting the medical robot legal personality" and "No-fault insurance", the issue of incentives plays a crucial role in liability rules. Any proposal will have a different impact on the incentives of different players. Thus, a regulatory impact assessment of the rules is required before we embark on promulgating new rules. In other words, we have to ask ourselves: "to whom should we direct our incentive?" Is it more important to direct our incentives to patients, physicians, hospitals, insurance companies, the manufacturers of robots or to their developers? This in itself depends on answering two questions: (1) Who can reduce the risks associated with the use of the robot the most? and (2) Would incentivizing one party on the chain reduce the incentive for another link on the chain?

Law and economics scholars have attempted to explore this aspect of the problem and to provide an economic perspective.[11] Talley, for example, suggests that product liability may tend to take a more central stage the more autonomous robots become (Talley, 2019). However, he advances to propose a model of "multilateral investment" by potential players involved. Another important solution is assigning residual liability where negligence or fault is not established, to the party who can avoid the risk in the more efficient way. Some scholars argue for assigning such liability to the manufacturer (Guerra et al., 2022). Their model applies to "robots operated with human intervention and shifts liability to manufacturers provided that operators and third-party victims have invested in due care." On its face, the proposal would probably apply more aptly to autonomous robots that do not involve human intervention, although probably more data is necessary to compare the actual efficiency of such a model. It is equally important to assess how this compares with no-fault insurance in terms of patients' welfare and their ability to obtain quick and easy compensation.

It should be noted here that the position is further complicated by the inherent problems in the Internet of Things (IoT), which can make it quite difficult to determine precisely to whom the wrongful act should be attributed. The emergence of the IoT may pose new questions on how to manage legal responsibility for all risks associated with the Internet of Things. It would thus be better if we, instead of placing full liability on the shoulders of one party, focus on how this technology will affect end users and what risk of loss is appropriate for end users to assume. We need also to determine who can best bear and manage such risks without creating the possibility of enormous potential liabilities that may discourage ongoing innovation and development of AI and IoT technologies. The future of such technologies will depend on how the relevant legal system deals with AI algorithms and their data.

Conclusion

This paper advocates that it has become necessary to re-evaluate the legal status and role of autonomous AI in the healthcare sector as traditional theories of civil liability (strict liability, negligence, product liability, and vicarious liability) are insufficient alone to present convincing answers to all questions raised by such technology. This necessitates that the ultimate legal scheme for civil liability should benefit from a collaboration among the medical industry, the AI industry, the insurance industry, and the legal community.

AI is developing so fast, and the legal environment surrounding it is responding before it can even comprehend the full implications of the technology. Even creators and promoters of AI are calling for legislative intervention to curb and regulate the tide. We believe that the ultimate legal scheme for civil liability for AI will greatly depend on many surrounding factors: regulations of the medical profession, data governance, medical insurance, on codes of ethics among AI creators. It will also depend on how AI itself develops and challenges those rules. It would be more productive, then, to attempt to build a research agenda that probes the surrounding factors affecting liability to come up with a final and ultimate liability rule. We believe such topics are worthy of further, interdisciplinary exploration.

Data availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Received: 19 June 2023; Accepted: 6 February 2024;
Published online: 23 February 2024

