DOI 10.69085/PSC2025352

MEDICAL LIABILITY AND ARTIFICIAL INTELLIGENCE: HOW TORT LAW MAY AFFECT THE USE OF AUTOMATED DEVICES*

Senior Research Fellow Alfio Guido Grasso, PhD
University of Catania, Italy

Abstract: *The field of medicine is among the areas most heavily affected by the development and application of artificial intelligence. However, this advancement raises ethical and legal issues that are difficult to resolve, especially in cases of harm caused by medical devices capable of autonomous decision-making. In assessing the behaviour of a physician who has used a faulty robotic or diagnostic medical device, the level of control over the device and the accuracy with which the patient has been informed are relevant. These difficulties must also be considered in light of recent legislative proposals in the European context, which seek to address the phenomenon of artificial intelligence from an integrated perspective – both by establishing the requirements it must meet in order to minimise the risk of harm, and from an ex post perspective to better address harm produced. The precision and accuracy of these devices must not be overlooked, and their supervised development should also be promoted through appropriate and targeted liability rules, rather than merely repressive ones.*

Key words: *artificial intelligence, medical equipment, robotics, clinical decision support systems, negligence, European law*

МЕДИЦИНСКА ОТГОВОРНОСТ И ИЗКУСТВЕН ИНТЕЛЕКТ: ВЪЗДЕЙСТВИЕТО НА ДЕЛИКТНОТО ПРАВО ВЪРХУ ИЗПОЛЗВАНЕТО НА АВТОМАТИЗИРАНИ УСТРОЙСТВА

ст. н. с. д-р Алфио Гуидо Грасо
Университет на Катания, Италия

Резюме: *Медицината е сред сферите, които са най-силно повлияни от развитието и внедряването на изкуствения интелект (ИИ). Тази технологична еволюция обаче поражда сложни етични и правни предизвикателства, особено в случаи на вреди, причинени от автономни медицински устройства. При*

* *Prodotto PIACERI 2024/2026 Università di Catania progetto SINPLI.*

преценка на отговорността на лекар, който се е доверил на погрешен роботизиран или диагностичен инструмент, от съществено значение са степента на контрол, упражнен върху устройството, както и адекватността на даденото информирано съгласие от страна на пациента.

Тези предизвикателства следва да се разглеждат и в светлината на последните законодателни инициативи в рамките на Европейския съюз, които целят цялостно регулиране на ИИ. Това включва както предварителни мерки – като технически и оперативни изисквания за минимизиране на риска от вреди, така и последващи подходи, улесняващи надлежното определяне на отговорността при настъпила вреда.

Не бива да се подценяват високите нива на прецизност и надеждност, които тези устройства могат да постигнат. Тяхното по-нататъшно развитие следва да бъде насърчавано чрез специално пригодени механизми за отговорност, които стимулират отговорни иновации, вместо да се разчита единствено на репресивни мерки.

Ключови думи: *изкуствен интелект, медицински устройства, роботика, системи за клинична подкрепа на решенията, медицинска небрежност, право на Европейския съюз*

1. Introduction

The now imminent fourth industrial revolution promises to break down the borders between the physical, digital, and biological worlds. Medicine is one of the areas most affected by these changes. Using medical devices and robotics that exploit artificial intelligence requires skills on the part of the medical profession that are often lacking. These are devices potentially able to make autonomous decisions, which pose very disruptive and innovative problems, because they have the potential to take away from the doctor not only material skills but also intellectual ones. At the moment, these devices are not yet widely used by physicians, who are often afraid of them and do not trust or know exactly how to use them.

These technologies have proved to be, to a certain extent, more precise than human beings but they have also shown that when they cause harm, it is usually very significant. Questions related to liability profiles range from the determination of who is responsible in the causal chain of damage to the concrete assessment of the percentage of risk and the allocation of the burden of proof.

In this study, I will argue that by amending the liability rules, physicians might be more inclined to use these devices. In fact, one of the reasons why doctors are reluctant to use the support of artificial intelligence in diagnostics and in clinical practice is the fear of being held liable for any harm they may cause. The hypothesis of this work, also in light of the experience gained from the economic analysis of law, is that by shifting liability rules from phy-

sicians to manufacturers or developers, or by preferring safety rules to liability rules, health professionals will be encouraged to use automated diagnostic systems or clinical radiodiagnostic devices more frequently (to the benefit of everyone's health).

2. Medical device autonomy and liability

The type of problems raised by the introduction of sophisticated automated medical systems depends on what the device is actually able to do 'by itself'. In this regard, if we think about harm caused by teleoperated medical systems, we encounter problems that are not particularly new or different from those raised by harm caused by any other electrical or electromechanical device. Many judicial cases have arisen in the United States, for example in relation to the da Vinci surgical system, which operates as a sort of *longa manus* of the operator who controls it during a given procedure.[1]

In these cases, it was either a malfunction of the da Vinci, which behaved unexpectedly, or much more commonly, an error by the physician in controlling it. Obviously, the use of a complex device like the da Vinci requires specific training on the part of the physician, which is sometimes lacking. In fact, what is most interesting in these U.S. court cases is that in some of them there was pressure from hospitals on physicians to use the da Vinci system even though they were not fully trained.[2]

The use of a complex medical device by an employee of the healthcare facility who is not a specialist could be a cause for the manufacturer's exemption from liability. This is in light of the detailed information the manufacturer provides, suggesting, for example, that the device should be used only by particularly skilled professionals. The healthcare facility may have a duty to prevent the use of such a complex device by those who are not adequately trained to handle it.

