Psychiatria Danubina, 2024; Vol. 36, Suppl. 2, pp 348-353
Conference paper

Review

ARTIFICIAL INTELLIGENCE IN HEALTHCARE AND REGULATION CHALLENGES: A MINI GUIDE FOR (MENTAL) HEALTH PROFESSIONALS

Allison Gilbert[1], Emanuela Pizzolla[2], Sofia Palmieri[3] & Giovanni Briganti[1]
[1]Department of Computational Medicine and Neuropsychiatry, Faculty of Medicine, University of Mons, Mons, Belgium
[2]University of Verona, Verona, Italy
[3]Faculty of Law and Criminology, University of Ghent, Ghent,Belgium

SUMMARY

Artificial intelligence (AI) offers new perspectives in the healthcare sector, ranging from clinical decision support tools to new treatment strategies or alternative patient remote monitoring. However, as a disruptive technology, AI is associated with potential barriers, limitations and challenges for appropriate integration in medical practice. To avoid potential patient safety risks and harm, a robust regulatory framework is crucial to guide health professionals in their AI adoption in clinical practice. The European Union offers a new legal framework for the development and deployment of AI systems, the AI Act. This regulation was approved in March 2024 and will be fully applicable by 2025 to ensure that AI technologies are safe, transparent, and respect fundamental rights. However, these new regulatory concepts may be obscure for clinicians. This article aims to provide health professionals with the preliminary key points of regulation needed to interact adequately with these new AI applications and consider the potential risks of AI systems to patient safety.

Key words: *Artificial Intelligence - AI Act – regulation - healthcare*

Abbreviations: *AI - Artificial Intelligence; GenAI - Generative Artificial Intelligence; Art – Article; FLOP - floating-point operations per second*

* * * * *

INTRODUCTION

Artificial intelligence (AI) has the potential to offer many innovative applications in the healthcare sector, ranging from administrative task management to the development of clinical decision support tools, precision medicine, and patient remote monitoring through virtual assistants (Alowayd et al. 2023). Recently, there has been increased interest in the use of AI in healthcare, with the emergence of generative AI (GenAI) models, which are characterized by their impressive ability to generate human-like texts, original images, videos, or music (Briganti 2024, Hu et al. 2023). While all these AI models offer promising opportunities to augment the skills of health professionals in clinical practice and may significantly impact the future development of original health applications, it is crucial to temper this enthusiasm with a thoughtful consideration of the new challenges they present. These include ensuring patient safety and privacy, as well as addressing significant ethical concerns inherent to their use in the medical field (Mesko & Topol 2023, Tan et al. 2024). Consequently, health professionals must operate within a well-defined regulatory framework to effectively balance the benefits and risks following the development of these emerging AI-based technologies.

The European Union's AI Act, approved on March 13, 2024, aims to establish a legal framework for the development and deployment of AI systems, including in the healthcare sector (Palmieri 2024a, European Parliament 2024a). This new regulation aims to ensure that AI technologies are safe, transparent, and respect fundamental rights while fostering innovation and competitiveness.

This paper aims to explore the regulatory landscape of AI applications in healthcare under the recent AI Act and address key questions frequently asked by clinicians when confronted with these technologies. The objective is to provide a mini guide of the main regulatory points to give clinicians the initial information they need to interact adequately with these new AI applications and consider potential risks to patient safety. To achieve this objective, the article is structured around different questions related to daily healthcare operations, which are subsequently analysed and discussed in the context of the AI Act regulation.

CLINICAL EFFICACY AND PATIENT SAFETY

The AI Act was recently published in the Official Journal of the European Union and will enter into force by February 2025. This regulatory framework aims to promote the implementation of human-centric and trustworthy AI, notably in healthcare (European Parliament 2024a). In medical practice, the integration of disruptive AI systems implies new perspectives to ensure the efficacy of these clinical applications and

patient safety. The evaluation of patient safety risks is a crucial step for ensuring appropriate quality of care following AI implementation (Ratwani et al. 2024). The first three questions of this mini guide address patient safety concepts, risk identification and clinical validation.

