Berrikusitako literatura oinarri hartuta, dokumentu honek Europan etorkizunean IA medikoa arautzeko proposatu diren orientabide nagusien laburpena eskaini nahi du, bai eta izan ditzakeen eragin kliniko, industrial eta sozialena ere.

Giltza-Hitzak: Adimen Artifiziala Medikuntzan. Osasun Datuak. Big Data. Erabaki Automatizatuak Hartzea. Etika. Zuzenbidea. Adimen Artifizialari Buruzko Europako Erregelamendua. Erantzukizun Zibila.

Sobre la base de la literatura revisada, este documento tiene la intención de ofrecer una síntesis de las principales orientaciones que se han propuesto para la futura regulación de la IA médica en Europa, así como sus posibles impactos clínicos, industriales y sociales.

Palabras Clave: Inteligencia Artificial en Medicina. Datos de Salud. Big Data. Toma de Decisiones Automatizada. Ética. Derecho. Reglamento Europeo de Inteligencia Artificial. Responsabilidad Civil.

S'appuyant sur la littérature revue, cet article entend proposer une synthèse des principales orientations qui ont été proposées pour la future régulation de l'IA Médicale en Europe ainsi que ses impacts potentiels cliniques, industriels et sociétaux.

Mots-clés : Intelligence Artificielle en Médecine. Données de Santé. Big Data. Prise de Décision Automatisée. Éthique. Droit. Réglementation Européenne de l'Intelligence Artificielle. Responsabilité.

Five crucial challenges for regulation of Medical Artificial Intelligence

Alkorta, Itziar
University of Basque Country. Faculty of Law. Manuel de Lardizabal, 2 – E-20018 - Donostia
Jakiunde, Zientzia, Arte eta Letren Akademia. Prim, 7. E-20007 - Donostia
itziar.alkorta@ehu.eus

Recep.: 2022-06-30
Acep.: 2022-11-21

Five crucial challenges for regulation of Medical artificial intelligence

Introduction

Medical technologies are actually going through an unprecedented digitization process. Many of these software systems are progressively incorporating algorithms as well as machine-learning technologies, whereby intelligent software is being used for a variety of purposes, ranging from clinical research and drug development to patient monitoring and decision support systems. In addition to a better health care, these technologies are deemed to respond to the increasing needs of an aging population facing chronical diseases in a scenario of scarce resources.

European Union authorities have issued some pragmatic approach to medical AI to foster innovation in health care under privacy and security safeguard requirements. Other than the General Data Protection Regulation and the Data Governance Act, new regulations are being drafted in Europe to frame the "Medical Artificial Intelligence", such as the Artificial Intelligence Act, which is expected to guide the uses and the limits of these new systems.

In this paper, we are addressing the ongoing reflection on the main requirements "medical AI" systems should meet to comply with the rights of the users. Privacy, safety, explainability and fairness principles will be explained. Also terms such as liability and accountability deriving from AI systems. We believe that not until these issues are fully discussed and assessed by the general public, should these technologies be offered on a regular basis.

In order to do so, we will first tackle the role of artificial intelligence in medicine from a chronological perspective, to be then able to discuss medical AI adoption patterns, including the benefits and the risks of the decisions that have been made so far on medical AI technology adoption. Once the context is laid, at the following sections we shall address the five main challenges medical AI systems are facing from a regulatory point of view, i.e., automated decision making patterns; identification of high risk and non-high risk systems and their authorization processes; liability systems that are suited to automated decision making, and, finally, regulation of non-discriminatory algorithm based decision systems.

Role of artifial intelligence in medicine

One of the reasons for the permeability of AI systems in the clinical sector is that current medicine is strongly based on **digitized data**.[1] In Europe as well as in the US, complete digitization of health records has been a major endeavor that has been going on since the mid-1990s.[2] From this moment on until the end of the first decade of the XXI century, the principal source of clinical data for research and treatment have been clinical records. However, in the last decade, other sources of data have appeared in the horizon, as for

[1] Cummins, N. & Schuller, B. W. (2020) "Five Crucial Challenges in Digital Health", *Front. Digit. Health* 2:536203.

[2] Evans R. (2016) "Electronic health records: then, now, and in the future". *Yearb Med, Inform*. 25:S48–61.

instance, a variety of **molecular** (e.g. genomics, transcriptomics, metabolomics…) analysis, including consumer driven genetic profiles[3]; different medical **imaging** modalities, very rich in phenotype information; data from wearable or implantable **sensors**[4], which have enabled mass data collection to aid public health and research efforts for the COVID-19.