Notes

1 Other popular surgical robots are "Sensei X Robotic, Vdrive Robotic, Rosa Spine, Senhance, Mako Rio robot, Smith & Nephew's Navio surgical system, and Mazor Robotics Renaissance Guidance System." For more information, see 'Villanueva (2020) The Legal Battle with the Future of Autonomous Surgical Robotics', 17 Ind. Health L. Rev. 367, 373.
2 *Robbins v Footer*, 553 F.2d 123, 126 (D.C. Cir. 1977).
3 *Register.com, Inc. v. Verio*, Inc 356 F.3d 393 (2d Cir. 2004), in which the court attributed the search robot's actions to Verio and held it liable for breach of contract.
4 Restatement (Third) of Torts: Products Liability, § 19 cmts d and f.
5 § 201(h) of the Food, Drug & Cosmetic Act provides that a 'medical device' is "an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component part, or accessory which is:
(A) recognized in the official National Formulary, or the United States Pharmacopoeia, or any supplement to them,
(B) intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or
(C) intended to affect the structure or any function of the body of man or other animals, and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes. The term "device" does not include software functions excluded pursuant to section 520(o)."
6 The action plan of the US Food and Drug Administration is published online at https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device.
7 See, for example, Solum (1992) 'Legal personhood for artificial intelligences', 70 N C L Rev 1275.
8 For a thorough discussion of this proposal and its different facets, see E Dahiyat (2021) 'Law and Software Agents: Are They Agents by the Way?' 29 Artificial Intelligence and Law. 59, 65.
9 For an interesting discussion, see Determann (2019). No One Owns Data. 70 Hastings Law J. 1; see also A Saidane, S Al-Sharieh (2021). 'A Compliance-Driven Framework for Privacy and Security in Highly Regulated Socio-Technical Environments: An E-Government Case Study' in Information R. Management Association (ed), *Research Anthology on Privatizing and Securing Data* (IGI Global, 2021). For more information about the interaction between the burden of proof and evidentiary discovery rules, see Parisi, F, Pi D & Guerra A (2022) Access to evidence in private international law. Theoretical Inquiries in Law 23(1):77–96. https://doi.org/10.1515/til-2022-0004.
10 The Corporation was originally established by the Accident Insurance Act 1998 and is continued and governed by chapter 7, article 259–286.
11 See for example Lemley M, Casey B (2019) Remedies for Robots. University of Chicago Law Review 86(5):1311–1396.

References

Allain J (2013) From jeopardy! to Jaundice: the medical liability implications of Dr. Watson and other artificial intelligence systems. La Law Rev 73:1049–1079
Bush S (2010) Formation of the physician-patient relationship: the Oregon Court of Appeals clarifies, but questions remain. Phys Org 13:11–13
Chan B (2021) Applying a common enterprise theory of liability to clinical AI systems. Am J Law Med 47:351–385. https://doi.org/10.1017/amj.2022.1
Clark W et al. (2019) Evaluating the use of robotic and virtual reality rehabilitation technologies to improve function in stroke survivors: a narrative review. J Rehabil Assist Technol Eng 6:1–7. https://doi.org/10.1016/j.imu.2021.100567