Question 1: *"How are AI systems classified in the AI Act?"*

Before fully responding to this question, it is important to understand how the AI Act defines an "AI system". In Title I, Article (Art.) 3 of the AI Act, an "AI system" is defined as *"a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments"* (European Parliament 2024b). If AI systems could thus influence the patient environment, this definition – when reflecting on AI in healthcare - must be read in conjunction with Hippocrates' principle of *"First, do no harm"*. The classification of AI systems by the AI Act relies on risk identification to assess potential harm based on the probability of occurrence and the severity of the adverse event (European Parliament 2024a). Different risk categories are identified, ranging from minimal or no risk, limited risk to high risk and unacceptable risk.

Title II, Art. 5 of the AI Act provides criteria for an AI system to be categorized as unacceptable. For example, manipulative AI applications or systems that exploit patient vulnerabilities will fall into this category.

Annex III of the AI Act offers a list of high-risk AI system criteria, including, for example, remote biometric identification systems or, in the healthcare sector, applications to triage emergency situations (European Parliament 2024a). Besides the implicit references made in Annex III to the healthcare sector, the classification of AI used in healthcare as high-risk is primarily due to the AI Act referencing the Medical Device Regulation to define the boundaries for this risk category (Palmieri 2023). Consequently, these systems are subject to rigorous scrutiny, including transparency, accountability, and human oversight requirements. This proactive stance highlights the importance of balancing AI's immense potential in healthcare with the imperative to mitigate associated risks.

Question 2: *"What are the requirements for an AI model to be approved for healthcare applications?"*

AI systems for healthcare applications are mainly considered high-risk and have a set of compliance requirements identified in Chapter 2 of Title III, which must be fulfilled before being put into service or entering the EU market (Figure 1). The requirements tackle risk management (Art. 9), data training and data governance (Art. 10), technical documentation (Art. 11), record-keeping of (adverse) events (Art. 12), transparency and instructions for use (Art. 13), human oversight (Art. 14), accuracy, robustness and cybersecurity (Art. 15) (European Parliament 2024b).

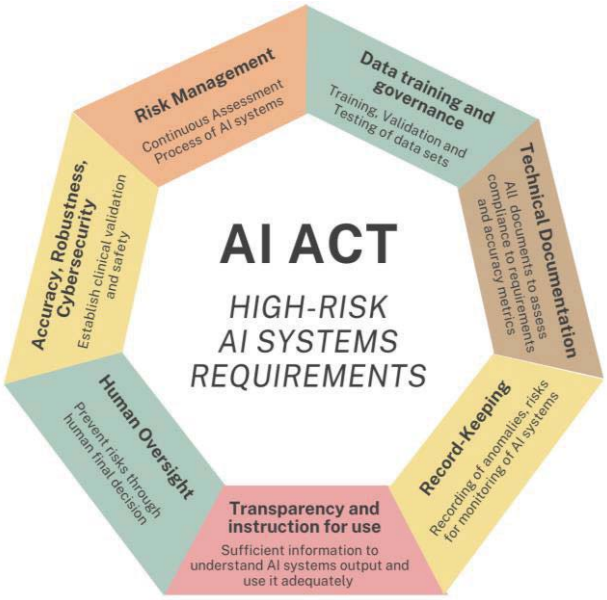

Figure 1. Requirements and their description for high-risk AI systems, based on the AI Act

Moreover, specific needs exist for regulation as certain AI models are dynamic (such as GenAI), possessing the capability to *"adapt and make decisions based on their experience and interactions within the system"* (Templin et al. 2024). Consequently, monitoring practices must also be dynamic to align with the evolving nature of the GenAI system. To ensure the safety and reliability of AI systems throughout their entire life cycle, the AI Act provides a post-market monitoring system to identify and mitigate foreseeable problems encountered during use (European Parliament 2024a).