Along with data, **algorithms**[5] have played a major role in the development of AI for health solutions. Medicine has been using algorithms for diagnosis, epidemiology and clinical essays for a long time[6] but with the advent of the computer sciences, algorithms have thrived in medicine, with an increased ability in task such as automating tasks, accelerating workflows, predicting disease or its progression, predicting treatment response, stratification of risk or classification of complex phenotypes, enabling precision medicine, or personalizing prescribed or self-managed interventions.[7] Epidemiology, for instance, relies on systems of algorithms to find correlated factors that explain the incidence and prevalence of diseases.[8] Clinical trials are another promising field for AI, as smart algorithms can enhance eligibility criteria and help identifying the right volunteers for trials, which is often a limiting factor in rare diseases.

Finally yet importantly, **hardware technology** is already in place to substitute many human tasks, including some that were highly specialized. Self-running **robots** are expected to be increasingly relied upon in a variety of different medical settings, ranging from diagnosis and medical treatment, to surgery, with sector-specific applications, such as radiology or virology.[9]

[3] Muse, E.D.; Torkamani, A; Topol, E.J. (2018) "When genomics goes digital". *Lancet*. 391:2405; Rehm H.L. (2017) "Evolving health care through personal genomics". *Nat Rev Genet*. 18:259–67.

[4] Increasing globalization of healthcare solutions means that more and more health, genomic and biometric data are being collected and curated by insurances and private firms outside the public health systems. The situation is rapidly changing to the point that private companies currently held more data than public administrations even in countries with universal health care public systems such as Europe.

[5] In Computer Science, an algorithm is a list set of instructions, used to solve problems or perform tasks, based on the understanding of available alternatives. Blass, A.; Gurevich, Y. (2003) "Algorithms: A Quest for Absolute Definitions". *Bulletin of European Association for Theoretical Computer Science.* October 9, 81.

[6] Tubiana, M. (1995), *Les chemins d'Esculape : histoire de la pensée médicale*, Paris : Flammarion, 45.

[7] For instance, IBM's Watson for oncology. AIS is meant to support clinical users and hence directly influence clinical decision-making. The AIS would then evaluate the information and recommend the patient's care. The use of AI to assist clinicians in the future could change clinical decision-making and, if adopted, create new stakeholder dynamics. The future scenario of employing AIS to help clinicians could revolutionize clinical decision-making and, if embraced, create a new healthcare paradigm. Clinicians (including doctors, nurses, and other health professionals) have a stake in the safe rollout of new technologies in the clinical setting. See, Smith H. (2020) "Clinical AI: opacity, accountability, responsibility and liability". *AI Soc*. 36:535–45.

[8] The COVID-19 sped up new software that is now being developed and adapted to other infections. See Randhawa, G.S., Soltysiak, M.P.M., El Roz, H., de Souza C.P.E., Hill, K,A., Kari, L. (2020) "Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study". PLoS ONE 16(1): e0246465.

[9] Bohr, A., Memarzadeh, K. (2020) "The rise of artificial intelligence in healthcare applications." *Artificial Intelligence in Healthcare*, 25-60.

Medical ai adoption patterns

The trifold factor, data, algorithms and robots, will make practice and research at the health factor unrecognizable for us in a decade time. Major benefits are expected from the medical AI technologies in terms of enhancing the health care quality and costs of an aging population.

Yet, adoption patterns of medical AI could vary depending on political and juridical boundaries that will be adopted. As the political and juridical conceptual frameworks confronts AI technologies in the form of medical innovations, they should also be able to say something about the **societal choices** (even implicit) made in the course of technological innovation and the grounds for making those choices wisely. Two decades ago, philosophers of science showed how the strong interweaving of our culture with technology makes us inclined to accept all kinds of innovations. Because technological objects and processes have practical usefulness, says Winner (1989), they tend to be considered to be neutral in their moral position.[10] Thus, judgments about technology are usually built on a narrow basis, paying attention to practical issues such as whether or not the new device meets particular needs, works more efficiently than its predecessor works, makes a profit or provides a convenient service. Only later is the deeper meaning of the adoption of the said technology revealed, usually as a series of surprising "side effects" or "secondary consequences". Therefore, technology adoption should not be based only on needs, efficiency or profit, but it should be the result of a societal decision to be taken after a deliberative and fully informed democratic process, leading to a regulatory framework where key concerns identified are duly targeted.

Building an AI system that involves human beings as operational parts of that system brings a reconstruction of social roles and relationships. Often, this is the result of a new system's own operational requirements: it simply will not work unless human behavior changes to suit its form and process. Therefore, the very act of using the types of machines, techniques, and systems available to us generates patterns of activities and expectations that soon become "second nature." In fact, we use smartphones and computers in the conventional sense, but our world soon becomes one in which data processing and computing are life forms in the most powerful sense: life would hardly be thinkable without them.