If we move into the field of telediagnostic systems, the situation becomes a little more nuanced. Consider, for example, devices that provide heart-condition monitoring for patients suffering from blood-pressure prob-

[1] For more details on these cases, see Ugalde, M. L. The Da Vinci System: A Feat in Medicine, a Newfound Dilemma in the Legal Field. – Richmond Journal of Law and Technology, 17 November 2023; Marks, P. Robo-surgeon da Vinci faces lawsuits. – New Scientist, 217, 2910, 2013, p. 20 (https://doi.org/10.1016/S0262-4079(13)60798-8); Dyer, C. Robot Assisted Surgery is Blamed for Heart Surgery is Blamed for Heart Patient's Death, 2018, BMJ 363.

[2] For further information, see Nicholson Price II, W., S. Gerke, I. G. Cohen, Liability for use of artificial intelligence in medicine. – In: B. Solaiman, I. G. Cohen, eds., Research Handbook on Health, AI and the Law, Edward Elgar Publishing Limited, Cheltenham, 2024, p. 157.

lems, which can automatically send vital data to the physician.[3] Or devices that monitor neurodegenerative disorders, in the form of a portable helmet that can detect brain activity and transmit signals.[4] These devices function as a mobile app that acts as a telemedicine platform, allowing doctors to monitor patients remotely.[5]

In the field of remote diagnosis procedure, there has been great development recently due to the pandemic. On one hand, a shortage of physicians, often engaged on the front line of care, and on the other hand the need to avoid human contact to reduce contagion, have led to an increase in demand for devices capable of assessing the condition of COVID patients without the physical presence of the physician. As a result, telediagnostic procedures have been developed to evaluate patients without direct physician contact. The physician is connected using a newly developed user interface and a lead robot with force-feedback control, which allows precise movements with the follower robot on the patient's side.[6]

They are all extremely useful applications for the early diagnosis of medical problems and are particularly innovative from a technical point of view, but from a legal point of view we need to assess the extent to which they raise issues other than those already raised by medical malpractice.

On the legal level, the most innovative problem might arise when the smart device gives an incorrect value and this erroneous value misleads the doctor who, as a result, adopts the wrong therapy or does not intervene in a timely manner. If we were dealing with a defect in a previous-generation de-

[3] Yenurkar, G., S. Mal, V. Nyangaresi, S. Kamble, L. Damahe, N. Bankar, Revolutionizing Chronic Heart Disease Management: The Role of IoT-Based Ambulatory Blood Pressure Monitoring System. – Diagnostics (Basel), 2024 Jun 19, 14(12), p. 1297 (doi: 10.3390/diagnostics14121297).

[4] For more details, see Rizzetto, M. Wearables e IOT per il monitoraggio costante del paziente a distanza: opportunità e limiti. – In: G. C. Feroni, ed., Le nuove frontere della medicina, Bologna, Zanichelli, 2024, p. 189. For many more examples of AI applications in healthcare, see Chiappini, D. Legal and ethics state-of-the-art of artificial intelligence in medicine. – Diritto e processo, 2020, 97–142 (https://iris.cnr.it/retrieve/fb294aca-93bd-4510-9006-c79562720cb4/LEGAL%20AND%20ETHICAL%20STATE%20OF%20THE%20ART%20OF%20ARTIFICIAL%20INTELLIGENCE%20IN%20MEDICINE.pdf).

[5] Shariq, A. B., N. Mudasser, A. Arshad, K. Abbas, J. Tauseef, N. Sumera, Remote mobile health monitoring frameworks and mobile applications: Taxonomy, open challenges, motivation, and recommendations. – Engineering Applications of Artificial Intelligence, 133, C, 2024, 108233, (https://doi.org/10.1016/j.engappai.2024.108233).

[6] See Fuchtmann et al, COVID-19 and beyond: development of a comprehensive telemedical diagnostic framework. – International Journal of Computer Assisted Radiology and Surgery, 2021, 16, 1403-1412.

vice, we would say without hesitation that the possible malfunction of the machine does not exclude the responsibility of the doctor or the hospital, since they must verify the functionality of the device or ensure that others have checked it. I believe that we can adopt the same solution if a smart device provides incorrect information, although the physician or the hospital could argue that the new tool was more reliable than the old one. However, these do not seem to me to be strong enough defenses to exclude their liability; at most, they will benefit from a reduction in liability compared with the manufacturer, whose liability will be aggravated.

This answer is justified on the grounds that the remote diagnostic procedures described are limited to performing material actions. Although these devices are technically extraordinary, the processing of the resulting data remains entirely in the hands of the physician, and from a legal perspective this should be given the utmost consideration. In essence, they only relieve the human physician of some material tasks, leaving him with purely intellectual tasks. So, fundamentally, these devices cannot be called properly intelligent, because no intellectual effort is required to perform the actions they execute. Thus, there is no need to adopt a different solution than if it were a 'conventional' or older-generation device that was defective.

However, the situation appears different in relation to those devices that are able not only to offer tools to support diagnosis, but also to make the diagnosis themselves (Diagnostic Decision Support Systems).[7] Clearly, these are machine-learning software systems, which can be embedded in a device or used as an App.[8] This is the specific type of artificial intelligence that deals with creating systems that learn – or improve their performance – based on the data they use. Machine-learning systems are assigned a

[7] See Sikma, Edelenbosch, Verhoef, The use of AI in healthcare: A focus on clinical decision support systems, RECIPES project, European Commission, April 2020; Montani, Striani, Artificial Intelligence in Clinical Decision Support: a Focused Literature Survey. – In: Yearb Med Inform., 2019, 28(1), 120; Miller, Medical Diagnostic Decision Support Systems–Past, Present, and Future: A Threaded Bibliography and Commentary. – J Am Med Inform Assoc, 1994, 1, 8–27.