It is important to note that if an AI system used in medicine also qualifies as a medical device under the Medical Device Regulation, it must adhere to the requirements set forth by the Medical Device Regulation.

Question 3: *"How do we, as healthcare professionals, consider AI systems as validated for use in clinical practice, and how is the validation process performed?"*

For high-risk models, regarding Title III, Art. 11 and 13, AI models must be put into service with sufficient information to users and transparency regarding their technical characteristics. Annex IV of the AI Act contains a list of what users must be aware of when they are confronted with the AI model. Other authors have previously claimed the importance of using a minimum

information checklist for AI applications (Norgeot 2020). With the recent AI Act, clinicians must check different points before implementing an AI model in practice. These include: (1) a description of the AI system, (2) a description of the elements and process of development of the AI system (methods), (3) a description of the monitoring process, (4) a description of the performance metrics or (5) a description of the risk management system (European Parliament 2024b).

However, several questions persist. Title III, Art. 15 argued for AI models that respond to appropriate accuracy and robustness (European Parliament 2024b). However, if accuracy metrics are cited as important elements to check, the methodologies used to obtain these metrics are not detailed. AI experts and researchers must think about the correct methods to validate AI models and achieve appropriate and general AI quality standards (Kuziemsky 2024).

The method of validation is also of the utmost importance. The AI Act suggests in Title VI, Art. 57, the establishment of at least one regulatory sandbox at a national level (European Parliament 2024b). These sandboxes are simulation techniques that allow appropriate validation of AI systems. However, real-world condition testing is also crucial for evaluating the robustness and reliability of these models. The AI Act, Art. 60 also aims to regulate real-world testing of AI even if it differentiates from entry into the market. A real-world testing plan would have to be adapted and validated before testing high-risk AI models in clinical practice (European Parliament 2024b).

AI COMPETENCIES AND RESPONSIBILITY FOR AI

Another fundamental question regarding AI use is the required competencies of health professionals. Zhang et al. described in their article two levels of factors that can influence trustworthy medical AI: the design level corresponding to the technical characteristics of the AI system and the application level related to the impact of humans on the use of AI (Zhang et al. 2023). This second section of the mini guide focuses on the competencies required to be adequately prepared for AI use and the associated responsibility for use and training.

Question 4: *"Should we, health professionals, acquire new expertise to use AI applications in our practice?"*

The requirements for high-risk AI applications in healthcare have been previously described with reference to Title III, Chapter 2, Art. 14 of the AI Act mentioning the concept of human oversight. However, point 4 of Art. 14 is of particular interest when considering the deployment of AI systems in healthcare. This point describes to what extent the person (or people) responsible for human oversight needs to acquire specific competencies to address AI challenges. In accordance with the AI Act, the expert responsible for human oversight or the "human in the loop" should (a) *"properly understand the relevant capacities and limitations of the high-risk AI system and be able to duly monitor its operation"*, (b) *"remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system (automation bias)"*, (c) *"correctly interpret the high-risk AI system's output", (d) "decide, in any particular situation, not to use the high-risk AI system or to otherwise disregard, override or reverse the output of the high-risk AI system"*, and (e) *"intervene in the operation of the high-risk AI system or interrupt the system through a 'stop' button or a similar procedure that allows the system to come to a halt in a safe state"* (European Parliament 2024b).

The question following these novel requirements presented by the AI Act is the level of competencies that health professionals should acquire to address the specific challenges of AI in clinical practice. The AI Act suggests in Title I, Art. 4, the need to ensure an adequate AI literacy among users of AI systems (European Parliament 2024b). However, there is a lack of a universally accepted set of expertise that health professionals should master in interactions with AI-based tools. Several authors suggested defining key elements required to help health professionals to answer fundamental questions about AI tools, such as learning basic knowledge of AI, discerning whether to use them or not and in which circumstances, being aware of their bias and limitations, and adequately considering how to communicate their use to patients or understanding the social and ethical implications of AI. However, according to Liaw et al., concrete expertise in AI should be limited to specifically trained professionals who may preserve the global process (Liaw et al. 2022, Russel et al. 2023).