In the same way, advances such as medical AI will bring about fundamental changes in the understanding of human life itself. Such advancements are not coming without challenges relating directly to the control of patient data and the safety of new AI systems, but more broadly, to the current model of health care that was construed after the II World War in Europe. This caveat suggests that we should pay attention not only to the certification of devices and algorithmic processes just in terms of practical benefits and risks of the devices (as the incoming EU regulations are proposing), but we should also consider the production of economical, psychological, social, and political conditions as a part of a very significant technical change.

[10] Winner, L. (1989) *The Whale and the Reactor. A Search for Limits in an Age of High Technology*, Second Edition, Chicago : University of Chicago Press, 13.

Privacy and control of the data

Recognizing the link between sharing of sensitive personal information and the due protection of the fundamental rights of citizens –the **traditional sphere of privacy** and health data protection– is crucial for the future of these technologies. Controllers of health data should bear in mind that health, genetic and biometric data are highly protected by the European General Data Protection Regulation (GDPR)[11] due to its sensitivity and potential harm that misuse can derive for the subjects.[12] However, in addition to biometric, health and genetic data –considered as medical data–, behavioral and psychological data should also be considered sensitive.[13] Medical AI is fed by data harvested from social media –such as images– and through wearable technologies.[14] Behavioral data, along with psychological ones captured by these means, are used to elaborate behavioral patterns and profiles that can be highly intrusive and lead to automated decision making on the subjects, as we shall latter see.

The General Data Protection Regulation prohibits health data processing unless the data subject is explicitly **consenting** to it.[15] Yet patients´ consent has been challenged as the most suitable legal base to treat medical and biometric data for medical treatment. Bearing in mind that processing data is necessary for treating a patient it is not much help to ask the patient if he or she consents on the use of her data for medical purposes. On the other hand, current data-consent procedures are dubious. Consent forms are supposed to be clear and easy to understand, but who has time to read so many words when the desired app is waiting to be downloaded in our device? Literature has highlighted the extreme devaluation of the informed consent forms for digital services through the Internet, which add discredit on the already poor quality consent that clinicians get from the patients.[16] To regain confidence on consent as a valid tool to control our data, scholars have proposed innovative consent procedures, which are adapted to the needs and expectations of the subjects, such as **dynamic consent**.[17]

---

[11] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, GDPR)

[12] However, medical records do contain some data that hold rights from third parties, such as observation and diagnostic remarks done by the clinicians, as well as administrative data about the treatment hold by the health providing service –insurer, clinic, etc.- and, also, metada about the processing systems which are protected by the IP rights of the data base owner. This is why sharing medical records information for treatment or research rises concerns relating to data ownership that are frequently discussed in the literature. See Mirchev, M. (2019) "Patient information ownership in the age of digital health and big data." *Eur J Public Health*. 29:ckz186.078. Kostkova, .P, Brewer, H., de Lusignan, S., Fottrell, E., Goldacre, B., Hart, G., *et al*. (2016) "Who owns the data? Open data for healthcare". *Front Public Health*. 4:7.

[13] GDPR, art. 4 (15)

[14] Wearables and self-quantifying apps measure ambience data and metadata such as geo localization, weather conditions, time, IP, cookies, etc. which can be also considered protected personal data when linked to the subject identification and her circumstances.

[15] GDPR, art. 9

[16] Capron, A.M. (2018), "The Future of Informed Consent." *The Journal of Law, Medicine & Ethics*, 46: 12-29

[17] Teare, H. J. A. *et al*. (2021) "Reflections on dynamic consent in biomedical research: the story so far". *European journal of human genetics: EJHG*, 29,4, 649-656.

As for the use of clinical data in research, other kinds of safeguards should be offered to discharge the clinicians of any illegitimate use of this data, instead of placing the burden of the responsibility on the consent of the patient. In particular, scientists have for long claimed open access to pseudonymised data. The General Data Protection Regulation, in spite of being more permissive than the previous Directive, does not specify the requirements that controllers must comply with to be authorized the processing of health data. Instead, the GDPR delegates this decision to Member States.[18] However, regulation issued by Member States is so disparate that it does not serve to promote international research and innovation in the field of Medical AI.[19]

