Special issue

Informed consent and artificial intelligence applied to RCT and Covid-19

Alberto Eugenio Tozzi, Giulia Cinelli*

ABSTRACT: Artificial intelligence (AI) tools allow to extract knowledge from big data and are increasingly used for research purposes applied to -omics, diagnostic images, complex patterns of diseases and system medicine, drug development, robotics, and other topics. The management of big data, largely made of individual clinical data, poses specific ethical challenges that must be addressed in research studies and that should be reflected in the informed consent process. Explaining the mechanisms used by AI algorithms in supporting clinical decision making may be particularly difficult because of the opacity of its process. Moreover, depending on the quality of data feeding their algorithms, AI applications may result in errors. As the General Data Protection Regulation (GDPR) includes the possibility that a patient withdraws his/her informed consent from a study, it may be challenging to update AI algorithms accordingly. On the other hand, AI tools may help support the recruitment and retention of participants in clinical trials matching eligibility criteria with individual data collected for clinical purposes in electronic health records, and improve data collection and analytics. The possibility to stream data from wearable devices offers the possibility to generate large data volumes relevant to Patient Reported Outcomes feeding AI predictive algorithms. The Covid-19 pandemic has promoted the application of digital tools and of AI in clinical trials in order to limit personal contacts. The pressure exerted by the pandemic will possibly speed up the adoption of AI solutions for clinical trials and will highlight their potential ethical implications.

KEYWORDS: Artificial intelligence, Covid-19, ethics, informed consent, randomized clinical trial

SUMMARY: 1. Introduction – 2. Applications of AI in healthcare – 3. Ethical challenges of AI – 4. The risks of AI in healthcare – 5. Emerging uses of AI in research studies – 6. The Covid-19 pandemic – 7. GDPR and AI in healthcare – 8. Shaping the informed consent for AI interventions – 9. Conclusions.

ISSN 2284-4503

* *Alberto Tozzi, Giulia Cinelli, Multifactorial and Complex Diseases Research Area, Bambino Gesù Children's Hospital, Rome, Italy. Mail: albertoeugenio.tozzi@opbg.net, giulia.cinelli@opbg.net.* *This essay is developed within the European project "Improving the guidelines for Informed Consent, including vulnerable populations, under a gender perspective" (i-CONSENT), funded by the European Union framework program H2020 (Grant Agreement n. 741856). The article was subject to a double-blind peer review process. The Authors thank the Reviewers for their comments.*

1. Introduction

The very essence of artificial intelligence (AI) applications in health care relies on the availability of big data generated by a network of multiple stakeholders engaged in data sharing. As AI is already pervasive in our daily life in other domains, we expect that this technology will become common in healthcare as well. The impact of AI in healthcare may be enormous if we will be able to properly share the largest amount of data possible as it happened when the exponential computing power of the internet skyrocketed in the early nineties and one million computing machines were connected on the web. As health care is typically organized in silos of data, overcoming this barrier to appreciate a significant impact on population health will be a challenge with a number of ethical implications, including the informed consent process.

Many health care problems could be addressed by AI applications that may resolve several high priority issues, from drug development, to prediction of clinical severe events, to what is one of the most abused terms of the last years: precision medicine. Indeed, a deep integration of AI in healthcare processes and research is yet to come, the main barrier being the lack of interoperability of data repositories.

The Covid-19 pandemic has increased the sense of urgency to address healthcare solutions that are still unresolved. On the other hand, the pandemic has accelerated and favored the application of digital tools using AI that may be suitable to support mitigation strategies for the pandemic. For this reason, we expect that the hype of digital and AI tools will soon result in using them in routine clinical care. Nevertheless, novel AI medical interventions will require proof of safety and efficacy through clinical trials. AI is expected to become one of the many tools available to deliver preventative, diagnostic and therapeutic interventions, but contrary to traditional medical interventions, being a software. Yet, the many specificities of AI tools and their intersection with human decisions require tailored ethical strategies. All these circumstances call for a deep reflection on the ethical implications of AI in healthcare to avoid that the promises of technologies and the hope for effective novel tools to combat the pandemic will obscure the basic principles of ethics in the interest of patients. Moreover, because of their nature, AI applications in healthcare may result in ethical and legal challenges regarding responsibilities and liability. We need to focus on the specificities that informed consent has in these particular circumstances as it is the most important process through which patients make their decisions in healthcare.