The main challenge of the new AI Act regulation in the healthcare sector is how to evolve the healthcare system to meet these requirements. The healthcare system will have to train specific AI experts to promote appropriate implementation across hospitals and the global sector and increase the basic knowledge of health professionals to prepare them for the AI revolution. Training opportunities must be integrated across all levels of medical education, including customized programs for undergraduates and graduates in the medical field (Liaw et al. 2022). Considering the rapid evolution of AI-based tools, dynamic training programs should improve over time and adapt to new challenges. Some initiatives already exist worldwide; for example, in Belgium, mandatory AI training is starting to be implemented in medical education (Pizzolla 2023).

However, at this point, one question remains: Is the system ready to invest massively in AI training to achieve the "human in the loop" objective? One solution suggested for the future is the conceptualization of a "team in the loop" with not only health professionals but also multiple professionals with different expertise to respond adequately to the human oversight objective (Palmieri 2024b).

Question 5: *"What should we be careful about when deploying AI in care settings?"*

Responsibility for AI in healthcare can be divided into three distinct aspects: the deployment of AI tools within hospitals, and the training of users to ensure proper use of AI tools.

Title II, Art. 29 of the AI Act outlines obligations for users of high-risk AI systems. Healthcare professionals, when intended as AI users, must ensure that these systems are operated according to the provided instructions. They should effectively monitor AI systems to guarantee proper functioning and report any malfunctions to the provider. This responsibility ensures the safe and effective use of AI tools in clinical practice.

AI providers have significant obligations to develop and provide AI tools. Art. 14 paragraph 1, Title II, states that *"High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which they are in use"*. This suggests that providers must ensure that high-risk AI systems are designed with adequate safeguards and control measures to minimize risks to health, safety, and fundamental rights. Additionally, Title II, Art. 17, requires providers to ensure compliance with the AI Act, conduct risk assessments, and verify that the systems are safe for use.

Finally, Title I, Art. 4 of the AI Act emphasizes the importance of ensuring that employees using AI systems possess the necessary technical knowledge, experience, education, and training to operate them safely and effectively. Providers must ensure that users receive clear information about the capabilities and limitations of AI systems, including providing training and resources to help users understand how to use the AI tools effectively and safely (Title II, Art. 25).

In conclusion, although the AI Act provides comprehensive parameters regarding the division of different responsibilities, there remains a significant gap regarding the responsibilities that AI users, providers, and experts should have, especially in a clinical context. Healthcare professionals should be considered fully responsible for patient safety and the decision-making processes. Still, in case the decisions are based on AI outputs, organizations should be accountable for providing adequate training and education on AI tools to ensure staff competency and confidence in using these technologies. Supporting this distinction of responsibilities, the NHS AI Laboratory report emphasizes that all healthcare staff should receive training in AI, with additional specialist training for those using AI tools in clinical practice (NHS AI Lab report 2022). Consequently, from an operational point of view, regarding the AI Act requirements, hospital managers will have to adapt to these new regulations and refer to specific expertise available in their staffing to supervise the deployment and training of the users.

CONCEPTIONS AND FEAR OF AI

The final section of this mini guide offers perspectives regarding health professionals' conceptions and fears of AI in the light of the recent AI Act recommendations. Indeed, authors have reported fears or misconceptions about AI with recurrent questioning about how we must now perceive the integration of AI and, if one day, AI will replace humans (Carpio 2023).