For tackling with this handicap, a new legal framework (NLF) has been announced in order to share health data for research and innovation by building a common European Health Data Space **European Health Data Space**.[20] The first legal instrument to be adopted at the NLF has been the **Data Governance Act**[21] **(DGA).** The DGA develops a new framework — "data intermediation services" — for companies and individuals to engage in data sharing in a secure environment. This framework, based on the creation of a trusted and independent intermediary, aims to facilitate data sharing between data subjects and data holders, on the one hand, and data users, on the other hand. To ensure that the shared data is not used by entities providing the intermediary services themselves, the measure introduces structural separation requirements, which obliges intermediaries to be structurally separated from any commercial activity rather than offering intermediation services. On this point, the DGA also recognizes "data cooperatives" as services intermediaries, and thus empowers individuals to choose if using intermediaries instead of consenting large corporations to access and use their data. By doing this, the DGA empowers user's GDPR rights. In addition, with this blueprint, the DGA paves the way for the development of data stewardship models rooted in the collective exercise of users' rights, which are expected to introduce models that are more democratic and participative bringing opportunities for citizens regarding the use of their data for research. The conditions for the participation of the citizens in the governance of their data are expected to put an end to the power imbalance and restore the power of data to the citizens. An emerging trend on this domain -not just in medical AI, but also in health research- is to seek increased patient engagement, treating the patient as a stakeholder in research, not just a data source.[22] The perspective of patients is vital for gaining a real understanding of what security and privacy mean in the context of connected health, especially about health

[18] Hansen, J., Wilson, P., Verhoeven, E., Kroneman, M., Kirwan, M., Verheij, R., van Veen, E., (2021) *Assessment of the EU Member States' rules on health data in the light of GDPR, Netherlands Institute for Health Services Research*, NIVEL, DG Health and Food Safety, Specific Contract No SC 2019 70 02 in the context of the Single Framework Contract Chafea/2018/Health/03.

[19] Some regulations are relying on specific and explicit consent, while others are just saying that scientific research with sensitive data should not be done without the knowledge of the data subject. Although, all the regulations are acknowledging that use of data without the consent of the subject to solve a public health problem, such as COVID-19, may be justified with certain conditions.

[20] European Commission. White Paper on Artificial Intelligence - A European approach to excellence and trust. Brussels, 19.2.2020. Available at https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

[21]Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance) *OJ L 152, 3.6.2022.* ELI: http://data.europa.eu/eli/reg/2022/868/oj

[22] De Wit, M., Cooper, C., Register, J.Y. (2019) "Practical guidance for patient-centered health research". *Lancet*. 393:1095–6.

and biometric data harvested by commercial **well-being apps**.[23] Here, the role of big tech companies, such as Google, Apple, Meta, Amazon or Microsoft, who have all entered the digital health domain, remains in question.[24] These companies have the ability to influence the delivery of public and essential medical services inclining the balance to the side of private healthcare, where market strategies shift more power into corporate hands through the control of sensitive information about the citizens.

The Regulation also introduces a new concept "data altruism in the general interest", based on the idea of voluntary sharing for the common good by data subjects and data holders. This is applicable for interests related to health care or scientific research in the general interest. Data altruism will be implemented through "data altruism organizations", namely data pools administered by entities of not-for-profit nature which will be subject to transparency and structural separation requirements akin to those falling on data intermediation services. Data altruism organizations require the appointment of a legal representative in one of the Member States where data collection takes place, but need not necessarily be established in the EU.

Automated decision making

Automated decision-making is defined by the GPDR as any decision made without meaningful human involvement that has legal effects on a person, or similarly significantly affects him or her.[25] This definition may partially overlap with, or result from, profiling, but this is not always the case. Under the GDPR, **profiling** is the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location, or movements.[26]

The GDPR imposes specific requirements on profiling and automated decision-making. Clinicians and patients have a right to **understand** how a particular AI decision on a health matter has been made.[27] The General Regulation also declares the right to human

[23] Taylor, M.J., Wilson, J. (2019) "Reasonable expectations of privacy and disclosure of health data". *Med Law Rev.* 27:432–60.

[24] Laato, S., Islam, A., Islam, M., Whelan, E. (2020) "What drives unverified information sharing and cyberchondria during the COVID-19 pandemic?" *Eur J Inform Syst*. 29:288–305.

[25] GDPR, art. 22.

[26] GDPR, art. 29.

[27] GDPR, Art. 22(3) requires controllers to implement suitable measures to safeguard the data subject´s rights and freedoms and legitimate interests when their data is used for solely automated decision making, and to provide data subjects at least with the right to obtain human intervention and to contest the decision. GDPR, arts. 13, 14 obligation for controllers to disclose the existence of solely automated decision making, together with the logic involved in the automated decision-making. The information about the existence of solely automated decision making pursuant both arts. 13, 14 must be given at the time of collection, art.13, or within a reasonable time after obtaining the data, art. 14. GDPR applies to the processing of personal data in the context of an EU establishment, or when offering goods or services to, or monitoring the behavior of, individuals in the EU. The GDPR applies regardless of the means by which personal data are processed, and therefore applies when an AI system is used to process personal data (e.g., when using an AI system to make decisions on the priority of a medical treatment).

intervention, enabling the individual to **challenge** the automated decision.[28] A data subject will always have the right to demand human review of a fully automated decision, which has profound implications from a technology policy point of view. Thus, following GDPR rule, Member States should develop suitable measures regarding the right to obtain human intervention (Art. 22 (2) (b) GDPR). When the individual pertains to a vulnerable group, processing requirements should be subjected to a more stringent national control.[29]

Yet, the GDPR does not overrule any decision-making system in itself. This is different under the proposal of **Artificial Intelligence Act**[30], which intends to fill the gap on automated decision-making algorithms in the actual regulatory landmark. The new AI Act will also prohibit a number of AI systems that are deemed too risky under any circumstances. Most of these prohibitions are limited to AI systems used by public authorities or law enforcement. There are also prohibited AI systems that are relevant to the private sector, as for instance those that cause physical or psychological harm to an individual by deploying certain techniques (see below).