2. Applications of AI in healthcare

AI techniques vary in their mechanism and deployment in healthcare. Algorithms may use machine learning (ML) or deep learning (DL), with different levels of complexity and performance. An area where AI is applied is interpretation of natural language with Natural Language Processing (NLP) techniques. Often, a combination of these technologies is applied in robotics[1].

[1] ANTHONY C. CHANG, *Intelligence-Based Medicine: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare*, Elsevier Science, Amsterdam, 2020, pp. 534.

The potential applications of AI in healthcare are almost limitless[2] and its impact has become clear for both clinicians, health systems and patients[3]. The most exciting advances have been observed in the field of diagnostic image interpretation. It is already possible to automatically classify retinal photographs of diabetic patients and diagnose macular edema or retinopathy[4,5]. Several companies have already developed systems that are embedded into the radiologic equipment and that support the detection of cardiovascular diseases and other conditions, such as lung nodules or breast mass[6,7]. A similar progress has been observed in dermatology, where AI applications proved to be more accurate than a human observer in the detection of many types of skin lesions[8]. AI has also been successfully applied to pathology for classification of biopsies and detection of oncologic anomalies[9].
A more popular application of AI regards the interpretation of complex genetic patterns that may be associated with specific diseases[10]. Moreover, AI may help understanding the role of even more complex patterns including gene expression, protein abundance levels and methylation profiles in predicting several diseases (omics approaches), with implications in the discovery of novel biomarkers of disease[11].
AI is also helpful in predicting the prognosis of patients based on information recorded in the electronic health records (EHR). With this approach, it is possible to predict mortality, readmission and length of stay of complex patients, helping to implement appropriate countermeasures[12].

---

[2] K. H. Yu, A. L. Beam, and I.S. Kohane, *Artificial Intelligence in Healthcare*, in *Nature Biomedical Engineering* 2, no. 10 (October 2018): 719–31.
[3] E. J. Topol, *High-Performance Medicine: The Convergence of Human and Artificial Intelligence*, in Nature Medicine 25, no. 1 (January 2019): 44–56.
[4] M. D. Abràmoff et al., *Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning*, in *Investigative Ophthalmology & Visual Science,* vol. 57, no. 13 (1 October 2016): 5200–5206.
[5] A. Grzybowski et al., *Artificial Intelligence for Diabetic Retinopathy Screening: A Review*, in *Eye* vol. 34, no. 3 (March 2020): 451–60.
[6] B. van Ginneken et al., *Off-the-Shelf Convolutional Neural Network Features for Pulmonary Nodule Detection in Computed Tomography Scans*, in *IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015, 286–89.
[7] R. K. Samala et al., *Mass Detection in Digital Breast Tomosynthesis: Deep Convolutional Neural Network with Transfer Learning from Mammography*, in *Medical Physics* vol. 43, no. 12 (December 2016): 6654.
[8] A. Esteva et al., *Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks*, in *Nature* 542, no. 7639 (February 2017): 115–18.
[9] G. Litjens et al., *Deep Learning as a Tool for Increased Accuracy and Efficiency of Histopathological Diagnosis*, in *Scientific Reports* vol. 6, no. 1 (23 May 2016): 26286**;** B. E. Bejnordi et al., *Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer*, in *JAMA* vol. 318, no. 22 (12 December 2017): 2199–2210.
[10] D. Quang, Yifei Chen, and X. Xie, *DANN: A Deep Learning Approach for Annotating the Pathogenicity of Genetic Variants*, in *Bioinformatics* (Oxford, England) 31, no. 5 (1 March 2015): 761–63; D. Quang and Xiaohui Xie, *DanQ: A Hybrid Convolutional and Recurrent Deep Neural Network for Quantifying the Function of DNA Sequences*, in *Nucleic Acids Research* 44, no. 11 (20 June 2016): e107.
[11] N. Biswas and S. Chakrabarti, *Artificial Intelligence (AI)-Based Systems Biology Approaches in Multi-Omics Data Analysis of Cancer*, in *Frontiers in Oncology* 10 (2020): 588221.
[12] J. M. Karnuta et al., *The Value of Artificial Neural Networks for Predicting Length of Stay, Discharge Disposition, and Inpatient Costs after Anatomic and Reverse Shoulder Arthroplasty*, in *Journal of Shoulder and Elbow Surgery* 29, no. 11 (1 November 2020): 2385-94.