Cooke BK, Worsham E, Reisfield GM (2017) The elusive standard of care. J Am Acad Psychiatry Law 45(3):358–364
Dahiyat E (2011) Towards new recognition of liability in the digital world: should we be more creative? Int'l J Law Inf Tech 19:224–242. https://doi.org/10.1093/ijlit/ear006
Dahiyat E (2021) Law and software agents: are they agents by the way? Artif Intell Law 29(1):59–86. https://doi.org/10.1007/s10506-020-09265-1
Determann L (2019) No one owns data. Hastings Law J 70(1):1–44
Dyer C (2018) Robot assisted surgery is blamed for heart patient's death. Br Med J 363:k4791. https://doi.org/10.1136/bmj.k4791
Duplechin R (2018) The emerging intersection of products liability, cybersecurity, and autonomous vehicles. Tenn L Rev 85:803–846
Fan Y, Hu Y, Jiang L et al. (2021) Intelligent disinfection robots assist medical institutions in controlling environmental surface disinfection. Intell Med 1:19–23. https://doi.org/10.1016/j.imed.2021.05.004
Frierson RL (2019) Malpractice law and psychiatry: an overview. Focus (Am Psychiatr Publ) 17(4):332–336. https://doi.org/10.1176/appi.focus.20190017
Glavanicová D, Pascucci M (2022) Vicarious liability: a solution to a problem of AI responsibility? Ethics Inf Technol 24(3):28. https://doi.org/10.1007/s10676-022-09657-8
Goyal K, Einstein D, Yao M et al. (2010) Cyberknife stereotactic body radiation therapy for nonresectable tumors of the liver: preliminary results. HPB Surg 2010:309780. https://doi.org/10.1155/2010/309780
Griffin F (2021) Artificial intelligence and liability in health care. Health Matrix 31:65–106
Guerra A, Parisi F, Pi D (2021) Liability for robots I: legal challenges. J Inst Econ 18(3):331–343. https://doi.org/10.1017/S1744137421000825
Guerra A, Parisi F, Pi D (2022) Liability for robots II: an economic analysis. J Inst Econ 18(4):553–568. https://doi.org/10.1017/S1744137421000837
Hand v. Tavera, 864 S.W.2d 678 (Tex. App. 1993)
Habuza T, Navaz AN, Hashim F (2021) AI applications in robotics, diagnostic image analysis and precision medicine: current limitations, future trends, Guidelines on CAD Systems for Medicine. Inform Med Unlocked 24:100596. https://doi.org/10.1016/j.imu.2021.100596
Hellborg S (2019) Liability for medical injuries in Sweden. J du Droit de la Sté et de l'Assurance - Maladie (JDSAM) 23:72–76. https://doi.org/10.3917/jdsam.192.0072
Hellner J (2001) Compensation for personal injuries in Sweden—a reconsidered view. Scand Stud Law 41:249–276
Hodge S (2022) The medical and legal implications of artificial intelligence in health care - an area of unsettled law. Rich J L Tech 28(3):405–468
Kessler DP (2011) Evaluating the medical malpractice system and options for reform. J Econ Perspect 25(2):93–110. https://doi.org/10.1257/jep.25.2.93
Lemley M, Casey B (2019) Remedies for robots. Univ Chic Law Rev 86(5):1311–1396
New Zealand's Accident Compensation Act (2001) Public Act No. 49 (2001)
Parisi F, Pi D, Guerra A (2022) Access to evidence in private international law. Theor Inq Law 23(1):77–96. https://doi.org/10.1515/til-2022-0004
Price WN, Gerke S, Cohen IG (2019) Potential liability for physicians using artificial intelligence. J Am Med Assoc 322(18):1765–1766. https://doi.org/10.1001/jama.2019.15064
Rachum-Twaig O (2020) Whose robot is it anyway? liability for artificial-intelligence-based robots. Univ Ill Law Rev 2020(4):1141–1176
Raskin M (2018) Designer babies, robot malpractice, and the cure for cancer: a legal survey of Some Medical Innovations. N. Y U J L Lib 12:151–188
Restatement (Second) of Torts (1977) American Law Institute Publishers
Restatement (Third) of Torts (2010) Liability for Physical and Emotional Harm, American Law Institute Publishers
Rodgers v. Christie, 795 F. App'x 878 (3d Cir. 2020)
Saidane A, Al-Sharieh S (2021) A compliance-driven framework for privacy and security in highly regulated socio-technical environments: an e-government case study. In: Research anthology on privatizing and securing data. IGI Global. p. 933-962. https://doi.org/10.4018/978-1-7998-8954-0.ch043
Schweikart S (2021) Who will be liable for medical malpractice in the future? how the use of artificial intelligence in medicine will shape medical Tort Law. Minn. J Law Sci Tech 22(2):1–22
Searle J (1980) Minds, brains, and programs. Behav Brain Sci 3(3):417–424. https://doi.org/10.1017/S0140525X00005756
Solum L (1992) Legal personhood for artificial intelligences. NC Law Rev 70:1231–1287
Sonny B (2009) An introduction to medical malpractice in the United States. Clin Orthop Relat Res 467(2):339–347. https://doi.org/10.1007/s11999-008-0636-2
Stein A (2012) Toward a theory of medical malpractice. Iowa Law Rev 97:1201–1257
Talley E (2019) Automatorts: how should accident law adapt to autonomous vehicles? lessons from law and economics. Hoover Institution Working Group on Intellectual Property, Innovation, and Prosperity. working paper series no 19002. https://www.hoover.org/sites/default/files/ip2-19002-paper.pdf. Accessed 9 Nov 2021
Villanueva A (2020) The legal battle with the future of autonomous surgical robotics. Ind Health Law Rev 17:367–392. https://doi.org/10.18060/25051
Wallis K (2013) New Zealand's 2005 'no-fault' compensation reforms and medical professional accountability for harm. N Z Med J 126(1371):33–44
Xue R, Liu R (2022) Statistical analysis of da Vinci procedure volumes of 2021 in the Chinese Mainland. Intell Surg 4:18–22. https://doi.org/10.1016/j.isurg.2022.06.003

Author contributions
All authors contributed equally to this work.

Competing interests
The authors declare no competing interests.

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

Informed consent
This article does not contain any studies with human participants performed by any of the authors.

Additional information
Correspondence and requests for materials should be addressed to Ahmed Eldakak.

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

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