REPORT
ON THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN HEALTHCARE AND ITS IMPACT ON THE "PATIENT–DOCTOR" RELATIONSHIP

Steering Committee for Human Rights
in the fields of Biomedicine
and Health (CDBIO)

COUNCIL OF EUROPE
75
1949 - 2024
CONSEIL DE L'EUROPE

REPORT
ON THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN HEALTHCARE AND ITS IMPACT ON THE "PATIENT–DOCTOR" RELATIONSHIP

Steering Committee for Human Rights
in the fields of Biomedicine
and Health (CDBIO)

September 2024

Council of Europe

Cover design and layout:
Documents and Publications Production Department (SPDP), Council of Europe

Photos: Shutterstock

Printed at the Council of Europe

Contents

I. INTRODUCTION

The Council of Europe aims to protect human dignity and the fundamental rights and freedoms of the individual regarding the application of biology and medicine. Major technological breakthroughs, like those involving artificial intelligence (AI), have the potential to advance biomedicine and benefit healthcare, yet uncertainty exists about their impact and the direction of developments.

The governance of these developments is more than just facilitating their application and containing their risks; it is the way their technological pathways are managed (and sometimes become irreversible). Governance is about embedding human rights in AI technologies that have an application in the field of biomedicine. This implies that developments are from the outset oriented towards protecting human rights. For that reason, governance arrangements need to be considered that seek to steer the innovation process in a way that connects innovation and technologies with social goals and values.

In the framework of its Strategic Action Plan on Human Rights and Technologies in Biomedicine (2020-2025), the Council of Europe Committee on Bioethics[1] set up a drafting group to prepare a report on the application of AI (hereinafter referred to as "AI systems" or "AI-enabled") in healthcare and its impact on the "patient–doctor" relationship, highlighting the role of healthcare professionals in respecting the autonomy, and right to information, of the patient, and in maintaining transparency and patient trust as critical components of the therapeutic relationship.

The drafting group comprised the following members: Dunja Pejović (Bosnia and Herzegovina); Emmanuel Didier (France); Joni Komulainen as Chair (Finland); Sabine Salloch (Germany); Evaristo Cisbani (Italy); Patricio Santillan-Doherty (Mexico); and Andreas Reis (World Health Organization).

The drafting group met on seven occasions from October 2022 to March 2024, for two in-person meetings and five online meetings. It held exchanges with experts from the Netherlands and France (Paris, 8-9 February 2022), took into account the views and proposals of young people who participated in the CDBIO pilot youth forum (Strasbourg, 6 June 2023), and integrated the feedback from a targeted consultation, held between December 2023 and February 2024, into the report.

1. Since replaced by the Steering Committee for Human Rights in the fields of Biomedicine and Health (CDBIO).

Purpose

The report is intended for decision makers, health providers, health professionals and patients (including patient associations), to:

- consider how AI systems are used in healthcare, having regard to their human rights implications;
- develop and strengthen the therapeutic relationship, especially in supporting doctors and, where appropriate, other healthcare professionals in promoting the agency and autonomy of patients, patient welfare and equitable access to healthcare.

Scope

The report focuses on selected human rights principles of particular relevance to the therapeutic relationship, namely consent (Article 5 of the European Convention on Human Rights and Biomedicine, or the Oviedo Convention), professional standards (Article 4, Oviedo Convention), private life and right to information (Article 10, Oviedo Convention) and equitable access to healthcare (Article 3, Oviedo Convention).

The report addresses AI[2,3] in healthcare, including applications that are used by healthcare professionals as well as applications that are used by the patients themselves (apps prescribed by a doctor, but also independently used apps such as symptom checkers or health data trackers). With its focus on the patient–doctor relationship, the report does not focus on AI development, nor on AI-related research that includes human subjects.

Understanding the patient–doctor relationship with regard to AI systems

The therapeutic relationship is a critical component of good patient care, which AI systems have the potential to improve or affect adversely. It must be acknowledged that AI systems are already becoming an essential tool for modern medicine. This requires attention to the design, development and application of AI systems used in health so that the "interests and welfare of the human being shall prevail over the sole interest of society or science".[4] There should be synergy in the progress and protections advanced by AI systems.

2. The Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law defines an "AI system" as a "machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment."
3. The European Union (EU) AI Act defines an AI system 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".
4. European Convention on Human Rights and Biomedicine of 1997, otherwise known as the "Oviedo Convention", Article 2.

Most importantly, the patient–doctor relationship is one of trust that, in turn, is based on the trustworthiness of the standards and ethics of healthcare professionals to help and support people facing illness and disease. This relationship is inherently human, taking different forms (noting that in contrast AI systems do not ask questions, so they will not engage in conversations unless prompted). It is a very special and historically valued type of relationship that has been evolving constantly from an outdated traditional paternalistic model to a more desirable deliberative model. To assert the character of this relationship, one that puts patients first, the drafting group decided to invert the reference from "doctor–patient" to "patient–doctor" for the entirety of this report.

AI systems have the potential to transform the "patient–doctor" relationship. The expertise of doctors could be challenged, but also significantly increased by high-performing decision support in various domains of healthcare. Patients who decide to use AI systems by themselves, however, might rely less on the advice of health professionals. Challenges lie in misplaced trust, overestimation of technological performance, and testimonial (in)justice as to the question of whom to trust in the patient–doctor encounter.

The patient–doctor relationship will need to be safeguarded as AI systems are integrated into healthcare. Doctors and patients should be obligated to work together to assert the importance of shared decision making and to seek information and explanations (from healthcare providers, providers of AI systems, technical collaborators) about how and why AI systems result in certain outputs so that professional standards and patient autonomy are maintained. To this end, there should be emphasis on information being provided about the logic, scope and consequences of recommendations made by AI systems. Investment in automation will necessitate commensurate levels of support and training, public dialogues and other efforts to render this change as transparent and understandable as possible. Openness about what is known about AI systems should be foregrounded, especially about the place, purpose and functioning of AI systems in health settings. Communication campaigns will help to foster appropriate social attitudes and behaviours, for example by explaining how patients will use AI systems (chatbots) when accessing and using healthcare to optimise the allocation of health resources. A further example would be to explain to patients why their health data are needed to train AI systems (and therein be shared across platforms) to help allay concerns about privacy and medical confidentiality.

What is AI and why is it important in the patient–doctor relationship?

For the purposes of this report, AI systems are considered to simultaneously integrate at least three concurrent components: data, computational hardware and computational software (algorithms).