[8] Even if they are not embedded in a device, they are still considered medical devices according to the new regulations. According to Article 2 of EU Regulation 2017/745, which regulates medical devices in the single European market, any software intended for: (1) diagnosis, prevention, monitoring, prediction, prognosis, treatment or mitigation of disease; (2) diagnosis, monitoring, treatment, mitigation or compensation of an injury or disability; (3) study, replacement or modification of anatomy or of a physiological or pathological process or state; (4) providing information through in vitro examination of samples from the human body, including donated blood and tissue, falls within the scope of the Regulation. This is a very important innovation that increases and clarifies the requirements that medical software must fulfil.

goal and given a large amount of data to use as examples of how the goal can be achieved or, alternatively, how decision models can be derived. The larger and more varied the amount of data made available to feed the machine-learning system, the greater the learning capacity of the system, and the more accurate and effective its response to the concrete problem posed to the algorithm.[9]

These are devices potentially able to make autonomous decisions, which pose much more serious and innovative problems. They are devices that may, in the future, be presumed to replace the human physician. In general terms, they raise at least three main problems:

1. The loss of the humanistic dimension of medical knowledge;
2. Excessive reliance on the diagnosis proposed by the device, leading to a decline in the physician's competence and expertise;
3. The loss of the habit (deskilling) of relating to the patient, autonomously perceiving their problems, and carrying out an anamnesis without the suggestion of the machine-learning device.[10]

In addition, they also pose specific questions in terms of liability in the event that harm is caused. Decisions taken statistically, even by a perfectly functioning algorithm, may not be correct, especially in relation to the individual case since statistics provide more reliable information about a global rather than a local effect.[11]

---

[9] Probably the best-known clinical decision support system is IBM Watson. The system is able to perform complex anamnestic and diagnostic tasks, defining personalised treatments, thanks to the use of highly developed cloud computing. A recent study compared its high accuracy rate with the treatments received in urology departments: Yu et al., Early experience with Watson for Oncology: a clinical decision-support system for prostate cancer treatment recommendations. – World J Urol., 2021, Feb, 39(2), 407-413. See Lagioia, Contissa, The strange case of dr. Watson: liability implications of ai evidence-based decision support systems in health care. – EJLS, 12, 2, November 2020, 245-289; Allain, From Jeopardy! to Jaundice: The Medical Liability Implications of Dr. Watson and Other Artificial Intelligence Systems. – Louisiana Law Review, 2013, 73, 4, p. 1062, for the problems that Watson poses in terms of liability.

[10] See Sutton et. al, An overview of clinical decision support systems: benefits, risks, and strategies for success. – Npj Digital Medicine, 2020, 3, 17, 1-10; Wasylewicz, Scheepers, Hoeks, Clinical Decision Support Systems, in Fundamentals of Clinical Data Science. – In: Kubben et al., eds., Springer, 2019, p. 153; Lysaght et al, AI-Assisted Decision-making in Healthcare. The Application of an Ethics Framework for Big Data in Health and Research. – Asian Bioethics Review, 2019, 11, 299–314.

[11] See Cabitza, Alderighi, Rasoini, Gensini, Potenziali conseguenze inattese dell'uso di sistemi di intelligenza artificiale oracolari in medicina. – In: Recenti Prog. Med., 2017, 108, p. 397; Sikma, Edelenbosch, Verhoef, Op. cit., p. 19.

Statistics, by definition, can never lead to certain conclusions, but only to probable ones.[12]

Current machine-learning approaches to diagnosis are purely associative, which involves learning correlations between patient data and disease occurrences, resulting in the identification of diseases that are strongly correlated with patient symptoms. But this inferential approach, which is sufficient for simple causal scenarios involving single diseases, can lead to sub-optimal or dangerous diagnoses when applied to differential diagnosis, where a clinician chooses between several competing disease hypotheses.[13] Diagnosis, in fact, is not so much an association between symptoms and disease, but is fundamentally a task of counterfactual inference, whereas algorithms based on associative learning merely determine the most likely diseases in the population to which the patient belongs.[14]

In legal terms, this way of proceeding with diagnostic algorithms poses several kinds of problems. Among the many, which also affect other areas of law such as health law, data protection law and criminal law, I have selected the five problems that are most relevant to tort law.

1. In relation to patients who are asymptomatic or have atypical symptoms. Let us consider a patient who, in the absence of symptoms, is not diagnosed by the algorithm, or who presents anomalous symptoms that lead the algorithm not to attribute them to the correct disease.
2. With reference to rare diseases in general or rare in relation to the population to which the patient belongs. Let us consider a disease that affects 90% of adults and 10% of young people. The risk is that these percentages will give rise to results in the decision-making phase that are calibrated only to a certain target group of people[15] – those most affected by

[12] See Mittelstadt, Allo, Taddeo, Wachter, Floridi, The ethics of algorithms: Mapping the debate. – Big Data & Society, December 2016; yet Gini, I pericoli della statistica, 1939, reedited in 2001 in Statistica e induzione/Induction and Statistics, supplement to Statistica, Rome/Bologna, CLUEB, 61 (1), 27-70.