AI is also used to interpret continuous data streams of vital signs and other information, such as voice and movement generated by wearable devices. We started only recently to collect these data and to use them for diagnostic purposes often in complement with traditional clinical data[13].
Robotic systems, that are gradually taking place in surgery, offer another example of application of AI, taking also into consideration that it is likely that in the future robotic surgery will become autonomous[14]. Automation, however, is also found in technologies that interact with humans through natural speech. The capacity of AI to process natural language allows to develop social robots or chatbots[15] that may be used for health interventions.
Finally, a new class of therapies is becoming available, namely digital therapeutics. These therapies include as an active principle an AI software instead of a chemical or biological compound. Their application is mainly in the domain of behavioral treatments and support of adherence to pharmacological therapies[16].

3. Ethical challenges of AI

One known issue of AI interventions is the "black box" problem. Although AI systems are fed with known input data, and they generate explicit output data, the logic behind an AI algorithm may be opaque. This makes it difficult to understand, explain how, and the reason why an algorithm arrived at a specific result[17]. Indeed, AI algorithms do not apply synthetic rules, but they are rather trained on examples that may be difficult to reduce to simple logic.
According to the latest European Commission guidelines, trustworthy AI should be "(1) lawful, respecting all applicable laws and regulations; (2) ethical, respecting ethical principles and values; (3) robust, both from a technical perspective while taking into account its social environment"[18]. Moreover, both the Food and Drug Administration (FDA) and the European Commission consider transparency a fundamental requirement for AI systems in healthcare[19]. Transparency is intended as

[13] A.C. CHANG, *Intelligence-Based Medicine: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare*, Elsevier Science, Amsterdam, 2020, pp. 534.
[14] A. SHADEMAN ET AL., *Supervised Autonomous Robotic Soft Tissue Surgery,* in *Science Translational Medicine 8, no. 337 (4 May 2016): 337ra64*; R. ELEK ET AL., *Recent Trends in Automating Robotic Surgery,* in *2016 IEEE 20th Jubilee International Conference on Intelligent Engineering Systems (INES), 2016, 27–32.*
[15] A.C. CHANG, *Intelligence-Based Medicine: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare*.
[16] O. SVERDLOV ET AL., *Digital Therapeutics: An Integral Component of Digital Innovation in Drug Development,* in *Clinical Pharmacology and Therapeutics* 104, no. 1 (July 2018): 72–80.
[17] M. CARABANTES, *Black-Box Artificial Intelligence: An Epistemological and Critical Analysis,* in *AI & SOCIETY 35* (1 June 2020).
[18] HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *Ethics Guidelines for Trustworthy AI. Shaping Europe's Digital Future*, April 2018, available at https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (last accessed May 10th, 2021); EUROPEAN COMMISSION, *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*, 2021, available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELLAR%3Ae0649735-a372-11eb-9585-01aa75ed71a1 (last accessed May 10th, 2021).
[19] U.S. FOOD AND DRUG ADMINISTRATION, *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)-Discussion Paper and Request for Feedback,* 2019, https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-

ISSN 2284-4503

clarity of the AI systems themselves, of the kind of data used and the processes they follow to reach a result, and should apply to the output and to the functions of the software and its modifications over time to create trust among users[20].

Explainability can be considered as "a characteristic of an AI-driven system allowing a person to reconstruct why a certain AI came up with the presented predictions"[21]. While traditional algorithms are characterized by an inherent explainability[22], in DL algorithms such as artificial neural networks, explainability is not inherent to the system, but is approximated, due to the complex characteristics of the system itself. When it comes to medical AI-driven procedures, such as those implied in supporting clinical decision, lack of explainability may have important legal and ethical implications[23].

It may happen that AI algorithms are biased because of a non-sufficiently diverse training set or missing data, leading to disparities when applied to healthcare. For example, individuals from disadvantaged communities may access multiple institutions (i.e. less likelihood to find patients who have data in different EHRs), or data about race/ethnicity, socioeconomic status may be missing in the EHR leading to underrepresentation and bias toward these communities[24]. Data availability, data collection from minorities and specification of the population for which the algorithms is developed, are examples of standards and measures that could address some of these issues[25].