- **Data** can be collected by a range of heterogeneous protocols and devices, such as wearable personal sensors, medical exams and clinical equipment. Data are exploited by AI systems for training, testing, validation, and potentially continuous updates and performance verification. Data should be

representative to avoid collection bias, diversified and heterogeneous; the accuracy and precision of the data should be known and considered by the AI system in question. Quality assurance and standardisation are key aspects for the proper exploitation of data by AI system algorithms. Most data collected and processed by AI system devices are personal data of a sensitive nature that are subject to specific regulations (Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, CETS No. 108; General Data Protection Regulation (GDPR)) and require anonymisation and/or pseudonymisation (or other safeguards such as data aggregation or the generation of synthetic data) to avoid undue traceability and identification of the patient. It is noteworthy that the protection of patient privacy may affect the availability of statistically diversified, heterogeneous and adequate data, and it is often considered an obstacle to data exploitation and sharing and therefore to the development of AI system applications.[5]

- **Computational hardware** is needed to store the large volumes of data required for AI systems. Algorithms run on computational nodes, the performance of which depend on tasks to be accomplished (for example train a model, use trained model). Among different types of computational hardware, hybrid cloud computing (where different storage and computing environments co-exist, such as in-house and remote) currently represents one of the most effective compromises between economic affordability, scalability and shared exploitation among clinical centres.[6,7] Security, privacy and compliance represent the most challenging aspects for the deployment of an information technology infrastructure in healthcare.
- **Computational algorithms** are able to perform tasks that are generally associated with human intelligence. The capability of AI systems to mimic (part) of human intelligence is, on the one hand, the main challenge for these technologies. On the other hand, it represents the main concern for their application. Deep neural networks are among the most discussed and expanding subclasses of AI systems (within machine learning algorithms) due to their intrinsic "black box" structure and possibilities to learn during computations; they can be trained on (large) datasets to calibrate millions (or billions) of internal parameters and then be applied to new input data to identify new correlations and findings.[8]

5. Italian Committee for Bioethics and Italian Committee for Biosafety, Biotechnology and Sciences of Life, "Artificial intelligence and medicine: ethical aspects", May 2020, Presidency of the Council of Ministers, available at https://bioetica.governo.it/en/opinions/joint-opinions-icbicbbsl/artificial-intelligence-and-medicine-some-ethical-aspects, accessed 18 October 2024.
6. Pritpal A. et al. (2019), "Technical health check for cloud service providers", *ArXiv*.
7. For instance, to train an AI system using 1 million genomes, even as synthetic data, would require 8 petabytes of storage and processing capabilities.
8. Image recognition technologies, for example, can decide what types of objects appear in a picture. The algorithm "learns" by defining rules to determine how new inputs will be classified. The model can be taught to the algorithm via hand-labelled inputs (supervised learning); in other cases, the algorithm itself defines best-fit models to make sense of a set of inputs (unsupervised learning). In both cases, the algorithm defines decision-making rules to handle new inputs. Critically, a human user will typically not be able to understand the rationale of decision-making rules produced by the algorithm. See Qu'est-ce que l'apprentissage automatique ?, Trend Micro, available at www.trendmicro.com/fr_fr/what-is/machine-learning.html, accessed 18 October 2024.

AI systems are a significant driver for progress in healthcare. They are being used throughout healthcare, from diagnostics through to prediction, prevention, therapy (including triage) and rehabilitation. In Europe, there is increasing reliance on AI systems used in healthcare. In the United States of America, the Food and Drug Administration (FDA) has cleared or approved more than 950 AI algorithms for medical use.[9,10,11] The vast majority of AI systems are being developed in the field of medical imaging, but other areas are evolving.

- In practice, AI systems must be reliable, convenient to use and easy to integrate into health workflows. This may not be the case where AI systems require expensive infrastructure investment, computing and storage capabilities. Such constraints raise the question as to whether viable cost-efficient alternatives to AI systems exist. Early health technology assessments of AI systems are a way forward (for example as undertaken in the Netherlands in the field of multiple sclerosis).[12]

AI systems have the potential to, *inter alia*, support doctors with diagnostics and health personnel with administrative workflows. Yet AI systems risk potential disruption to care responsibilities. For example, radiologists competing with increasingly accurate AI system outputs could be displaced or otherwise substituted by other health professionals and/or feel compelled to use AI systems as a supporting tool, perhaps to gain a competitive advantage, notwithstanding the possibility of error. Doctors will need guidance in dealing with incidental findings that AI systems produce, especially on whether and how to communicate these to patients.[13,14]

AI trends and examples that have a bearing on the patient–doctor relationship

AI systems have the potential to bring about a significant transformation of the patient–doctor relationship, although the effects of its deployment are yet to be seen. The following examples illustrate the breadth of development of AI systems used in healthcare.

9. Heindl A. (2024), "The step-by-step guide to getting your AI models through FDA approval", *Encord*, available at https://encord.com/blog/ai-algorithm-fda-approval, accessed 18 October 2024.
10. Muehlematter U. et al. (2021), "Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): a comparative analysis", *Health Policy* Vol 3(3), E195-E203, available at www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30292-2/fulltext, accessed 18 October 2024.
11. Artificial intelligence and machine learning (ai/ml) – enabled medical devices, FDA, available at www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices, accessed 18 October 2024.
12. Wester V., Huygens S. and Versteegh, M. (2021), "Early HTA on the value of an AI-based decision support system in multiple sclerosis".
13. A recent study shows that chest X-ray AI systems can predict age, self-reported gender, self-reported ethnicity and insurance status: Adleberg J. (2022), "Predicting patient demographics from chest radiographs with deep learning", *Journal of the American College of Radiology* Vol. 19(10), pp. 1151-61.
14. Feathers T., Palmer K. and Fondrie-Teitler S. (2022), "Out of control": dozens of telehealth startups sent sensitive health information to big tech companies, The Markup, available at https://themarkup.org/pixel-hunt/2022/12/13/out-of-control-dozens-of-telehealth-startups-sent-sensitive-health-information-to-big-tech-companies, accessed 18 October 2024.

AI imaging systems use machine learning algorithms to help improve image quality (and/or for producing the same images of similar quality with less radiation dose to the patient). This is not diagnostics per se but does contribute to later (human) decisions that can identify patterns and anomalies used, for example, to detect tumours, diagnose heart conditions and identify other abnormalities.

- AI systems are being evaluated for use in radiological diagnosis in oncology (thoracic imaging, abdominal and pelvic imaging, mammography, brain imaging and dose optimisation for radiological treatment),[15] in non-radiological applications (dermatology), in diagnosis of diabetic retinopathy, and for RNA and DNA sequencing to guide immunotherapy.[16]

AI-enabled diagnostic systems use machine learning algorithms to help medical professionals identify symptoms, compare them to medical records and recommend treatments based on the data. Used in hospitals and clinics to help diagnose and treat a variety of diseases and conditions, AI systems can help to identify disease progression (assisting doctors in diagnosing with confidence by filling in gaps and helping them arrive at key insights).