[13] See Alther, Reddy, Clinical Decision Support Systems, in Healthcare Data Analytics. – In: Reddy, Aggarwal, eds., CRS Press, 2015, p. 646; Jiang, Qiu, Xu, Li, The Research of Clinical Decision Support System Based on Three-Layer Knowledge Base Model. – Journal of Healthcare Engineering, 2017, p. 1.

[14] See Richens, Lee, Johri, Improving the accuracy of medical diagnosis with causal machine learning. – Nat. Commun., 2020, 11, p. 3923; Prosperi, Guo, Sperrin et al., Causal inference and counterfactual prediction in machine learning for actionable healthcare. – Nat. Mach. Intell., 2020, 2, 369–375; Papadopoulos et al., A systematic review of technologies and standards used in the development of rule-based clinical decision support systems. – Health and Technology, 2022, 12, 713–727.

[15] See Sikma, Edelenbosch, Verhoef, Op. cit., p. 17.

the disease – thus causing damage to the minority of people who, statistically speaking, are less likely to suffer from the disorder.[16]

3. With respect to patients with comorbidities or a complex clinical picture. Let us consider a patient whose algorithm correctly diagnoses one disease but in fact hides another, more serious one.[17]
4. With regard to the characteristics of the target environment, which may be very different from those of the laboratories designing or testing the device. On the one hand, the incidence rates of a disease may vary from one area to another, due to the history of that region, climatic or natural factors, or industrial facilities. On the other hand, the surrounding health environment affects the performance of the device.[18]
5. A problem arises with regard to diseases as yet unknown and therefore not included in the knowledge base of the device.

3. Automation bias

The assessment of the situation is made worse by the fact that the physician has no way of verifying how the system achieved certain results, let alone which data fed it and which programmers designed it. The opacity that characterizes these systems entails the risk that the health professional, in trying to make his or her own decision, cannot reasonably validate – or, on the contrary, reasonably reject – the system's proposal. This raises a dilemma for the physician when confronted with the machine (deciding whether or not to rely on the indications of the algorithm).

But the physician, in the face of uncertainty, will tend to follow the advice of the algorithm, because he is guided by the so-called automation bias, which consists in the tendency to rely on automated decision-support systems, underestimating and overshadowing the decision-making capabilities of human beings. The importance of this bias is also underlined by its mention in Article 14(4)(b) of the AI Act (Regulation 2024/1689). When the producer sets out the measures to ensure that the device always works under human control, they must 'remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias'), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons'.

[16] See Winkler et al., Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. – JAMA Dermatol., 155, 2019, 1135–1141.

[17] Cfr. Comito, Falcone, Forestiero, AI-Driven Clinical Decision Support: Enhancing Disease Diagnosis Exploiting Patients Similarity. – IEEE Access, 10, 2022, 6878-6888, p. 6879.

[18] Panch, Mattie, Celi, The "inconvenient truth" about AI in healthcare. – Npj Digit. Med., 2, 2019, p. 77.

There is also another reason why the physician, in conditions of uncertainty, will be inclined to follow the indication of the algorithm, and that is a legal reason. If the physician decides not to follow the machine's diagnostic proposal, they should be prepared to give specific explanations as to why the machine – usually so efficient – should be considered unreliable in the specific case. Since this might be a very difficult task, due to the fact that algorithmic decision-making systems can be extremely complex and opaque, the human professional involved would be more likely to confirm the decision made by the machine.

4. The efficiency of artificial intelligence in medicine and the economic analysis of law

However, on the other hand, we must bear in mind that in the vast majority of cases the algorithm does not miss the therapeutic indication. So when we talk about errors stemming from incorrect algorithmic indications, we must give the problem its proper dimension. Otherwise, we risk pushing a whole series of highly useful instruments out of the market.[19]

In light of this picture, we can examine the consequences that may arise in terms of liability if the recommendation provided by the algorithm turns out to be wrong. Starting from this premise, we must avoid placing all the responsibility on the physician. This for two reasons: first, because the doctor is usually not the most financially solvent party, and second, because we should not discourage medical practitioners from using these tools, which in the vast majority of cases prove to be an efficient, precise support, and therefore of enormous benefit to public health.

On the other hand, we should avoid overly simplistic solutions, such as always excluding the doctor's liability. It is argued that the internal decision-making processes of intelligent medical products are often not recognizable to the physician. As a result, the risks posed by the product usually become apparent only after a harmful event. Given this lack of predictability, the physician could not reasonably be held fully liable.[20]

[19] See Montani, Striani, Artificial Intelligence in Clinical Decision Support: a Focused Literature Survey. – In: Yearb Med Inform., 2019, 28(1), p. 120; Kunhimangalam, Ovallath, Joseph, A clinical decision support system with an integrated EMR for diagnosis of peripheral neuropathy. – Neuroimaging Clin N Am. 2014 Feb, 24(1), 49-65, who developed a particularly successful DDSS application for the diagnosis of peripheral neuropathy, achieving 93% accuracy over experts in identifying motor, sensory, mixed or normal neuropathies; Sutton et al, Op. cit.

[20] See Brutti, Intelligenza artificiale e responsabilità in ambito medico: la prospettiva statunitense. – Resp. med., 2018, 4, 473-474.

While it is certainly true that product control affects the extent of the doctor's liability, the unfathomable 'black box' nature of AI should not be used as a pretext to exclude the doctor's liability a priori. The recommendations on the use of the AI device issued by the producer should be taken into account. If these recommendations are not transparent, the doctor's erroneous decision must be viewed with goodwill. However, if the manufacturer's guidance is clear – requiring a certain standard of preparation in the physician, or indicating, for example, a certain percentage of diagnostic error – then liability on the part of the physician and/or the hospital should still apply.