Despite AI already having achieved remarkable results in a range of decision-making systems, such as triage of patients in emergencies or surgical robots interventions, **explainability** of the decision-making process remains a challenging task, as many of these results have been achieved using "opaque algorithmic processing".[31] The use of **black-box techniques** where the system does not provide any information concerning how it arrived at the predicted value can lead to defenselessness of subjects. The traditional information imbalance between the clinician and the patient risks to be aggravated by algorithmic decision-making systems.[32]

---

[28] Bygrave, L.A. (2020), "Art. 22, automated individual decision making, including profiling, in The EU General Data Protection Regulation (GDPR): A Commentary", by Christopher Kuner, Lee A. Bygrave, Christopher Docksey, and Laura Drechsler, Oxford, 522-542.

[29] Brkan, B. (2019) "Do algorithms rule the world ? Algorithmic decision making and data protection in the frame of GDPR and beyond", *International Journal of law and information technology*, 27(2) 101

[30] Proposal for a Regulation of the European parliament and of the council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts COM/2021/206 final. On 21 April 2021, the European Commission issued a Proposal for a Regulation of the European Parliament and of the Council, laying down harmonized rules on artificial intelligence (said AI Act). This proposal is part of a broader legislative package known with the acronym NLF (New Legislative Framework) that aims to generate a new market for safe and reliable industrial products. The NLF frame is based on the risk-benefit analysis of the proposed technology. EU New Legislative Framework. Available at: https://ec.europa.eu/growth/single-market/goods/new-legislative-framework_en.

[31] As BALKIN has remarked behind algorithms are governments and businesses organized and staffed by human beings. The dream of the Algorithmic Society is the omniscient governance of society. It is way of governing populations. A characteristic feature of the Algorithmic Society is that new technologies permit both public and private organizations to classify and govern large populations of people. By governance, I mean the way that people who control algorithms analyze control, direct, order and shape the people who are the subjects of the data. People use algorithms to classify, select, comprehend, and make decisions about entire populations of people. This relationship is not simply a relationship of market profit. It is also a relationship of power and governance. Balkin, J. M. (2017) "The Three Laws of Robotics in the Age of Big Data", *Ohio State Law Journal*, Vol. 78, 1217-32

[32] "The AI knows a lot about you but you don´t know a lot about AI" (BALKIN, 2017). Moreover, the patient cannot monitor very well what the AI agent or algorithm does. There is an asymmetry of power and an asymmetry of information between operators and those acted on. This asymmetry is a central feature of the AI society –it is the asymmetry of knowledge and of power between the people and

High risk ai systems

The Draft AI Act defines AI system as **software** which is able to generate outputs (such as content, predictions, recommendations or decisions) influencing the environment it interacts with.[33] This broad definition of AI systems brings under its scope a very large number of software-based technologies, which so far have stayed out of the existing EU laws.[34] The proposed regulation assigns new, specific, legal and regulatory obligations to those manufacturing, importing or distributing AI system on the EU market. It further introduces technical and ethical standards for software, which are able to offer guidance as to liability of businesses involved (see more below) either in the AI systems themselves or in their use as components of products.

This new regulation relies on a risk-based classification of AI systems posing an 'unacceptable', 'high' or 'medium-low' risk to the Union's values and public interests. The first category ("unacceptable") creates a **blacklist of practices** ranging from AI systems that deploy subliminal techniques beyond a person's consciousness or that exploit any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, to certain 'real-time' remote biometric identification systems used in publicly accessible spaces. The prohibitions stem from the liberties´ undermining effects AI could have in terms of mass vigilance systems or public opinion manipulation strategies.[35]

The second category is represented by **high-risk systems**, which could nevertheless be accepted under stringent conditions. It covers, for example, AI systems used for emergencies´ triage; also, predictive policing uses of AI at the border through automated recognition of sensitive identity traits (like race, gender identity, disability); uses of AI to

---

the governors (public and private) of the Algorithmic society-. Balkin, Jack M. (2017) "The Three Laws of Robotics in the Age of Big Data", *Ohio State Law Journal*, Vol. 78, 1217-32

[33] The following techniques and approaches (used for software development) are covered by Annex I of the Proposal: machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; statistical approaches, Bayesian estimation, search and optimization methods.