- Developmental dysplasia of the hip has a significantly better prognosis when caught early. However, there are often few to no symptoms during the early stages.[17]

AI-enabled clinical decision support systems use machine learning algorithms to assist medical professionals in making decisions related to patient care. These systems are being used to provide personalised health recommendations, flag potential drug interactions and alert medical professionals to potential errors. As AI systems evolve, social determinants and lifestyle choices (social media data) may also be incorporated into AI systems. This means an AI system that evaluates an individual's combined genetic and behavioural/social data could greatly improve a doctor's ability to choose the best treatment path/medication.

- Mathematical representations of patients by AI systems from deep analysis of electronic health and medical records (EHR and EMR) are becoming a relevant field of application for deep neural network algorithms.[18] AI systems trained on those data may support clinical professionals in personalised patient risk stratification, diagnosis, prognosis and triage.[19]

---

15. Farina E. et al. (2022), "An overview of artificial intelligence in oncology", *Future Science OA* Vol. 8(4), available at www.ncbi.nlm.nih.gov/pmc/articles/PMC8965797, accessed 18 October 2024.
16. Yang Y. et al. (2022), "Artificial intelligence for prediction of response to cancer immunotherapy", *Seminars in Cancer Biology* Vol. 87, pp. 137-47, available at www.sciencedirect.com/science/article/abs/pii/S1044579X22002309?via%3Dihub, accessed 18 October 2024.
17. Tabata R. C. and Forbes Technology Council (2022), "How AI could predict medical conditions and revive the healthcare system", *Forbes*, available at www.forbes.com/sites/forbestechcouncil/2022/01/25/how-ai-could-predict-medical-conditions-and-revive-the-healthcare-system/?sh=11ae69a46c47, accessed 18 October 2024.
18. Yuqi S. et al. (2021), "Deep representation learning of patient data from Electronic Health Records (EHR): a systematic review", *Journal of Biomedical Informatics* Vol. 115, 103671.
19. Hao K. (2020), Doctors are using AI to triage covid-19 patients. The tools may be here to stay, *MIT Technology Review*, available at www.technologyreview.com/2020/04/23/1000410/ai-triage-covid-19-patients-health-care, accessed 18 October 2024.

Surgery using AI-enabled tools can support health professionals in making health interventions both safer and more effective. For instance, brain surgery is precise and painstaking, and it is vital to avoid damaging critical structures.[20]

- In treatment support to clinicians,[21] AI systems have the potential to assist surgeons in removing malignant tumour tissue more effectively, fusing biopsies with full-scale interventions.[22]

Virtual nursing assistants use AI systems to automate tasks such as monitoring vital signs, providing medication reminders and managing patient records. These systems are designed to reduce the workload of nurses and other medical personnel while providing more accurate results.[23]

- Clinical decision support systems can assist clinicians in everyday problem solving for systemic inflammatory response syndrome, sepsis and associated organ dysfunctions in paediatric intensive care, as they summarise, analyse and present clinically relevant data at the point of care.[24]

AI-enabled wearable medical devices use machine learning algorithms to track and monitor health data. These devices are being used to monitor heart rate, blood pressure and other vital signs, as well as to provide actionable insights into an individual's health, including patients.

- In the detection of disease, AI-enabled medical devices have given the electrocardiogram (ECG) and clinicians reading them diagnostic abilities. This transforms the ECG, a ubiquitous, non-invasive cardiac test that is integrated into practice workflows, into a screening tool and predictor of cardiac and non-cardiac diseases, often in asymptomatic individuals.[25]

AI-enabled telemedicine platforms use machine learning algorithms to provide remote[26] medical care (in line with the Sustainable Development Goals of the United

---

20. Mazumdar T. (2023), Safer brain surgery using AI possible within two years, BBC, available at www.bbc.com/news/health-66921926?at_medium=RSS&at_campaign=KARANGA, accessed 18 October 2024.
21. Other examples include European start-ups such as Cerenion, a company that introduced a device that monitors the brain function of intensive care patients, and Omnidermal, which uses AI algorithms in dermatology. Global tech companies such as Apple are developing a device that detects tremors afflicting sufferers of Parkinson's disease, and Samsung has introduced AI breast screening solutions.
22. Project Classica uses AI-based algorithms to differentiate between cancerous and non-cancerous tissues in real time. This helps to clinically validate a novel AI-guided intraoperative decision support technology in the surgical care of cancer patients.
23. Jesus A. de (2019), "Machine learning for nursing – 8 current applications", Emerj, available at https://emerj.com/ai-sector-overviews/machine-learning-for-nursing-8-current-applications, accessed 18 October 2024.
24. Böhnke J. et al. (2022), "Prediction models for SIRS, sepsis and associated organ dysfunctions in paediatric intensive care: study protocol for a diagnostic test accuracy study", *BMJ Paediatrics Open* Vol. 6(1), available at https://bmjpaedsopen.bmj.com/content/6/1/e001618, accessed 18 October 2024.
25. Zachi I. et al. (2021), "Application of artificial intelligence to the electrocardiogram", *European Heart Journal* Vol. 42(46), pp. 4717-30, available at https://pubmed.ncbi.nlm.nih.gov/34534279, accessed 18 October 2024.
26. Ahmed A. et al. (2023), "The effectiveness of wearable devices using artificial intelligence for blood glucose level forecasting or prediction: systematic review", *Journal of Medical Internet Research* Vol. 25, e40259, available at www.ncbi.nlm.nih.gov/pmc/articles/PMC10131991, accessed 18 October 2024.

Nations[27] concerning equal accessibility to treatment and rehabilitation also in rural/remote areas). These systems are being used to provide virtual consultations, provide medication reminders and adherence tracking, and monitor patient vital signs. Eventually, AI systems will likely assist patients in self-managing their medical conditions, especially chronic diseases such as cardiovascular diseases, diabetes and mental problems. AI systems already assist in self-care, including through conversation agents (chatbots), health monitoring and risk prediction tools and technologies designed specifically for individuals with different problems[28] and disorders.[29]

- Patients using AI-enabled apps are assisting in real-time monitoring of biometrics. For instance, AI-enabled wearables for people with diabetes help to monitor and maintain glucose levels via an automated insulin delivery system worn on the body. Using a self-learning algorithm, such medical devices embed AI-enabled treatment mechanisms to assist people in managing their daily insulin levels.

AI-enabled drug discovery systems use machine learning algorithms to more rapidly identify (novel) potential treatments for diseases and conditions. These systems are helping to expedite the drug discovery process and can be used to identify potential treatments for conditions that were previously thought to be untreatable.

- The development of AI systems to predict the three-dimensional shape of proteins, such as RoseTTAfold and AlphaFold,[30] are helping to speed up the development of new medicines and to improve the repurposing of existing medicines for use against new viruses and diseases.