The introduction of AI in medicine must lead to a change in the criteria or guidelines used so far to determine the level of control exercised by the physician and the ways in which accurate patient information influences his or her assessment (we will see this in the following paragraphs). AI is not a passive technology, completely in the hands of the medical practitioner, rather, it gives rise to (semi-)autonomous systems which the physician does not always control or master, and which are instead in the hands – also in terms of greater efficiency – of the manufacturer and developer.

The economic analysis of law, through its many examples – from environmental law to expropriation – shows that civil liability is the private law's chosen instrument for affecting societal change. Recognizing or not recognizing compensation for damages, or even imposing so-called punitive damages – liquidated sums far exceeding the victim's actual loss – fundamentally influences the way practitioners and entrepreneurs operate within a sector.[21]

5. Product defect and manufacturer's liability

A violation of the medical standard does not necessarily lead to product liability. In a medical liability case, the manufacturer is not liable for an AI's treatment error, but only for the possible production of a defective device or for providing insufficient or erroneous information to the user.[22]

Although Directive No. 2024/2853 on liability for defective products ('PLD' – Product Liability Directive) addresses developments related to new technologies, particularly artificial intelligence, it is still grounded in the notion of

[21] Calabresi, G., P. Bobbitt, Tragic Choices. New York, Norton & Co Inc, 1978; Calabresi, G., The Cost of Accidents: A Legal and Economic Analysis. Yale University Press, 1970. For a complete overview on the subject, see Kaplow, L., S. Shavell, Economic analysis of law, Chapter 25. – In: Handbook of Public Economics, Elsevier, 2002, 3, 1661-1784.

[22] The EU has a new product liability regime for defective products: EU Directive 2024/2853. The previous regulation was COUNCIL DIRECTIVE of 25 July 1985 on the approximation of the laws, regulations, and administrative provisions of the Member States concerning liability for defective products.

defect. A product may function perfectly and still cause harm. In such cases, the manufacturer would not be liable, and the physician's liability would only arise if their conduct were grossly negligent.

We can probably state, therefore, that in the event of a misdiagnosis by a perfectly functioning algorithm and a doctor's conduct that is not grossly negligent, there is a gap in protection. In this respect, the general principle may be considered that products which do not meet mandatory safety standards are regarded as defective, irrespective of the producer's fault or the existence of an actual defect. On this point, the Report from the Commission to the European Parliament, the Council, and the European Economic and Social Committee on the safety and liability implications of AI, the Internet of Things, and robotics, echoing the position of the expert group, notes that adaptations of national laws to facilitate the burden of proof for victims of AI-related damage could be considered. For example, the burden of proof could be linked to compliance (by a relevant operator) with specific cybersecurity or other safety obligations set by law: if one does not comply with these rules, a change to the burden of proof regarding fault and causation could apply.[23]

6. Artificial intelligence and the physician's gross negligence and informed consent

In light of this premise, we can affirm that when the algorithm rarely makes mistakes, it does so in cases of great complexity. We can therefore consider such instances as special cases that exceed the average professional background, even that of a physician particularly experienced in the relevant medical area.

If we take a closer look at national laws, we see, for example, that in Italy there is a rule, Article 2236 of the Civil Code, which could be useful in this regard because, in cases involving the solution of particularly difficult technical problems, it limits the professional's liability to cases of malice or gross misconduct.

However, following an algorithmic decision in itself does not seem to constitute gross misconduct; on the contrary, following algorithmic instructions could even be considered good medical practice.

What constitutes gross negligence, for which the physician would be held liable, is the failure to pay attention to signs that should lead him to question the algorithmic indication – signals that only he could detect in front of

[23] Brussels, 19 February 2020, COM(2020) 64 final. Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics.

the patient, provided these signals were observable with the diligence appropriate to his profession. A different approach to reducing the physician's liability in the event of an algorithmic error, and thus encouraging the use of these devices within the medical community, could be to inform the patient about the use of such systems and their associated risk rates. This information would later be taken into account in the liability assessment in the event of harm.

However, doubts arise in this regard, because Italian law on informed consent (Law 22 December 2017, no. 219) does not provide any guidance on this point (Art. 1), and Italian case law considers that patients should be informed only about common, recurring outcomes, not about abnormal or rare ones, so as not to discourage them from undergoing useful and generally low-risk procedures or treatments. If this view is followed, there would be no obligation to inform the patient about the use of these artificial intelligence devices and their error rates.[24]

However, this approach can also be seen as paternalistic, whereas the law affirms the right of every person to be fully, accurately, and comprehensibly informed about their diagnosis.

Other legal systems, such as Germany's, interpret consent in a more detailed and egalitarian manner, while the Italian approach is more paternalistic. The debate on whether a physician must inform the patient that AI systems are being used for diagnosis and treatment therefore remains open.[25]

If we examine § 630 BGB more closely, particularly letters e and h, we can note the importance of the information that the doctor provides to the patient in terms of responsibility and the right to self-determination. The treating party is obliged to inform the patient of all circumstances relevant to consent. This includes, in particular, the nature, extent, implementation, anticipated consequences, and risks involved in the procedure, as well as its necessity, urgency, suitability, and prospects for success regarding the diagnosis or therapy.

We can therefore assume that the doctor must first inform the patient that a machine learning device is being used. It can also be assumed that the

[24] See De Menech, C. Intelligenza artificiale e autodeterminazione in materia sanitaria, in BioLaw Journal, 2022, p. 181; Tozzi, Cinelli, Informed consent and artificial intelligence. – Biolaw Journal, Special Issue, 2, 2021, p. 106; Muto, Tosco, Forum AI and Law. – Biolaw Journal, 2020, 1, p. 507.