Big data needed for artificial intelligence applications in healthcare may fall under the control of various organizations, which may share these data for health research. Moreover, especially when data is collected from multiple data sources, identification or re-identification, which is technically feasible, may be useful for AI to recognize and exploit associations among data sources. If we consider the example of research in genomics or in multiple -omics, where AI techniques are largely applied, it should be considered that genetic information carries data relevant to relatives and may be re-identified[26].

Machine-Learning-Discussion-Paper.pdf (last accessed May 10th, 2021); EUROPEAN COMMISSION, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust,* February 2020*,* https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (last accessed May 10th, 2021).

[20] H. FELZMANN ET AL., *Towards Transparency by Design for Artificial Intelligence,* in *Science and Engineering Ethics* vol. 26, no. 6 (December 2020): 3333–61.

[21] J. AMANN ET AL., *Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective,* in *BMC Medical Informatics and Decision Making* vol. 20, no. 1 (30 November 2020): 310.

[22] C. RUDIN, *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,* in *Nature Machine Intelligence* vol. 1, no. 5 (May 2019): 206–15.

[23] J. AMANN ET AL., *Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective,* in *BMC Medical Informatics and Decision Making* vol. 20, no. 1 (30 November 2020): 310.

[24] M. A. GIANFRANCESCO ET AL., *Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data,* in *JAMA Internal Medicine* vol. 178, no. 11 (1 November 2018): 1544–47.

[25] S. GERKE, T. MINSSEN, AND G. COHEN, *Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare,* in *Artificial Intelligence in Healthcare,* 2020, 295–336.

[26] J. S. WINTER AND E. DAVIDSON*, Governance of Artificial Intelligence and Personal Health Information, Digital Policy, Regulation and Governance* 21, no. 3 (1 January 2019): 280–90.

4. The risks of AI in healthcare

AI interventions should be progressively adapted to and incorporated into clinical research, but they require an adaptation of study designs developed for drugs. Robust clinical trials to understand the potential benefits and risks of an algorithm are needed[27]. Moreover, risks that are specific of AI should be considered in the informed consent process in clinical studies involving this intervention. In particular, AI systems may come across specific potential risks such as error, bias, cyberattacks and re-identification of anonymous data.

As it may happen with a navigation system that fails to indicate the right way during a car trip, an AI algorithm may incur in error for various reasons, including the possibility that the AI system has not been properly trained with cases similar to a specific situation. Such a situation may cause severe events if AI is applied in high-risk settings like intensive care units. Although promising, these applications still need to be studied to reach the highest level of confidence in patients with high risk of complications and death. It must be noted that not only AI systems may be not appropriately trained, but also AI is still not able to adequately recognize a causation compared to a correlation, and this could lead to recommendations that do not correspond to the actual patient's background and needs[28].

Ideally, AI systems should be fed with the largest amount of data available. Still, data available for training an algorithm may be biased and may not represent all the individuals to whom AI algorithms will be applied. Bias could derive from a non-heterogeneous training data set[29], missing data and patients not identified by algorithms, small sample size and underestimation, misclassification and measurement errors[30]. Although humans are subject to bias as well, these potential problems should be taken into account when developing algorithms for clinical support and medical decision, and preventive measures should be adopted[31].

Cyberattacks are becoming a real threat for the health sector[32] and their frequency is increasing, especially during critical situations, such as a pandemic[33]. A particularly dangerous threat is represented by "input attacks", consisting in manipulating the data used by the AI system in order to alter the output of the system[34]. For example, a MRI image can be modified so that the AI will detect

---

[27] E.J. TOPOL, *Welcoming New Guidelines for AI Clinical Research*, in *Nature Medicine* vol. 26, no. 9 (September 2020): 1318–20, https://doi.org/10.1038/s41591-020-1042-x.

[28] M. KIENER, *Artificial Intelligence in Medicine and the Disclosure of Risks,* in *AI & Society,* 22 October 2020, 1–9.