27. The 17 Goals, United Nations, available at https://sdgs.un.org/goals, accessed 18 October 2024.
28. Milne-Ives M. et al. (2022), "Artificial intelligence and machine learning in mobile apps for mental health: a scoping review", *PLOS Digit Health* Vol. 1(8), e0000079.
29. Kiluk B. D. et al. (2019), "Technology-delivered cognitive-behavioral interventions for alcohol use: a meta-analysis", *Alcohol, Clinical and Experimental* Research Vol. 43(11), pp. 2285-95.
30. Editorial (2022), "Method of the year 2021: protein structure prediction", *Nature Methods* Vol. 19(1).

II. HUMAN RIGHTS IMPLICATIONS OF AI IN THE PATIENT–DOCTOR RELATIONSHIP

Autonomy

Context

According to Article 5 of the Oviedo Convention, "[A]n intervention in the health field may only be carried out after the person concerned has given free and informed consent to it. This person shall beforehand be given appropriate information as to the purpose and nature of the intervention as well as on its consequences and risks. The person concerned may freely withdraw consent at any time."

Consent empowers a patient's decision process by contrasting their personal preferences with a proposed medical intervention. Paragraphs 35 and 36 of the Explanatory Report to the Oviedo Convention provide further details on the specific requirements for consent, which include requirements concerning the quality, breadth and clarity of information provided.

> 35. The patient's consent is considered to be free and informed if it is given on the basis of objective information from the responsible health care professional as to the nature and the potential consequences of the planned intervention or of its alternatives, in the absence of any pressure from anyone. … In order for their

consent to be valid the persons in question must have been informed about the relevant facts regarding the intervention being contemplated. This information must include the purpose, nature and consequences of the intervention and the risks involved. Information on the risks involved in the intervention or in alternative courses of action must cover not only the risks inherent in the type of intervention contemplated, but also any risks related to the individual characteristics of each patient, such as age or the existence of other pathologies. Requests for additional information made by patients must be adequately answered.

36. Moreover, this information must be sufficiently clear and suitably worded for the person who is to undergo the intervention. The patient must be put in a position, through the use of terms he or she can understand, to weigh up the necessity or usefulness of the aim and methods of the intervention against its risks and the discomfort or pain it will cause."

Yet autonomy is more than consent. It engenders a more active role for patients in shared decision making, one that is not restricted to being informed and agreeing to options presented to them. It encompasses, for example, the choice to take preventive measures, to ask for a second opinion and to exercise the "right not to know" as provided in paragraph 67 of the Explanatory Report to the Oviedo Convention. Most importantly, autonomy is the ability for patients to introduce their own values, preferences and perspectives in patient–doctor communications.

"67. The right to know goes hand in hand with the "right not to know". Patients may have their own reasons for not wishing to know about certain aspects of their health. A wish of this kind must be observed. The patient's exercise of the right not to know this or that fact concerning his health is not regarded as an impediment to the validity of his consent to an intervention; for example, he can validly consent to the removal of a cyst despite not wishing to know its nature."

In research settings, autonomy to consent to diagnostic and treatment interventions (as characterised above) differs from consent by the patient (or health proband) to become a research subject. Everyone involved needs to consider the different "logics" of treatment directed to the individual's needs and research that primarily aims at producing knowledge that serves the interests of future patients. Not least as there is not always the potential for study participants to benefit from their participation, the ethical and legal requirements on informed consent are usually higher than in the therapeutic setting. Notably, in data-intensive research, for example in the development of clinical decision support systems, the borderline between clinical care and medical research becomes increasingly blurred, generating a need for new legal and ethical frameworks for ensuring informed consent (dynamic consent

models are commonly used in such circumstances).[31] In addition, there is more and more online research for which obtaining appropriate informed consent poses additional difficulties.[32,33]

Challenges

Patients may find it difficult to understand what AI systems are, why they are being relied on, and how they are being used in the realisation of their care. While some might argue that doctors have discretion to decide whether to inform their patients about their reliance on AI systems,[34] this should not preclude patients from seeking out information and explanations to consent to health interventions.

As AI systems merge into medical practice, the patient–doctor relationship could even be overthrown or increasingly frustrated as patients are rendered unable to withhold consent to the use of AI systems in their treatment or realisation of their care when other options that do not rely upon them are not easily available, or if the clinician, who has handed over responsibility for such functions to an AI system, is unable to provide care without the use of an AI system.

Recommended action

Patient autonomy necessitates more information, explanation and transparency rather than less. This includes patients knowing when they are interacting with an AI system, and knowing how to consent, especially in cases when the deployment of AI systems leads to healthcare administered with less and/or without recourse to the therapeutic hand of the doctor. Action should be taken to determine when and how to support and empower patients in consenting to interventions recommended by health professionals, supported by AI systems. Clinical ethics support services could play an increasing role in providing guidance and orientation.

When the risks to the patient are high, there should be the possibility to distance human consent from AI system outputs. Patients should be given the opportunity to seek a second opinion without reliance on an AI system and/or oppose treatment decisions and care that depend on (or go beyond mere support of) AI system outputs, should they prefer.

Efforts should also be made to encourage patients to become more active in and critical of decisions regarding their health. Patient organisations can play a useful role in sharing knowledge and good practice on health literacy, which should include AI literacy.

---

31. See guidance from the Central Ethics Commission at the Federal Chamber of Physicians in Germany.
32. Kleinsman J. and Buckley S. (2015), "Facebook study: a little bit unethical but worth it?", *Journal of Bioethical Inquiry* Vol. 12, pp. 179-82.
33. It is noteworthy that increasingly, national or Europe-wide regulations referring to the processing of health data and training of algorithms are based on legislation, not consent.
34. See de Miguel Beriain I. (2020), "Should we have a right to refuse diagnostics and treatment planning by artificial intelligence?", *Medicine, Health Care and Philosophy* Vol. 23(2), pp. 247-52, and Ploug T. and Holm S. (2020) "The right to refuse diagnostics and treatment planning by artificial intelligence", *Medicine, Health Care and Philosophy* Vol. 23(1), pp. 107-14.

Summary

Patients need:

- to be more aware and critical of AI systems used in their healthcare;
- to know when they are interacting with an AI system;
- to be clearly informed about the use of AI in cases that matter to them;
- to be able to accept or refuse AI in their care or treatment, and to this end seek a second opinion without the use of an AI system, where possible.

Health professionals and/or healthcare providers[35] have the responsibility to:

- provide guidance and training on what and how to inform patients about the use of AI systems in their treatment and care,[36] and who can be supported, when needed, by more specific technical collaborators;
- inform patients in clear and simple language why and how AI systems are being used (benefits and risks);
- review procedures to facilitate informed consent when AI systems are relied upon;
- promote health literacy, including AI literacy, and facilitate public dialogues where appropriate.

35. "Healthcare providers" should be interpreted broadly to comprise various actors and assistants in the care process.
36. The European Data Protection Board and the European Data Protection Supervisor have released useful guidelines on transparency (informing) and consent when informing patients about AI, and how to obtain their consent.