[25] See Droste, Wiebke, Intelligente Medizinprodukte: Verantwortlichkeit des Herstellers und ärztliche Sorgfaltspflichten, MPR 2018, 109–115; Helle, Intelligente Medizinprodukte: Ist der geltende Rechtsrahmen noch aktuell? – Medizinrecht, 2020, 38(12), 993-1000; Spindler, Telemedizin; Roboter und Künstliche Intelligenz, sub § 823 BGB, VIII. Besondere Fälle. – In: Gsell, Krüger, Lorenz, Reymann, eds., Beck-online. Grosskommentar zum BGB, Munich, 2021, Rn. 1052-1057.

doctor must inform the patient of the associated risk rates of misdiagnosis. If the physician provides detailed information and the patient gives consent, this should influence the assessment of the physician's liability.[26]

In this case as well, however, the physician remains liable if serious misconduct has occurred, particularly if the committed error constituted a general treatment risk that was fully manageable by the practitioner and resulted in harm to the patient's life or health (see § 630h BGB, nos. 1 and 5).

Hindering this approach are the limits of transparency associated with the operation of artificial intelligence devices, which may make it difficult in practice to communicate adequate information to patients. However, as will be discussed below, the information requirements imposed on providers by the Proposal for a European Regulation on Artificial Intelligence (in particular Articles 12–14) could help compensate for the difficulty of adequately informing patients about the functioning and limitations of machine learning devices.

In both Italian and German law, there is a parallel in the treatment of gross negligence regarding the reduction of the physician's liability.[27]

An alternative approach, also followed in the German legal system (see § 630h BGB, nos. 1, 4, 5),[28] consists in the use of relative presumptions. These presumptions can provide effective protection for the injured patient from the combination of human negligence and the fallibility of AI medical devices, without unduly penalising the physician.

---

[26] See Ebers, Heinze, Krügel, Steinrötter, Künstliche Intelligenz und Robotik, 1. Auflage 2020, München, 100, who seem to exclude such an obligation only when the AI is mature, sufficiently tested, and actually safe to use.

[27] Also significant in this respect is the French reform of the law on bioethics. Law No. 2021-1017 of 2 August 2021 inserted a new article (Art. L4001-3) into the Public Health Code, which lays down the obligation for physicians who use medical devices that involve algorithmic data processing for the purpose of carrying out preventive, diagnostic or therapeutic treatment to inform the person thereof and, where appropriate, to warn him or her of the interpretation of the results obtained. Healthcare professionals shall, in turn, be instructed in the use of data processing, and the patient data used in the processing, as well as the results obtained, must be accessible to them. Developers and designers ensure that the relevant operation can be explained to users.

[28] According to § 630h BGB, nos. 1, 4, 5, if a physician was not qualified to perform the treatment he or she performed, it must be presumed that the lack of qualification was the cause of the injury. This is a presumption that the physician can overcome if one considers, precisely, that, as we have argued, when the algorithm is wrong it means that it is a rare and difficult case. Moreover, it is difficult to claim that the reason for the error was solely the physician's lack of preparation. Even if, as we have seen before in connection with American court cases, it happens that the hospital imposes the use of a certain means despite the fact that the clinician is not sufficiently prepared.

7. The emerging European system on AI liability

Resorting to presumptions turns out to be particularly useful if one considers the characteristics of autonomy and opacity possessed by artificial intelligence systems. The European legislator, not by chance, has chosen presumption as the preferred instrument of harmonisation in the recent Proposal for a Directive on Artificial Intelligence Liability.[29]

This regulatory framework is based on the application of the requirements for high-risk systems contained in the EU Artificial Intelligence Act.[30] Depending prevalently on the compliance or non-compliance with such requirements, the burden of proof may be eased. Let us begin, therefore, by taking a closer look at the contents of the so called AI Act, which is placed not only temporally but also logically before the Proposal for a Directive on Artificial Intelligence Liability, because the former aims at the prevention of possible damage from artificial intelligence, while the latter deals with damages that have already occurred and for which compensation must be given.

7.1. Artificial Intelligence Act (Regulation (EU) 2024/1689

The Artificial Intelligence Act takes a preventive stance on damage, seeking to prevent it. The European Commission aims to adopt a balanced approach to AI regulation that ensures effective protection of fundamental rights without hampering the socio-economic benefits of AI. With a view to ensuring a functional balance between the interests to be preserved and the techno-economic development of the single market, the European Commission has identified four levels of risk, each referring to certain AI systems and their applications.

The Commission considers it an 'unacceptable risk' – and therefore prohibited – when AI is applied in systems that are considered a threat to security. This category includes applications that manipulate human behaviour to circumvent the free will of users (e.g. voice-assisted toys that encourage minors to behave dangerously).

The Commission instead includes within 'high risk' thresholds those systems and technologies of AI entrusted with the preservation or implementation of certain fundamental rights, as in the case of critical infrastructures or

[29] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), COM/2022/496 final, Bruxelles, 28.9.2022.

[30] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

scholastic and professional training; for example, the management of private and public credit services in which an AI algorithm can deny access to a loan, or systems assuring public security.

Thirdly, a 'limited risk' is is assigned to systems such as chatbots that may create an erroneous awareness in individuals, i.e. uncertainty as to whether the interaction is with a machine rather than a human being.