Question 6: *"Should we, health professionals, be more afraid of GenAI than other AI models?"*

Generative AI has different specificities and challenges than to other AI models and may lead to other anomalies or adverse events (Templin 2024). Accordingly, referring to the AI Act, the distinction between AI systems and general-purpose AI models must be detailed. The AI Act, in Title I, Art. 3, provides a definition of general-purpose AI models, identifying these as AI systems that can perform generally applicable functions such as image/speech recognition, audio/video generation, pattern detection, question answering, translation, etc., and can have multiple intended and unintended purposes (George 2023). For example, an application such as ChatGPT, is considered a general-purpose AI model and will require more attention than other non-general-purpose models, even those considered as high-risk. This is why following the launch of ChatGPT in late 2022, the EU legislator had to adapt the text of the AI Act to the unique features that set GenAI apart from the AI systems seen thus far and subject these to a different regime under the AI Act (Novelli 2024).

The AI Act acknowledges that models created by GenAI with "high-impact capabilities" could be a potential threat to public health, safety, security, fundamental rights, and society (Madiega 2024). These "high-impact capabilities" are related to the computing power of the model or, in technical words, the floating-point operations per second (FLOPs). The classification for GenAI is thus not based on risk classes but rather concerns the possibility that general-purpose AI presents "high-impact capabilities" and could lead to

"systemic risks". Therefore, GenAI providers must notify the European Commission if their model is trained using a total computing power exceeding 1025 FLOPs. When they exceed this threshold, it will be presumed that the model is a GenAI model posing systemic risks. In this case, the manufacturer will have to comply with a different set of requirements. Thus, systemic-risk GenAI model providers must continuously assess and reduce the risks they pose and ensure cybersecurity protection. This entails various actions, such as monitoring, documenting, and reporting severe incidents such as violations of fundamental rights. Additionally, corrective measures must be implemented to address any issues that may arise.

Notably, general-purpose AI models are neither exempt from the requirements of high-risk AI nor exempt from the regulations designed for the features of GenAIs when they are medical devices and, thus, high-risk. Instead, they are subject to two layers of regulation.

Question 7: "*What does the Act AI say about AI replacing humans?"*

The recent AI Act (Art. 14) highlights the capital importance of human oversight regardless of the level of autonomy of the AI systems. Based on this regulation, human expertise will always be mandatory for patient safety. As AI evolves in medicine, health professionals should adopt the vision of augmented medicine with AI as a brilliant tool to augment their skills in daily practice.

CONCLUSION

Artificial intelligence in healthcare will undoubtedly lead to significant changes in clinical practice. However, these changes require an appropriate regulation process to ensure patient safety. The recent European AI Act offers new perspectives for understanding the complexity of AI regulation in the healthcare sector. However, as AI applications continue to spread among the medical field, a multidisciplinary collaboration between AI developers, healthcare providers, regulators, and patients will be essential to fully realize its benefits while considering and mitigating potential harm. Through such collaborative efforts, AI can become a cornerstone of modern healthcare, improving outcomes and enhancing the quality of care for patients worldwide. There is still a long road ahead for the optimal integration of AI and regulation. We are now at a crossroad where we need to implement these regulation strategies in practice and think about how these strategies will be integrated into the healthcare system alongside other existing regulations, such as medical device regulation or general data protection regulation.

Acknowledgements: *None.*

Conflict of interest: *None to declare.*

Contribution of individual authors:

Allison Gilbert: conceptualization, writing - original draft, writing - review & editing.

Emanuela Pizzolla: writing - original draft, writing - review & editing.

Sofia Palmieri: writing - original draft, writing - review & editing, supervision.

Giovanni Briganti: conceptualization, writing - review & editing, supervision.