Professional standards

Context

The patient–doctor relationship is founded on several prerequisites in healthcare (safety, efficacy, quality) encapsulated in key policy and legal documents, such as the Declaration of Geneva and the Oviedo Convention. Article 4 of the latter states: "Any intervention in the health field, including research, must be carried out in accordance with relevant professional obligations and standards."

Professional obligations, whatever form they take (codes of conduct, legal obligations, etc.), are necessary to ensure standards and quality of care. They create a commitment for doctors over time. They are a bulwark of protection for patients because they compel doctors to pay careful attention to the needs of each patient, as underlined in paragraphs 32 and 33 of the Explanatory Report to the Oviedo Convention.

> "32. It is the essential task of the doctor not only to heal patients but also to take the proper steps to promote health and relieve pain, taking into account the psychological well-being of the patient. Competence must be determined primarily in relation to the scientific knowledge and clinical experience appropriate to a profession or speciality at a given time. The current state of the art determines the professional standard and skill to be expected of health care professionals in the performance of their work. In following the progress of medicine, it changes with new developments and eliminates methods which do not reflect the state of the art.
>
> 33. Further, a particular course of action must be judged in the light of the specific health problem raised by a given patient. … Another important factor in the success of medical treatment is the patient's confidence in his or her doctor. This confidence also determines the duties of the doctor towards the patient. An important element of these duties is the respect of the rights of the patient. The latter creates and increases mutual trust. The therapeutic alliance will be strengthened if the rights of the patient are fully respected."

These paragraphs reinforce the duty of doctors to take care of patients, and to determine the proper treatment within a joint decision-making framework with the patient. Action taken is based on competence, scientific knowledge and clinical experience, which includes the assessment of risk (of AI systems).[37]

- In the EU, it is noteworthy that any AI systems intended to be used for a medical purpose shall comply with the Regulation (EU) 2017/745 on medical devices,[38] which determines different levels of risk-based compliance, taking into account the vulnerability of the patient and the risks associated with the devices.[39] The EU regulation (and similarly, the FDA) classify medical devices according to their intended use and risk into three main classes: I (lowest risk), IIa[40] and IIb,[41] and III[42] (highest risk).

In assessing such risks, further thinking is needed about the attribution of responsibilities for AI systems. AI system developers are responsible for ensuring they are designed in a responsible and ethical manner. Health professionals and/or healthcare providers have the responsibility to use them in a way that aligns with ethical and legal guidelines, and in accordance with the intended uses and instructions provided by manufacturers of AI systems (including what oversight and control by humans is possible to ensure fairness and accuracy of results).[43]

The adoption of a new technology can result, on the one hand, in a loss of skills (deskilling), which should not happen too fast. On the other hand, it requires the acquisition of new skills (upskilling). What a health professional could lose in skills may still be valuable to medical practice (noting that medical practice that largely depends on a technology may be more vulnerable). To this end, specific and continuous professional training on a new technology, such as AI, needs to be carefully evaluated.[44]

---

37. For example, the EU AI Act and the Council of Framework Convention on AI, human rights, democracy and the rule of law.
38. EU Regulation of medical devices: "For the purposes of this Regulation, the following definitions apply: (1) 'medical device' means any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes: diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease, diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability," available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02017R0745-20200424&from=EN, accessed 20 October 2024.
39. Directorate-General for Health and Food Safety (2021), "MDCG 2021-24 - Guidance on classification of medical devices", available at https://health.ec.europa.eu/latest-updates/mdcg-2021-24-guidance-classification-medical-devices-2021-10-04_en, accessed 20 October 2024.
40. For example catheters, hearing aids, short-term contact lenses.
41. For example incubators, insulin pens, long-term contact lenses and ventilators.
42. For example pacemakers, prosthetic heart valves, surgical mesh, breast implants and other devices that require permanent monitoring throughout their lifetimes.
43. Comité Permanent de Médecins Européens (2021), "CPME feedback on Commission proposal for a regulation on artificial intelligence".
44. Cabitza F. et al. (2017), "Unintended consequences of machine learning in medicine", *Journal of the American Medical Association* Vol. 318(6), pp. 517-18.

Challenges

A major challenge lies in ensuring that AI systems (that is, their data and models) are empirically sound, robust and accurate, and that their results are consistent and reproducible (for example based on independent standards or expertise, such as independently tested algorithms). Standards for clinical trials involving AI systems are necessary to ensure the safety, quality and reliability of trial results. This will allow AI systems to be more easily appraised by investigators and others, such as regulators,[45] forming part of a chain of responsibility for AI systems (especially as they self-improve).

Care and standards should not waver when AI systems are introduced. The challenge lies in understanding how these standards are applied and should adapt to safeguard the therapeutic relationship. This includes addressing the benefits and risks of AI systems considering their opacity ("black box"[46] algorithms) and other shortcomings in transparency and reproducibility.

Professional standards should strive to protect and enable doctors and other healthcare professionals to use AI systems with discretion in the best interests of patients, to understand the essential elements of AI systems, and to help explain and support patients. Notwithstanding the potential for AI systems to be effective supporting tools (that is, a form of augmented intelligence for health professionals),[47] the critical thinking and expertise of health professionals should not be underestimated.[48] This should include appropriate professional oversight of the clinical validation of AI systems[49] to ensure that AI in healthcare is subject to high standards and empirically evaluated evidence, as with any other digital device.

There are also challenges in assessing the risks of AI systems. By integrating ethics and responsibility into AI use and, in turn, into the strategic implementation and organisational planning processes of AI systems, risks can be reduced, and trustworthiness maintained. A responsible approach to AI should place humans (that is, patients) at the centre and should align with stakeholder expectations as well as applicable regulations and laws.[50]

---

45. AI algorithms can undergo updates or changes so it will be necessary to determine which version of the AI system was deployed in the clinical trial in question.
46. In computing, a "black box" is a device, system or program that allows you to see the input and output but gives no view of the processes and workings in between.
47. Comité Permanent de Médecins Européens (2019), "CPME policy on AI in health care", CPME/AD/Board/16112019/062_Final/EN, p. 2; American Medical Association (2018), "Augmented intelligence in health care", content derived from Augmented Intelligence (AI) in Health Care (Annual Meeting 2018); World Medical Association (2019), *WMA statement on augmented intelligence in medical care*, adopted by the 70th WMA General Assembly, Tbilisi, Georgia, October 2019, available at www.wma.net/policies-post/wma-statement-on-augmented-intelligence-in-medical-care, all accessed 20 October 2024.
48. For more on the Babylon Triage app, see https://youtu.be/FQm-wnUJNrU?t=74, accessed 20 October 2024.
49. Comité Permanent de Médecins Européens, "CPME policy on AI in health care", p. 6.
50. Sivarajah U. et al. (2023), "Responsible artificial intelligence (AI) for digital health and medical analytics", *Information Systems Frontiers*, available at www.ncbi.nlm.nih.gov/pmc/articles/PMC10240104, accessed 20 October 2024.