[29] R.B. PARIKH, S. TEEPLE, A. S. NAVATHE, *Addressing Bias in Artificial Intelligence in Health Care,* in *JAMA*, 22 November 2019.

[30] M.A. GIANFRANCESCO ET AL., *Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data,* in *JAMA Internal Medicine* vol. 178, no. 11 (1 November 2018): 1544–47.

[31] M.A. GIANFRANCESCO ET AL., *Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data*.

[32] S.T. ARGAW ET AL., *The State of Research on Cyberattacks against Hospitals and Available Best Practice Recommendations: A Scoping Review, BMC Medical Informatics and Decision Making* 19, no. 1 (11 January 2019): 10.

[33] H. S. LALLIE ET AL., *Cyber Security in the Age of COVID-19: A Timeline and Analysis of Cyber-Crime and Cyber-Attacks during the Pandemic, ArXiv E-Prints* (1 June 2020), arXiv:2006.11929, available at https://arxiv.org/abs/2006.11929 (last accessed May 10th, 2021).

[34] M. COMITER, *Attacking Artificial Intelligence. Paper, Harvard Kennedy School, Belfer Center for Science and International Affairs, August* 2019, available at https://www.belfercenter.org/sites/default/files/2019-08/AttackingAI/AttackingAI.pdf (last accessed May 10th, 2021).

a false malignant diagnosis. These attacks do not interfere with the AI algorithms themselves, hence they are difficult to detect. Moreover, they should be considered as "inherent" risks in AI-based medical procedures[35].

Privacy and confidentiality are two of the fundamental principles in managing personal data. Patients have the right of control and decide on their health information and physicians have the duty to guarantee privacy[36]. Even if data are anonymized the risk of re-identification in particular circumstances, i.e. getting the patient's identity by matching and combining different data, is real and has to be taken into account[37].

5. Emerging uses of AI in research studies

AI may significantly improve the efficiency of clinical trials in all phases, from early design, to recruitment, conduction of the study, and reporting activities. AI may help to improve patient selection during recruitment, to reduce heterogeneity and even to select patients with a higher probability to have the outcome under study and more likely to respond to a treatment[38]. It has been suggested that NLP techniques may be used to analyze data from EHRs and social networks to balance the underrepresentation of certain groups such as black people or older adults[39]. Moreover, AI and other digital tools have been indicated as potential solutions for running decentralized or virtual/hybrid clinical trials. These kinds of trials are characterized by a limited in person interaction between trial participants and the investigation site. Digital technologies, including NLP, wearable devices and biosensors, can ideally support the tasks of usual clinical trials with significant resource savings. Decentralized trials, moreover, allow patients to report their outcomes autonomously, which has implications for data quality and for the success of clinical studies. This possibility has raised interest in the use of digital markers that are mostly linked to personal behaviors but that can be collected through simple devices such as a smartphone[40].

Digitized clinical trials should apply appropriate privacy, safety, and regulatory measures. First, the enormous amount of data exchanges require security measures to prevent data breaches. This observation is relevant to cybersecurity, which is a component of medical device certification[41]. Blockchain has been indicated as an helpful technology to prevent breaches of databases containing sensitive information[42]. A blockchain is a digital transaction archive, which guarantees immutability and privacy of transactions through duplication over the entire computer network. Its application is

[35] M. KIENER, *Artificial Intelligence in Medicine and the Disclosure of Risks,* in *AI & Society,* 22 October 2020, 1–9.
[36] P. BALTHAZAR ET AL., *Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics,* in *Journal of the American College of Radiology: JACR* vol. 15, no. 3 Pt B (March 2018): 580–86.
[37] S. CHOUDHURY ET AL., *Big Data, Open Science and the Brain: Lessons Learned from Genomics,* in *Frontiers in Human Neuroscience* vol. 8 (2014): 239.
[38] S. HARRER ET AL., *Artificial Intelligence for Clinical Trial Design,* in *Trends in Pharmacological Sciences* vol. 40, no. 8 (August 2019): 577–91.
[39] O.T. INAN ET AL., *Digitizing Clinical Trials,* in *NPJ Digital Medicine* vol. 3, 101 (2020).
[40] O.T. INAN ET AL., *Digitizing Clinical Trials,* cit.
[41] O.T. INAN ET AL., *Digitizing Clinical Trials*, cit.
[42] P.V. KAKARLAPUDI AND Q. H. MAHMOUD, *A Systematic Review of Blockchain for Consent Management,* in *Healthcare* (Basel, Switzerland) 9, no. 2 (1 February 2021).