References

1. *Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohared SN et al.: Revolutionizing healthcare : the role of artificial intelligence in clinical practice. BMC Medical Education 2023;23:689*
2. *Briganti G: How ChatGPT works: a mini review. European Archives of Oto-Rhino-Laryngology 2024; 281:1565-1569*
3. *Carpio EJT: Overcoming Fear, Uncertainty and Doubt: Artificial Intelligence (AI) and the value of trust. Cureus 2023; 15:e45576*
4. *Epstein Z, Hertzmann A, Investigators of Human C, Akten M, Farid H, Fjeld J, et al.: Art and the science of generative AI. Science 2023; 380:1110–1*
5. *European Parliament: 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) Text with EEA relevance. Official Journal of the European Union 2024; Available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj*
6. *European Parliament: 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)Text with EEA relevance. 2024; Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_EN.pdf*
7. *George A, George AS, and Gabrio Martin AS: A Review of ChatGPT AI's Impact on Several Business Sectors. Partners Universal International Innovation Journal 2023*
8. *Hu K: ChatGPT sets records for fastest-growing user base – analyst notes. Reuters 2023; Available at: https://www.reuters.com/ technology/chatgpt-sets-record-fastest -growing-user-base-analyst-note-2023-02 -01/*
9. *Liaw W, Kueper JK, Lin S, Bazemore A, Kakadiaris I: Competencies for the Use of Artificial Intelligence in Primary Care. Ann Fam Med 2022; 20:559-563*

10. *Madiega AT: Artificial Intelligence Act. European Parliamentary Research Service 2024. Available at https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021) 698792*
11. *Mesko B, Topol EJ: The imperative for regulatory oversight of large language models (or generative AI) in healthcare. Npj Digital Medicine 2023; 6:120*
12. *Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco R, et al.: Minimum information about clinical artificial intelligence modeling: the MI-CLAIM Checklist. Nat Med 2020; 26:1320-1324*
13. *Novelli C, Hacker P, Morley J, Trondal J, Floridi L: A robust governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities. 2024; Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4817755*
14. *Palmieri S: A brief reflection on the approved AI Act and Healthcare AI. MetaMedica 2024a; Available Online at: https://www.metamedica.ugent.be/sample-page/blog/ai-act-and-healthcare/*
15. *Palmieri S: Ensuring the trustworthy use of Medical AI: a legal perspective. Ghent University 2024b; In press*
16. *Palmieri S, Goffin T: A blanket that leaves the feet cold: Exploring the AI Act Safety Framework for medical AI. European Journal of Health Law 2023; 30:406-427*
17. *Pizzola I, Aro R, Duez P, De Lièvre B, Briganti G: Integrating Artificial Intelligence into Medical Education. Jl. of Interactive Learning Research 2023; 34:401-424*
18. *Ratwani RM, Bates DW, Classen DC: Patient Safety and Artificial Intelligence in Clinical Care. JAMA Health Forum 2024; 5:e235-514*
19. *Reddy, S: Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. Implementation Sci 2024; 19:27*
20. *Russell RG, Lovett Novak L, Patel M, Garvey KV, Craig KJT, Jackson J, et al: Competencies for the Use of Artificial Intelligence-Based Tools by Health Care Professionals. Acad Med 2023; 98:348-356*
21. *Tan S, Xin X, Wu D: ChatGPT in medicine: prospects and challenges: a review article. International Journal of Surgery 2024; 110:3701-3706*
22. *Templin T, Perez MW, Sylvia S, Leek J, Sinnott-Armstrong N. Addressing 6 challenges in generative AI for digital health: A scoping review. PLOS Digit Health 2024; 3:e0000503. Published 2024 May 23. doi:10.1371/journal.pdig.0000503*
23. *Zhang J, Zhang Z: Ethics and governance of trustworthy medical artificial intelligence. BMC Medical Informatics and Decision Making 2023; 23:7*

Correspondence:
Giovanni Briganti, MD, PhD
Department of Computational Medicine and Neuropsychiatry
Faculty of Medicine, University of Mons
Avenue du Champs de Mars 6, 7000 Mons, Belgium
E-mail: giovanni.briganti@hotmail.com