Finally, a 'minimal risk' is associated with AI applications that are not as invasive as those described above. For example, video games or spam filters applied to e-mail services are included in this category.

A higher level of risk is associated with different and more restrictive compliance rules prior to placing the device or application on the European market. Among high-risk AI systems are also medical devices. They are permitted, of course, subject to the EU AI proposal's requirements, including a conformity assessment. In this regard, I would like to highlight a few concerns. The Proposal places most medical uses of AI in its highest risk class; however, medical devices should not always be considered high risk.[31] This could have a negative effect on increasing the use of these devices, because they are subject to greater regulatory burdens, which would therefore also increase their price and push them out of the market. The definition of high

[31] Article 6 of the AI Act determines that an AI system is "high-risk" when it is intended to be used as a safety component for a product, or is itself a product covered by the Union harmonization legislation in Annex II and is required to undergo a third-party conformity assessment pursuant to that legislation. Annex II includes the EU Regulation 2017/745 on Medical Devices and In Vitro Diagnostic Medical Devices (see note 4 above). Accordingly, most software as medical device will undergo a third-party assessment and fall under the definition of a "high-risk AI system." On the basis of these indications, hospital information systems that support the patient management process and serve the purpose of admission are qualified as medical devices. Similarly, radiation treatment planning systems, whose purpose is to calculate the dosage of ionising radiation to be applied to a specific patient, are considered medical devices, as they directly control, monitor, or influence the source of ionising radiation. Medication planning systems, whose purpose is to calculate the dosage of drugs to be administered to a specific patient, also fall under the category of medical devices. Prehospital ECG systems – software-based systems intended for ambulance services to store and transfer information from patients connected to an ECG monitor to a physician at a remote location – are likewise included in this definition. From this articulated list of examples, it can be seen that decision support systems all fall within the highest risk level and, as such, are subject to stringent safeguards. These include post-certification monitoring, annual surveillance audits of manufacturers by notified bodies, and procedures for verifying the conformity of software with the manufacturer's quality management system, with the level of control increasing according to the classification of the devices.

risk should probably apply only to those medical devices capable of producing significant harm despite the existence of substantial human control.

In addition, Article 14 (in particular paragraph 4.d) states that the user of high-risk devices – thus including medical devices – must be able to decide, in any particular situation, to disregard, override or reverse the output of the high-risk AI system. This sacrosanct principal risks remaining a dead letter if it is not accompanied by a correlative exclusion or reduction of the physician's responsibility. In a world dominated by defensive medicine, it is unrealistic to think that the physician will take responsibility for going against the output of the machine.

7.2. Proposal for a Directive on Artificial Intelligence Liability

Turning instead to the Proposal for a Directive on Artificial Intelligence Liability, which addresses these issues from an ex post perspective – i.e., the appropriate attribution of harm generated by a medical device or by a physician's conduct – and not from a pre-emptive harm perspective, as in the before analysed Regulation on AI.

The Proposal for a Directive on Artificial Intelligence Liability, which has a scope that is as narrow as it is crucial, establishes common rules on the disclosure of evidence concerning high-risk artificial intelligence systems suspected of having caused harm, and on the burden of proof in fault-based claims.

First of all, a presumption of breach of a duty of care is introduced when the provider or user of a high-risk artificial intelligence system refuses to comply with a court order to disclose or preserve evidence at their disposal (Article 3(5) of the Draft Directive). This presumption appears welcome in light of the reluctance often shown in practice by providers and users to disclose information and evidence concerning the artificial intelligence systems for which they are responsible.

The provider must inform the user of the residual risks associated with each danger that the device may generate, as well as the overall residual risk it may present (Art. 9(4) of the Proposal for a Regulation on Artificial Intelligence). The instructions for use of the device must also be brought to the user's attention, along with a range of information concerning the provider, the performance and purpose of the system, the groups of persons on whom it is intended to operate, the specifications of the input data or of the training, validation and test datasets, the expected lifetime of the system, and all maintenance and care measures required to ensure its proper functioning, including updates (Article 12 of the Proposal for a Regulation on Artificial Intelligence). All this information is aimed at making the principle of human oversight (human-in-the-loop) effective (Article 14 of the Proposal for a Regulation on Artificial Intelligence).

Consider, on the other hand, the information deriving from the obligations that the AI Act places directly on the user and which the physician, in his capacity as a user of artificial intelligence devices, might therefore be required to disclose. This is evidence deriving from activities carried out personally by the user or related to the supervisory activity performed by the user on the device. Users must first and foremost ensure the relevance of the data entered into the device (input data) to the intended purpose of the high-risk artificial intelligence system (e.g. a patient's data entered into a clinical or diagnostic decision support system). They must also monitor the operation of the device in accordance with the instructions received and are required to retain the logs automatically generated by the system for a period of time proportionate to the relevant purpose and legal obligations Article 26 of the AI Act.

Major concerns are raised by the presumption regarding the existence of a causal link.

The presumption applies, apart from other minor cases, to physical harm caused by artificial intelligence systems, provided that the injured party proves: the defendant's negligent disregard of the duties of care laid down by European or national law to prevent the harm from occurring; the reasonable likelihood, inferred from the concrete circumstances, that such conduct affected the performance of the system; and the origin of the harm from the output of the device ((Article 4(1)(a-c)). However, when it comes to a provider or user of a high-risk artificial intelligence system, the object of proof narrows and, as far as failure to exercise due diligence is concerned, boils down to proof that (1) the provider has failed to comply with one of the requirements to which it is subject under the AI Act (Chapters II and III), or that (2) the user (a) has failed to comply with the obligation to use or monitor the system in accordance with the instructions received or, where appropriate, to suspend or discontinue its use, or (b) has exposed the system to irrelevant input data under his control.