Recommended action

It is necessary to establish clear guidelines and regulations for the development and use of AI systems. This includes ensuring that developers are held accountable for errors that the systems may have, and that users are educated on the potential risks and ethical considerations when using AI systems.

AI systems in healthcare should be governed through "human meaningful control"[51] and the assertion of human values with recourse to verified information, evidence, approval and guidance by regulatory bodies, based on transparent information and adequate quality controls (that is, validating, certifying and regulating medical devices). This includes ways to question AI system outputs (for example which AI system is being used and what its intended use is). Furthermore, healthcare professionals who use AI systems, in accordance with the training provided and adherence to relevant instructions and guidelines, could be indemnified against adverse outcomes. New rules are needed to address liability for self-learning algorithms and to clearly identify who is responsible for what. There should be clarity as to whom a healthcare professional should address in case of a defective product, wrong diagnosis or wrong treatment caused by AI systems.[52]

The development of AI systems at all stages, from problem selection to deployment, should consider potential ethical implications. This requires multidisciplinary teams of experts (sociologists, psychologists, philosophers) and representative groups (patients, clinicians) who can detect and mitigate potential problems in AI system applications (for example by developing user interfaces that encourage critical thinking and assessment).[53]

- In the Netherlands, guidelines for high-quality diagnostic and prognostic applications of AI systems in healthcare[54] are an example of good professional conduct in the development, testing and implementation of AI systems.

Health professionals should know about the risks of AI systems (as assessed by providers of AI systems), for example the sources of possible bias when interpreting the plausibility of AI system outputs. Guidance, training and capacity building are needed to support them in addressing whether, what and how AI system information and outputs are communicated to patients. This comprises the standardisation of training and the certification of "high risk" AI systems, including, where appropriate, collaboration with AI developers to foster understanding, confidence and transition towards AI-enabled healthcare. In doing so, skills are developed, anxiety about

51. European Commission: Directorate-General for Research and Innovation and European Group on Ethics in Science and New Technologies (2018), "Statement on AI, robotics and autonomous systems", Publications Office of the EU (europa.eu), available at https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en, accessed 20 October 2024.
52. Comité Permanent de Médecins Européens (2021), "CPME feedback on Commission proposal for a regulation on artificial intelligence", p. 3.
53. Kostick-Quenet K. M. and Gerke S. (2022), "AI in the hands of imperfect users", *npj Digital Medicine* Vol. 5(197); available at www.nature.com/articles/s41746-022-00737-z, accessed 20 October 2024.
54. Smeden M. van (2022), "Guideline for high-quality diagnostic and prognostic applications of AI in healthcare", available at https://www.researchgate.net/publication/365410447_Guideline_for_high-quality_diagnostic_and_prognostic_applications_of_AI_in_healthcare, accessed 20 October 2024.

"deskilling" is reduced (that is, the concern that tasks become largely automated, in this way reducing a physician's ability to autonomously interpret data when AI fails)[55] and, importantly, an equilibrium can be struck between human care and AI-enabled support systems.

Summary

Patients need:

- support from doctors and other health professionals in understanding and/or being involved in the decision-making process when AI systems make important (probabilistic) determinations about their health.

Health professionals and/or healthcare providers have the responsibility to:

- support and empower health professionals in the transition towards AI-enabled healthcare, reinforcing their place, purpose and accountability vis-à-vis AI systems. This includes:
  - maintaining and adapting professional standards and training that help to understand and overcome shortcomings in the results and operations of AI systems;
  - providing guidance regarding the "essential elements" of AI systems;
  - encouraging discretion to distance themselves and/or otherwise oppose AI system outputs should there be a sufficient degree of uncertainty (for example "false positives" in AI-assisted radiology, AI-driven incidental findings in genomics);
- foster multidisciplinary collaboration with AI developers towards the establishment of shared responsibility (and liability) for AI-enabled healthcare;
- promote minimum standards of information and explainability for AI systems (regarding the acquisition, authorisation, practical use and iterative learnings of AI systems).

55. Cabitza F. et al. (2017), "Unintended consequences of machine learning in medicine"; Hoff T. (2011), "Deskilling and adaptation among primary care physicians using two work innovations", *Health Care Management Review* Vol. 36(4), pp. 338-48.

Self-determination regarding health data

Context

Article 10 of the Oviedo Convention states that everyone has the right to respect for private life in relation to information about their health. Furthermore, everyone is entitled to know any information collected about their health, which is likely to include its collection and processing by AI systems. This is underlined in paragraphs 66 and 70 of the Explanatory Report to the Oviedo Convention.

> "66. A person's "right to know" encompasses all information collected about his or her health, whether it be a diagnosis, prognosis or any other relevant fact.
>
> …
>
> 70. Furthermore, it may be of vital importance for patients to know certain facts about their health, even though they have expressed the wish not to know them. For example, the knowledge that they have a predisposition to a disease might be the only way to enable them to take potentially effective (preventive) measures. In this case, a doctor's duty to provide care, as laid down in Article 4, might conflict with the patient's right not to know. It could also be appropriate to inform an individual that he or she has a particular condition when there is a risk not only to that person but also to others. Here too it will be for domestic law to indicate whether the doctor, in the light of the circumstances of the particular case, may make an exception to the right not to know."

Health data are personal data that should be afforded greater protection, especially when healthcare providers pioneer the use of new technologies, like AI systems. Healthcare providers and healthcare professionals act as gatekeepers in what health data are collected and therefore have responsibilities in safeguarding patient data and preserving the medical confidentiality of the patient–doctor relationship.

Health data are becoming more dynamic and complex. As the health data space expands, larger amounts of data from patient records and other sources (signal data, medical notes and speech, lab results and data collected from prescribed wearables) will likely be aggregated into data platforms and shared with and between other AI systems and healthcare settings across borders. In this context, patients will likely play a more altruistic and active role in managing their health data, to freely determine whether and how it is used by AI systems.

Challenges

How AI systems use health data is a challenge requiring further reflection. As AI systems train, share and infer new patterns and findings from these data, greater implications for privacy are likely to emerge. At its root, the role played by health professionals and healthcare providers in collecting, generating and enriching, as well as safeguarding health data, becomes ever more important.

- The protection of and respect for privacy in and out of healthcare settings is a cause for concern. One arrangement between a company and healthcare institution resulted in the company gaining access to over 1 million pseudonymised patient data files. Following a review of the deal, a court found that there was a violation of the right to privacy. The patients concerned had not been properly informed by the healthcare institution that their data had been shared. Without patients' knowledge or consent, this deal resulted in non-compliance with data protection law.[56]

While the protection of data is crucial, a degree of openness in the data collected from databases and aggregated into data platforms used to pre-train, train and validate AI systems may be crucial, as it would help to evaluate potential biases in the patterns and findings generated by AI systems, thereby mitigating any unfair treatment (discrimination) of different populations.