particularly interesting because it allows patients to manage their personal data with smart contracts[43]. Blockchain allows different permission levels for different users when accessing personal data. This means that a patient may authorize a medical doctor to access her sensitive data but only a fraction of them may be available for a medical researcher. The same flexibility may be applicable to Internet of Things (IoT) devices where a patient may decide to revoke her decision at any time. Application of blockchain to AI interventions paves the way to future development of research based on data sharing which will require consideration in the ethical dimension and the informed consent process. Indeed, blockchain has implications in data authentication, storage, and privacy, which should be properly reflected in the informed consent process[44].

6. The Covid-19 pandemic

The Covid-19 pandemic has represented a stimulus in the implementation of electronic consent in clinical research. Several existing digital strategies have been reinforced to cope with difficulties in enrollment, including re-consent due to travel restrictions, and other measures that favored the transition from traditional to virtual clinical trials. The FDA, for example, has developed a platform allowing investigators to replace traditional consent with an electronic version[45]. Similarly, the health authorities in Singapore issued recommendations supporting the use of eConsent[46]. Finally, the European Medicine Agency has proposed alternatives to the traditional in person process for clinical trials, when participants are unable to consent in person, including electronic alternatives[47]. On the other hand, the pandemic emergency has limited research activities and development of AI tools because they cannot be trained on large numbers of observations and from homogeneous populations, resulting in bias[48].

Indeed, the Covid-19 pandemic has stimulated the development of a number of tools based on AI to monitor, model, control, mitigate, diagnose and treat the SARS COV2 infection. Among those directly inherent to patients, attempts have been made to develop algorithms for correctly interpreting diagnostic images. The results of these efforts were very good, but with little practical application as diagnostic images may be normal in early phases of infection, and very easy to interpret once pneumonia is established. Simple AI based models have been proposed to develop efficient strategies for viral diagnostic tests. AI has also been used for drug discovery for Covid-19, which led to significant

[43] Y. ZHUANG ET AL., *Applying Blockchain Technology to Enhance Clinical Trial Recruitment,* in *AMIA Annual Symposium Proceedings* 2019 (4 March 2020): 1276–85.

[44] P. V. KAKARLAPUDI AND Q. H. MAHMOUD, *A Systematic Review of Blockchain for Consent Management*.

[45] U.S. FOOD AND DRUG ADMINISTRATION, *COVID MyStudies Application (App),* 29 May 2020, https://www.fda.gov/drugs/science-and-research-drugs/covid-mystudies-application-app (last accessed May 10th, 2021).

[46] SINGAPORE HEALTH SCIENCE AUTHORITY, *Guidance on the Conduct of Clinical Trials in Relation to the COVID-19 Situation,* 29 July 2020, https://www.hsa.gov.sg/docs/default-source/hprg-io-ctb/hsa_ctb_covid-19_guidance_for_clinical_trials_29jul2020.pdf (last accessed May 10th, 2021).

[47] EUROPEAN MEDICINES AGENCY (EMA), *Guidance on the Management of Clinical Trials during the COVID-19 (Coronavirus) Pandemic,* 4 February 2021, https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/guidanceclinicaltrials_covid19_en.pdf (last accessed May 10th, 2021).

[48] M. KIENER, *Artificial Intelligence in Medicine and the Disclosure of Risks,* in *AI & Society,* 22 October 2020, 1–9.

ISSN 2284-4503

hypotheses for candidate therapies[49]. These examples underline how important is to balance the urgency to discover new solutions for an emerging disease with the ethical implications.

7. GDPR and AI in healthcare

The General Data Protection Regulation (GDPR) is currently the strongest data protection regulation in the world[50]. The GDPR aims at preventing unconsented and secondary uses of personal data, but also reduces administrative formalities before accessing and using health data. Among the rights highlighted by the GDPR, "Non-Discrimination Right", the "Right to Explanation", and the "Right to Be Forgotten" are included[51]. A biased algorithm applied to healthcare can result in discriminatory decision-making. As discrimination is forbidden according to EU law, even when it is indirect, there is a need to prevent that as underlined in article 22 of the GDPR[52,53]. For what concerns the "Right to be Forgotten", as explained in the article 17 of the GDPR, a subject has the right to have her personal data eliminated and no longer available and processed. This principle is not easy to apply to AI, because data deletion in AI contexts is complex and algorithms do not forget the way humans do[54]. Therefore, looking for potential solutions to protect this right in a way doable for machine learning environments is one of the challenges of AI in healthcare.

Another rule that is included in GDPR is that de-identification, a technique to remove personal data leading to identification of an individual, is not considered sufficient to prevent re-identification. Therefore, de-identified data are still considered protected personal data, while only anonymous data are excluded from its application. According to GDPR, anonymous data is "information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable". Inadequate anonymization could be applied due to several misunderstandings[55].

---

[49] A. C. Chang, *Artificial Intelligence and COVID-19: Present State and Future Vision,* in *Intelligence-based medicine* 3, 100012. https://doi.org/10.1016/j.ibmed.2020.10001.

[50] European Parliament, *Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 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),* available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN (last accessed May 10th, 2021).

[51] J. S. Winter and E. Davidson, *Governance of Artificial Intelligence and Personal Health Information, Digital Policy, Regulation and Governance* 21, no. 3 (1 January 2019): 280–90.

[52] European Union Agency for Fundamental Rights, *#BigData: Discrimination in Data-Supported Decision Making, Focus Paper,* 28 May 2018, available at https://fra.europa.eu/en/publication/2018/bigdata-discrimination-data-supported-decision-making (last accessed May 10th, 2021).

[53] European Parliament, *Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 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*).

[54] E. F. Villaronga, P. Kieseberg, T. Li, *Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten,* in *Computer Law & Security Review* vol. 34, no. 2 (1 April 2018): 304–13, https://doi.org/10.1016/j.clsr.2017.08.007.

[55] Agencia Española de Protección Datos (AEPD), European Data Protection Supervisor (EDPS), *AEPD-EDPS Joint Paper on 10 Misunderstandings Related to Anonymisation,* 27 April 2021, https://edps.europa.eu/data-protection/our-work/publications/papers/aepd-edps-joint-paper-10-misunderstandings-related_en (last accessed May 10th, 2021).

The GDPR also uses a clear definition of the informed consent that should be "freely given, specific, informed and unambiguous", that it should be explicit and informed and that "silence, pre-ticked boxes or inactivity should not therefore constitute consent". This information must be taken into account when considering the use of electronic consent for studies on AI interventions. Regarding the management of sensitive data, the GDPR states that data processing should be limited to predefined purposes although secondary use of data is allowed for scientific research purposes. As AI algorithms need large amounts of data to be trained, it is likely that this process will rely on the reuse of personal data collected for other purposes, and this may be a limitation in their development[56]. One of the most relevant rights affirmed by the GDPR, finally, is that the individual will not be subject to a decision based solely on automated processing, including profiling which leaves decision making to humans. These principles may challenge the application of AI interventions, which work on a complex intersection of data that may be difficult to justify[57].

8. Shaping the informed consent for AI interventions

The aforementioned observations pose unique challenges in the development of informed consent for AI based interventions. In order to make decisions regarding medical procedures to undergo, patients should be properly informed about risk and benefits and should be acknowledged about different options. The obvious benefit of AI intervention is the potential to be superior to routine clinical practice. The downside include potential cyberattacks, which became more frequent during the Covid-19 pandemic, and error, due to bias in the development and training of algorithms, which is also exacerbated by the Covid-19 pandemic. However, an appropriate communication on the risk/benefit analysis should be based on probabilities that a specific event occurs which, unfortunately, is very difficult to estimate for these risks.

Another challenge in developing an informed consent for AI interventions is explaining how certain tools work. The "black box" problem makes the explanation of the mechanism difficult. Moreover, the majority of medical doctors are not knowledgeable on the basic principles of AI[58]. This observation calls for the need by clinicians to disclose to which extent they can interpret the recommendations provided by AI systems[59].

If explainability may be an issue in adult patients, it becomes even more difficult when patients are in pediatric age and assent should be obtained from them. Many of the existing AI therapeutic interventions (digital therapeutics) are developed for diseases in childhood and will increasingly need to be tested in clinical trials. An effort should be made in improving explainability of AI interventions tailoring the informed consent to different age groups and profiles.

[56] J. Sunrise Winter, E. Davidson, *Governance of Artificial Intelligence and Personal Health Information, Digital Policy, Regulation and Governance* 21, no. 3 (1 January 2019): 280–90.

[57] M. Bourassa Forcier et al., *Integrating Artificial Intelligence into Health Care through Data Access: Can the GDPR Act as a Beacon for Policymakers?,* in *Journal of Law and the Biosciences* vol. 6, no. 1 (2019), 317-335.

[58] D. Schiff and J. Borenstein, *How Should Clinicians Communicate With Patients About the Roles of Artificially Intelligent Team Members?,* in *AMA Journal of Ethics 21,* no. 2 (1 February 2019): E138-145.

[59] S. Gerke, T. Minssen, and G. Cohen, *Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare,* in *Artificial Intelligence in Healthcare,* 2020, 295–336.

While the focus of informed consent is providing the best information to patients in order to make their own decisions, the possibility that a patient withdraws from a study including training algorithms, creates a potential problem. Indeed, patients no longer participating in clinical trials have the right to have all their data deleted. This may create a problem in updating algorithms when they have been already trained and deployed.

It is likely that the use of electronic tools based on AI will be increasingly accompanied by electronic consent processes. On one hand, these tools have the potential to improve the understanding of patients using multimedia and verification tools[60]. On the other hand electronic consent without the supervision of a health professional in person, should be carefully developed to include all details relevant of AI interventions which may be challenging[61].

Blockchain may be an interesting solution for data transactions, not necessarily specific for AI interventions, which secures privacy of users, immutability of transactions, and personalization of data access through smart contracts. As AI systems may result from a combination of data from different sources and there is the potential for reuse of data, blockchain may represent a safety net in clinical studies. Moreover, blockchain is increasingly considered as a technology supporting the informed consent process that allows tracking of the process itself and that may help to provide consent at multiple levels and for different purposes[62].

9. Conclusions

The explosion of AI technologies for healthcare will require an increasing number of clinical trials to compare AI based tools with existing best practices. At present, the existing recommendations and tools for the informed consent process are not perfectly tailored for these studies. A crucial role in satisfying the needs of these studies should be played by Institutional Review Boards and from Ethical Committees. It will be essential that these boards and committees will include participants with both a technical knowledge of AI systems, and a specific ethical background in this area. Such a result not only requires a multidisciplinary approach, but also a new impulse to acquire a proper knowledge in AI by diverse professionals involved in ethical decisions of AI studies. This approach should improve the content and the understanding of informed consent forms for these studies. A second need will be providing precise and thorough information to participants in clinical trials through personal communication, which is usually made by a clinician. The complexity of AI mechanisms and their difficulty to be explained may add to the scarce familiarity of clinicians with AI topics. It is clear that education activities focused on the use of AI in healthcare and on ethical implications are strongly needed for addressing these challenges since they carry completely different characteristics compared with traditional diagnostic and therapeutic interventions.

Due to the need to limit personal contacts, the Covid-19 pandemic has created pressure on the deployment of digital tools for clinical trials that can be managed remotely. This pressure has raised

[60] F. Gesualdo et al., *Digital Tools in the Informed Consent Process: A Systematic Review',* in *BMC Medical Ethics* vol. 22, no. 1 (27 February 2021): 18.

[61] S. Gerke, T. Minssen, and G. Cohen, *Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare.*

[62] P. V. Kakarlapudi and Q. H. Mahmoud, *A Systematic Review of Blockchain for Consent Management,* in *Healthcare* (Basel, Switzerland) 9, no. 2 (1 February 2021).

the interest in virtual and hybrid clinical trials and will possibly accelerate the transition from in person activities to digital trials helping to streamline costs, efforts and time, conducting to a wider patients' engagement and faster results. One interesting implication of remoting clinical trials is the opportunity to collect a wide array of data reported by patients or from sensors, creating rich databases that will require a specific effort for the analysis. On the other hand, the lack of personal encounters in virtual clinical trials may create communication problems between the patients and the investigators.

ISSN 2284-4503