While this latter requirement is properly under the user's control and, as such, is enforceable against the user (e.g. a physician who enters indications in the programme's dialogue box that can be traced back to another pathology, thereby misleading the system), the same user often finds himself in a position of not being able to comply with the obligation to use or monitor the device in accordance with the instructions received without thereby causing damage for which he could be held liable under national liability provisions. The instructions received may, in fact, be ambiguous, partial, or antinomian to the case brought to his attention; nevertheless, the physician should comply with them in order to avoid the risk of triggering the pre-

sumption of the existence of a causal link between his conduct and the output produced by the system.

Despite the fact that Article 14 of the AI Act (in particular paragraph 4d) stipulates that the user of high-risk devices – such as medical devices – must be able, in any particular situation, to exclude, ignore, or reverse the output of the system, in reality, as presented, the physician is not always able to orientate themselves by following their own science alone. Neither does it seem plausible to suppose that, faced with an uncertain or ambiguous algorithmic indication, the physician bears the burden of derogating from the algorithm; moreover, this is particularly unrealistic in a sector characterised by marked defensive mechanisms such as healthcare. If the use of artificial intelligence is to be encouraged without running the risk of deactivating the 'human-in-the-loop' principle, it is necessary to avoid basing interventions to tighten the burden of proof on bare causality, limiting them to cases in which a (serious) reproach can be addressed to the physician as malpractice.

Unfortunately, it was recently announced that the European Commission has decided to withdraw the Directive on the Responsibility of Artificial Intelligence after the Paris Summit on Artificial Intelligence in February 2025, for reasons that are mainly political rather than legal (the majority needed to approve it would probably not be found).[32]

Nevertheless, the AILD is still an important model for study and comparison, and it is hoped that it will be reconsidered, since it represents the *longa manus* of AI law (the question of responsibility certainly cannot be ignored).

8. Conclusions

The emergence of artificial intelligence systems involved in diagnosis raises the issue of resolving conflicts between humans and machines, in the sense that a specific event can be objectively attributed to the doctor even in the absence of a fault factor. In this context, in order not to discourage the use of artificial intelligence in the medical field – which in the vast majority of cases has demonstrated the ability to produce extremely successful results – there is a need to curb the punitive excesses that would arise from placing the blame solely on the basis of causality.

In particular, the drafted proposal is that, in order to promote the use of artificial intelligence in medicine, the physician should only be held liable when there is gross negligence in the use of the device or when the damage was caused by a professional error. In other words, outside the case where the practitioner committed an error pertaining to his or her knowledge (such

[32] Andrews, European Commission withdraws AI Liability Directive from consideration (https://iapp.org/news/a/european-commission-withdraws-ai-liability-directive-from-consideration).

as incorrectly detecting vital parameters and causing the machine to malfunction), the physician would only be liable when he or she failed to intervene to correct the device or deactivate it, even though he or she was able to do so. In other cases, when the health provider was only slightly at fault (for example, he or she did not realize that the device was malfunctioning but it would have been very difficult to notice), the manufacturer or developer will be objectively liable. They will then be liable not only when the product is defective, but also when it has failed to provide the safety and self-correction that would have been expected of them.

This change in the liability rules, to which national legislators (the subject of civil liability is not a European competence, except for some specific areas) and legal practitioners are called upon, is ahead of the developments that one can envisage artificial intelligence in medicine will have. Just think of radiomics, that emerging field of research focusing on the extraction and analysis of quantitative features from medical images. From radiomics, and thus from the extraction and analysis of quantitative characteristics from medical images, it is possible to create clinical decision support models capable of drastically reducing the work of radiologists, as they do not require manual image selection or pre-processing.[33] Nevertheless, these are activities that are only partly taken away from doctors, since radiomics allows information to be obtained that the human eye could not perceive and that the human mind itself could not process because of its quantity and complexity at the time of reporting, but which must be interpreted and translated into an operational response. So human error can always lurk, however much the possibility of it happening is reduced with the help of technology.

We should not forget that the challenge is also to strengthen the system of technological risk prevention in healthcare in a proportionate and bal-

[33] Medical images (MRI, CT, X-ray) and/or their features are analyzed with machine learning or deep learning programs to develop predictive models of pathology or treatment outcomes. This represents a quantitative approach to medical imaging, aiming to enhance the information available to doctors through advanced mathematical analysis. Thanks to this technology, medical images obtained from examinations such as CT, MRI, or PET scans are converted into numerical data. These data are then processed using specialized computational tools and managed and analyzed with advanced techniques, including artificial intelligence systems and 'big data' approaches. See Sahiner, B., N. Petrick, Evaluation of CAD and Radiomic Tools– In: The Handbook of Medical Image Perception and Techniques. Cambridge, Cambridge University Press, 2018, 389-406; Retico, A. Applicazioni di radiomica e intelligenza artificiale all'imaging medico: potenzialità e criticità, in 1060 Congresso nazionale società italiana di fisica, 14–18 September 2020 (https://static.sif.it/SIF/resources/public/files/congr20/ri/Retico.pdf).

anced way, by identifying the most suitable protective measures to deal with the risks that the use of machine learning algorithms actually poses. The focus should be, first and foremost, on how to prevent errors in healthcare delivery, but according to a less strict risk management model that leaves more room for 'accountability' on the part of producers. A model whose functionality should, however, be verified when the error is made by an algorithm with self-learning capacities.