Recommended action

If data processing is based on explicit consent, subject to any requirements in national legislation, patients' consent to the use of their data by AI systems should be free, express and informed.[57] They should know what is collected and how it may be shared. They should be provided with assurances and possibly different types of

56. Hern A. (2017), "Royal Free breached UK data law in 1.6m patient deal with Google's DeepMind", *The Guardian*, available at www.theguardian.com/technology/2017/jul/03/google-deepmind-16m-patient-royal-free-deal-data-protection-act, accessed 21 October 2024.
57. See Recommendation CM/Rec(2016)8 of the Committee of Ministers to member States on the processing of personal health-related data for insurance purposes, including data resulting from genetic tests, available at https://search.coe.int/cm?i=09000016806b2c5f, accessed 21 October 2024.

consent options[58] to facilitate the collection and processing of data for AI systems. They must be able to decide not to be subject to automatic decision making by AI systems that significantly affect them.[59]

At the same time, healthcare providers should ensure that safeguards are in place to protect the privacy of patients throughout their healthcare journey, especially at the source of collection of health data. To this end, there should be ever more vigilance with patient data, mitigating any inadvertent or otherwise ambiguous data sharing with third parties.

Summary

Patients need:

- to decide what personal health data are being collected by AI systems;[60]
- assurances that privacy and personal data are being safeguarded, including not being subject to automatic decision making by AI systems.[61]

Health professionals and/or healthcare providers have the responsibility to:

- provide patients with options for consent in the collection and processing of their data by AI systems;[62]
- review what, how and when patient data are safeguarded, in accordance with data protection standards;
- train health professionals in understanding and managing the benefits and risks associated with AI systems, including the safeguarding of patient data.

58. The most common type of consent is informed consent for a treatment and/or intervention, research consent (biomedical consent), consent for the processing of personal data and a safeguard consent for the processing of health data. There are also other types of consent, as follows: Explicit Consent (consent required by the GDPR for the processing of special categories of personal data); Implied Consent (used in some countries also in healthcare); Granular Consent; General Consent (used for various things at the same time); Conditional Consent; Ongoing Consent (also can be considered to be a wide consent and used in some countries in healthcare and secondary use); Presumed Consent (used in some countries for the processing of health data when providing healthcare); Revocable Consent; and Dynamic Consent (which is the most preferred consent model for changing situations).
59. See Article 9 of the Modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data (Convention 108+), and Article 22 of the GDPR.
60. Subject to any requirements in national/European legislation regarding consent.
61. Subject to any requirements in national/European legislation regarding consent.
62. Subject to any requirements in national/European legislation regarding consent.

Equitable access to healthcare

Context

Article 3 of the Oviedo Convention refers to the provision of equitable access to healthcare of appropriate quality. Subject to health needs and available resources, this likely comprises equitable access to the benefits of AI systems used in healthcare. It is foreseeable that, as AI systems develop, the principle of equitable access to AI systems will become more important. This can be inferred from paragraph 24 of the Explanatory Report to the Oviedo Convention that refers to "a fitting standard in the light of scientific progress".

> "24. The aim is to ensure equitable access to health care in accordance with the person's medical needs. "Health care" means the services offering diagnostic, preventive, therapeutic and rehabilitative interventions, designed to maintain or improve a person's state of health or alleviate a person's suffering. This care must be of a fitting standard in the light of scientific progress and be subject to a continuous quality assessment.
>
> 25. Access to healthcare must be equitable. In this context, "equitable" means first and foremost the absence of unjustified discrimination. Although not synonymous with absolute equality, equitable access implies effectively obtaining a satisfactory degree of care."

It is expected that AI systems will help to redress and resolve issues concerning equitable access to healthcare, especially in countries with stretched healthcare services.

- For example, breast cancer is an increasing problem in low and middle-income countries where screening programmes for early detection are rare. In some countries, this has resulted in the development of cheaper, non-invasive

alternative tests that use thermal imaging and AI systems. Although considered less reliable than mammography, it is hoped that this test will help spot some early cancers in people who might not otherwise have access to mammography screenings.[63]

Measures taken by states to ensure equitable access may take many different forms and a wide variety of methods may be employed. Assessing the pros and cons of AI systems deployment may include conducting a cost-benefit analysis of AI system adoption, and the availability of affordable alternatives. At this stage, key questions to be posed include whether AI systems tie up less resources than traditional processes, whether they generate more health benefits for the patient, and whether the institution in question uses outdated systems technology, the overhaul of which has been neglected. Here, targeting limited resources and increasing equitable access are also essential when discussing the use of AI systems.

There is an opportunity for AI systems to mitigate pre-existing biases in modern medicine. In the future, AI might provide options to redress differences in access to healthcare that have historically disadvantaged certain groups such as older people, those with lower socio-economic status and ethnic minorities.

- Sex is an important determinant of health. Women and men can manifest different symptoms and react differently to treatments (it is well known that the pharmacokinetics and pharmacodynamics of pharmaceutical agents differ between sexes, resulting in differential adverse event profiles and further affecting treatment outcomes). This can be attributed to the gap in representation of women in clinical trials, leading to bias favouring male subjects.

Considering the therapeutic nature of the patient–doctor relationship, it is foreseeable that for reasons of cost and access medical consultations could shift from face-to-face encounters to online (telemedicine) meetings, some of which may substitute and/or complement a human doctor with an AI-enabled chatbot. It remains to be seen whether such a shift will result in a satisfactory degree of care because, arguably, the therapeutic relationship is one that is inherently human. In other words, the distance created by AI-enabled virtual assistants should not be overlooked. For example, AI systems (health checker apps) are unlikely to discern a patient's symptoms where underlying (that is, hidden, unquantifiable) causes (psychological, social, cultural) that require a greater understanding and trust-building process to be identified are at play.

Challenges

In any given country, equitable access to AI-enabled care will be challenged where its deployment is geographically uneven across health settings. Other inequities may result when such access is dependent on the financial means of individuals. A "two-tier" access to care could develop, with the wealthier having access to human doctors, and the less wealthy having access to AI-enabled chatbots or conversely, not having access to the latest AI developments, such as robotic surgery. These differences in access could be compounded should AI systems act as gatekeepers in determining care needs and treatments.

63. BBC News (2022), "Using artificial intelligence to spot breast cancer", available at www.bbc.com/news/av/stories-63755128, accessed 21 October 2024.

The challenge will be for different healthcare providers to define responsibilities for the management of AI systems as a whole and seeing whether adjustments to healthcare culture and practice are likely to be significant. It will take time for AI systems to become an integrated feature in "back office" infrastructure and "front office" patient care.