[34] Under the proposed regulation, AI systems may be regulated in two different ways. As components of products that are already covered by the EU New Legislative Framework, such as medical devices and in vitro diagnostic medical devices, for those AI systems intended to be used as safety component of products that are subject to third party ex-ante conformity assessment, irrespective of whether the AI system is physically integrated into the product (embedded) or serves the functionality of the product without being integrated therein (non-embedded). As regards this first category and, more specifically, high-risk AI systems related to products covered by the EU New Legislative Framework, the requirements for AI systems set out in the Proposal will be checked as part of the existing conformity assessment procedures under the relevant legislation (e.g. Regulation (EU) 2017/745 on medical devices). This leads to an interplay of legal and regulatory requirements, which the proposed regulation solves as follows: safety risks specific to AI systems are meant to be covered by the requirements of the Proposal, whereas existing New Legislative Framework aims at ensuring the overall safety of the final product and therefore may contain specific requirements regarding the safe integration of an AI system into the final product.

[35] Harari, Y. (2020) "The world after coronavirus", *Financial Times,* https://www.ft.com/content/19d90308-6858-11ea-a3c9-1fe6fedcca75 .

determine access to essential public services; algorithms used for the risk assessment and pricing of health insurance, etc. These systems should undergo a strict certification procedure before being commercialized at the market. Particular attention is paid to post-market monitoring: the high-risk AI systems should be equipped with automatic recording of all events while the system is operating and throughout its lifecycle. In addition, high-risk AI systems are to be designed and developed to ensure sufficient transparency enabling users to interpret the outputs and to use them appropriately. For that purpose, the systems falling into this category of risk should be accompanied by detailed and comprehensible instructions in a digital format.

To indicate conformity with the proposed regulation, providers should obtain a specific CE marking for high-risk AI systems before they hit the market. This follows a conformity assessment procedure led by a notified body (designated under the proposed regulation and only for AI systems intended to be used for the remote biometric identification of persons) or by the manufacturer himself (for all high-risk AI systems) and the drawing up of an EU declaration of conformity. For instance, *ChestLink* is the first fully autonomous AI medical imaging product obtaining a CE mark.[36]

The third category is composed of **medium-low risk or non-high risk AI systems**, such as wellbeing wearables. In line with proportional approach, the regulation does not envisage an as strict regime. With regard to non-high-risk products, manufacturers are given the opportunity to self-regulate via non-binding codes of conducts and an overall lighter regime to comply with.

Health technology industry should thus prepare for the fact that development and market entry of high-risk AI systems are going to be significantly impacted -and likely delayed- by this new requirement following the entry into force of the proposed regulation. However, this regulation leaves a worrying gap for many discriminatory and surveillance practices used by governments and companies, in public events and other massive encounters, often with extremely harmful consequences for people. The proposal also fails to protect people from other harmful biometric methods (such as categorization), despite the significant threats posed to people's dignity and right to non-discrimination. Such practices can perpetuate harmful stereotypes by putting people into boxes of "male" and "female" based on a biometric analysis of their facial features, or making guesses about people's future behavior based on predictions about their race or ethnicity.[37] Moving forward, legislators must focus

[36] The producer of *ChestLink* claims that it dentifies CXRs with no abnormality and produces finalized patient reports without any intervention from the radiologist, reducing radiologist workload and enabling them to focus on cases with pathologies. The technical explanation is available at: https://oxipit.ai/products/chestlink/

[37] As *Acces Now* activist, Daniel Leufer has put it: "The proposal's treatment of 'biometric categorization,' defined as assigning people to categories based on their biometric data is very problematic. It lumps together categories such as hair or eye color, which can indeed be inferred from biometrics, with others like sexual or political orientation that absolutely cannot. Inferring someone's political orientation from the shape of their facial features is a revival of 19th century physiognomy, a dangerous, discredited and racist pseudoscience. Applications like this should be banned outright, but are currently only assigned minimal transparency obligations." Whilst the inclusion of rules to ensure that remote biometric identification (RBI) systems cannot be self-certified by whoever has developed them, and the fact that the exceptions to the ban on law enforcement uses don't apply unless each country specifically puts them into their national law, the AI Act nevertheless falls seriously short of safeguarding people's faces and public spaces. Shortly after the proposal launched, the European Commission's own supervisory authority, the

on preventing harms and clarifying the legal limits on unacceptable uses of AI. Legislators should strive for wide consultation with civil society and affected communities to set clear redlines in order to protect human rights.

Do artifacts have responsibilities?

Artificial Intelligence powered systems raise fundamental questions of liability in the event of errors and **harm for the patient.**

However, under the label "Medical AI" fall a broad variety of devices, which differ among one another for their technical features, diffusion, function and use; thus, no one-size-fits-all approach should be adopted.

Most used devices such as general expert systems and radiology computer assisted diagnosis tools **are not meant to substitute the doctor, but only work as intelligent assistants.** Under existing liability rules, medical practitioners remain the only persons in charge, and hence liable, for the negative consequences suffered by the patient.[38] This puts such systems in a peculiar situation, since when discussing liability for damages, we should consider both the existing rules of product liability, as well as those related to medical malpractice.[39] In these cases the system is not responsible for the final decision, but merely provides an analysis, which the doctor may rely upon. As a matter of fact, the expert systems and computer assisted diagnostic tools are employed by a human agent, who may be held liable for a behavior that was influenced, or should have been influenced, by the intelligent assistant.[40]

Nevertheless, the practitioner could argue that the malfunctioning or error in the system should be considered the main cause of the damage (algorithmic nuisance). In this case, producers may shield themselves from liability by claiming that the doctor was responsible for the final choice, if the doctor has relied on the system even if it would have been reasonable for him or her not to do so.

EDPS, strongly criticized the failure to sufficiently address the risks posed by RBI, urging that: "A stricter approach is necessary given that remote biometric identification [...] presents extremely high risks of deep and non-democratic intrusion into individuals' private lives." Legislators must reject the idea that some forms of remote biometric identification are permissible whilst others are impermissible; and instead implement a full prohibition on all forms of biometric mass surveillance practices in publicly accessible spaces by all public authorities and private actors. Leufer, D. « Artificial Intelligence », https://daniel-leufer.com/artificial-intelligence/

[38] Bertolini, A. (2020), "Artificial Intelligence and Civil Liability", requested by Policy Department for Citizens' Rights and Constitutional Affairs, European Commission, Directorate-General for Internal Policies, July, PE 621.926.

[39] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products. For some considerations on the applicability of the PLD to AI, see European Parliament Resolution on a civil liability Regime for artificial intelligence (2020/2014, INL). See also Bernhard A Koch, (2021) "Product Liability 2.0 – Mere Update or New Version?" in: *Liability for Artificial Intelligence and the Internet of Thing*s , Münster Colloquia on EU Law and the Digital Economy IV, 97 - 116

[40] Bertolini, A. Artificial Intelligence and Civil Liability, *op. cit*.

Yet, under contract law, the doctor/hospital may still sue the producer claiming that the system does not perform as expected and thus there is a case of lack of conformity or breach.

Against this sloppy background, medical practitioners will tend to adopt defensive medicine practices and decision-making, recommending a diagnostic test or medical treatment that is not necessarily the best option for the patient, but that mainly serves the function to protect the physician against the potential claims for damages by the patient or his family. This could happen both regarding the choice of using computer assisted diagnostic systems, as well as the decision to conform to the suggested diagnosis and treatment – in particular when requesting additional tests –, because not doing so may lead to increased chances of being exposed to liability, especially in systems, where the doctor is primarily sued in cases of medical malpractice. The physician could also use disclosure about her choice to conform or not to conform to the system's diagnosis as a way of shifting on the patient's informed consent the risk of adopting a specific curing strategy. Even the fact that the intelligent device used by the patient is private, could play an important role in shifting the burden of proof onto the patient or the producer.

This scenario is highly undesirable. To counteract it, it may be appropriate to revise existing liability rules – or anyway limit their exposure – as to shield doctors from responsibility connected to the use of medical systems, and rather hold strictly and absolutely liable either the hospital, or the producer or provider of the medical assistive technology, eventually under a specific form of producer liability.

Avoiding bias

Medical AI should be designed, developed and used in accordance with the fairness principle, by considering individuals´ reasonable expectations, by ensuring the use of AI systems remains consistent with their original purposes, and by taking into consideration not only the impact that the use of AI may have on the individual, but also the collective impact on groups and on society at large; thus, recognizing the need for delineation and boundaries on certain uses.

Efforts to adopt AI solutions moving away from traditional face-to-face medicine may increase already existing socioeconomic gaps between people who can easily access and use such services and those who cannot. As for instance, low levels of **digital literacy** in the general population, especially in the elderly, is a major contributing factor. The Spanish Organic Law for the Protection of Data and Digital Rights[41] refers to the minors and the elderly as especially vulnerable groups. As for the latter, the lack of digital knowledge can make it very difficult for them to have access to some services as health treatment or social security aids. This is why the Spanish Act has enunciated a list of new digital rights of citizens, including the right to have access to the Internet and the Digital Services Society in equal terms[42], and the right to be educated in the new Digital Technologies.[43]

[41] Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD).
[42] LOPDGDD, art. 81
[43] LOPDGDD, art. 83

There are also generalization issues associated with AI models made on "**limited" training sets**. A recent systematic review of deep learning solutions in medical images found that only a minimal number of studies in this field were of sufficient quality for clinical implementation.[44] Some data sets used to train machine learning based AI systems have been found to contain bias resulting in diagnostics or decisions, which can damage or discriminate against certain individuals or groups, potentially restricting the availability of certain health services or treatments. Medical AI systems could reinforce existing gender, racial and social bias in the research, detection and treatment of disease.[45] Take, for instance, the COVID-19 case. There is evidence that AI can help guide treatment decisions within COVID crisis; yet given the pervasiveness of biases, a failure to develop comprehensive mitigation strategies during the COVID-19 pandemic risks exacerbating existing health disparities.[46] The bias is quite apparent in terms of under representability of minorities in terms of infection rate, hospitalizations, and mortality. Many believe AI guided clinical decision-making for this novel disease, could result in the rapid dissemination of underdeveloped and potentially biased models, which may exacerbate the disparities gap. This is the case for countries which did not generate large amount of COVID data and could not share them accordingly in international diseases repositories.[47]

Unlawful biases or discrimination that may result from the use of AI should be reduced and mitigated by ensuring the respect of international legal instruments on human rights and non discrimination, investing on research into technical ways to identify, address and mitigate biases, taking reasonable steps to ensure personal data and information used in automated decision making is accurate, up to date and as complete as possible, and elaborating specific guidance and principles addressing biases and discrimination, and promoting individuals´ and stakeholders´ awareness.[48] Avoiding bias would also imply adopting measures to select data and to conduct studies on AI based systems, to compensate the hegemony and over representation of certain data.

---

[44] Barclay, L. and Zuanazzi, V. (2021), "Bias in medical imaging AI: Checkpoints and mitigation", November 24. Available at https://www.aidence.com/articles/bias-in-medical-imaging-ai/?cn-reloaded=1.

[45] Sonja Kelly and Mehrdad Mirpourian, primer on opening up new credit to women in emerging economies Women's World Banking February 2021.

[46] As Gutierrez *et al.* have shown Africa has produced very little amount of data in comparison with Europe. Gutierrez, B., Xu, B., Mekaru, S. (2020) "Epidemiological data from the COVID-19 outbreak, real-time case information". *Sci Data* 7, 106. See also, Röösli, E, Rice, B, Hernandez-Boussard, T. (2021) "Bias at warp speed: how AI may contribute to the disparities gap in the time of COVID-19." *J Am Med Inform Assoc.* 28(1):190-192.

[47] Schneeberger, D., Stöger, K., Holzinger, A. (2020). "The European Legal Framework for Medical AI". In: Holzinger, A., Kieseberg, P., Tjoa, A., Weippl, E. (eds) *Machine Learning and Knowledge Extraction*. *Lecture Notes in Computer Science*, vol 12279, 89.

[48] ICDPPC, *Declaration on ethics and data protection in AI. Declaration on Ethics and Data Protection in Artificial Intelligence*. 40 th International Conference of Data Protection and Privacy Commissioners, Tuesday 23 rd October 2018, Brussels

Conclusion

While technology is at the heart of any digital health system, the related transformations on health care that Medical AI is proposing cannot be viewed purely through a technological lens. Technology adoption patterns tell us that when a sophisticated new technique or instrument is adopted in medical practice, it transforms not only what doctors do, but also the ways people think about health, sickness, and medical care.

Our main concerns in terms of respect of human rights should be, first, **privacy protection and respect for the preferences of the citizens** regarding the use of their sensitive data. While the practices of informed consent and privacy by design are well established in digital health, there are still concerns surrounding patients' understanding of how their data are being processed and by whom in AI systems. Next generation of AI based digital technologies should support and engage users in such a way that fosters equality and inclusivity, resulting in AI healthcare solutions for all.

Second, in acknowledging that AI system (defined as software) may be a product, the AI Act Proposal strengthens **product liability risks for manufacturers or providers**. These requirements could lead to meaningful modernization of certain national product liability regimes, which have historically limped when regulating software products. Health technology industry should thus prepare for the fact that development and market entry of high-risk AI systems are going to be significantly impacted by this new requirement following the entry into force of the proposed regulation.

Third, risk based approach could be a solution to assess the admissibility of those devices. Yet, the Draft AI Act is excessively permissive with some AI systems, which could derive in **mass surveillance and citizen's behavioral control**, such as remote biometric recognition in public events.

Finally yet importantly, Medical AI systems could reinforce existing **gender, racial and social bias** in the research, detection and treatment of disease. Avoiding this bias would imply adopting measures to select data and to conduct studies on AI based systems, to compensate the hegemony and over representation of certain data.