Computational infrastructures, where data are stored and/or algorithms run, may require additional investments. This could lead to inequity in AI system deployment and exploitation, exacerbating inequalities: wealthy territories could benefit more from AI system potentialities, collected data could be biased towards more wealthy social groups, and those less well-off could be penalised and marginalised from the use of AI-enabled devices that have not been adequately trained on representative populations. Herein lies a significant cause of concern and a challenge to be addressed: the design, training and validation of AI systems, using data.

There is a problem of bias throughout the design, development and training of AI systems. There may be bias in algorithms powering systems. There may also be bias in the data used to train, test and validate AI systems. Many other types of bias, such as contextual bias, should also be considered. The concern is that such (upstream) biases could foster (downstream) discrimination that adversely affects equitable access to healthcare, especially for under-represented people and groups.

- AI systems entail risks to individuals as well as to groups, communities and wider populations due to various biases (for example "automation bias" and racial bias in training datasets, as asserted in academic research). Bias that is found in AI systems deployed by healthcare providers may increase the propensity for accentuated and disproportionate risks to health for certain population groups.[64]
- In a study published in Science in October 2019, researchers found significant racial bias in an algorithm used widely in the US healthcare system to guide health decisions. The algorithm was based on cost (rather than illness) as a proxy for needs; however, the US healthcare system spends less money on black patients than on white patients with the same level of need. Thus, the algorithm incorrectly assumed that white patients were sicker than equally sick black patients. The researchers estimated that the racial bias reduced the number of black patients receiving extra care by more than half.[65]

Recommended action

A sustainable approach to providing access to healthcare using AI systems should be one that includes a human rights perspective to safeguard well-being and protect the dignity of everyone. This human rights perspective should be "end-to-end" throughout a patient's healthcare journey, from initial consultation through to treatment and care at home. This includes, in certain situations such as AI systems diagnostics, the possibility for patients to oppose the offer of AI-enabled care.

64. Adleberg J. (2022), "Predicting patient demographics from chest radiographs with deep learning".
65. World Health Organization (2021), "Ethics and governance of artificial intelligence for health", WHO Guidance, p. 54, available at www.who.int/publications/i/item/9789240029200, accessed 21 October 2024.

In their design, development and training phases, action is necessary to address biases in AI systems to mitigate the potential for discriminatory access to healthcare affecting people and groups (based on, for example, race, gender, age or disability). "Ethics by design" in the early stages of AI systems development and the value of human evaluations (impact assessments) can help to mitigate the effects of bias. More representative training datasets, bias benchmarking frameworks, and diversity in those tasked with assessing data quality should be considered.

Summary

Patients need:

- access to a human doctor;
- equitable access to the benefits of AI-enaabled health services with assurances that they promote patient well-being;
- the choice of opting for a blend of access to a human doctor and/or AI-enabled health services as well as "in-person" support when using them.

Health professionals and/or healthcare providers have the responsibility to:

- ensure that AI systems are clearly identifiable and distinguishable from human care;
- ensure that any tiered or blended options for access to AI health services are transparently explained;
- mitigate bias in all its forms (for example human impact assessments of AI systems, efforts to ensure the representativeness of training datasets, bias benchmarking frameworks, and diversity in those tasked with assessing data quality);
- ensure that medically prescribed AI-enabled apps do not act as gatekeepers in determinations about access to healthcare;
- maintain patient well-being, not dilute it with AI system offerings devoid of human interaction.

III. LOOKING AHEAD

Considering the trends in AI systems being developed and deployed in healthcare, there are many opportunities to improve the therapeutic patient–doctor relationship. However, concerns about how they might disrupt this relationship should not be overlooked. For example, a patient's journey might start with support from an AI chatbot (an AI app symptom checker), which is followed up by AI systems with diagnostic capabilities, all done prior to any involvement of a human health professional. Should AI systems be used in such ways, having a bearing on who, how, when and even whether access to healthcare is granted, the effects on the therapeutic relationship could be considerable.

That said, there must be trustworthiness in the professional standards that scrutinise the safety, quality and efficacy of AI systems; should this falter (for example when AI systems are considered to be inscrutable, inconclusive and even misguided),[66] patient autonomy will be weakened.

The trustworthiness of AI systems in healthcare depends on human oversight and the explainability of AI outputs. The "black box" character of AI systems has been criticised as having a bearing on the risk of bias and discrimination without good options for detecting such failures in performance. The responsible use of AI in healthcare relies at least on a basic understanding of the strengths of AI recommendations, the design features of the AI system and the technology's limitations with respect to specific groups of patients.

Positioning AI systems in the therapeutic relationship will require a shared approach as to their governance and application, including "bottom-up" public engagement and dialogues on their design, development and application, with an active role to be played by patient associations. In this connection, national bioethics committees can provide guidance to health decision makers, care facilities and patient associations.

Yet AI systems should never be considered solely as a means to improve the provision of cost-efficient healthcare, to the detriment of patient-centred care. The therapeutic relationship is a human construct, involving people making decisions based on cognition and social behaviour.[67] AI-enabled care should never be a substitute for people (in vulnerable situations) who need human professional contact and guidance. On the other hand, care should be taken to not put the patient in a worse position because AI systems are not used or otherwise denied.

66. Mittelstadt B. (2021), "The impact of artificial intelligence on the doctor–patient relationship", Council of Europe Publishing, Strasbourg.
67. WHO guidance (2021), "Ethics and governance of artificial intelligence for health".

Doctors and other healthcare professionals will require support in adapting to AI systems that guide their actions. They will need to be informed and trained accordingly, with considerable emphasis on their critical role in protecting and safeguarding patient well-being and quality of care. Introducing this in undergraduate education and specialised training for health professionals (for example to enable specialised medical teams to embrace new AI systems in their work) will be important.

Above all, AI systems should never undermine the therapeutic relationship, however good the intentions are. They must be made transparent to patients and doctors so that they are aware of what is running in the background. Patient autonomy and agency, coupled with human oversight by health professionals, are the path forward to strengthening the therapeutic relationship affected by AI systems.

The report explores the human rights implications of artificial intelligence systems used in healthcare, with a view to supporting doctors and other healthcare professionals in promoting the agency and autonomy of patients, patient welfare and equitable access to healthcare. Aimed at decision makers, health providers, health professionals and patients (including patient associations), the report focuses on selected human rights principles of particular relevance to the "patient–doctor" relationship, namely consent (Article 5, European Convention on Human Rights and Biomedicine (Oviedo Convention)), professional standards (Article 4, Oviedo Convention), private life and the right to information (Article 10, Oviedo Convention), and equitable access to healthcare (Article 3, Oviedo Convention).

PREMS 172924

ENG

www.coe.int

The Council of Europe is the continent's leading human rights organisation. It comprises 46 member states, including all members of the European Union. All Council of Europe member states have signed up to the European Convention on Human Rights, a treaty designed to protect human rights, democracy and the rule of law. The European Court of Human Rights oversees the implementation of the Convention in the member states.

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE