Ensure the deployment of trustworthy Artificial Intelligence (AI) in healthcare

While Artificial Intelligence can represent a clear advantage for healthcare in France and worldwide, only when it is designed with the public interest and Human Rights at the center can it become equally beneficial to all. For the past two years, Global Health Advocates has been researching risks and opportunities around the implementation of EU regulations and gaps in the existing frameworks in France. In the battle to shape AI policies, all stakeholders including civil society need to contribute to the debate. With the report "Soignons nos Algos", GHA presents its concrete proposals to ensure optimal conditions for the development of a trustworthy AI in healthcare.

I - Risks and opportunities of AI in healthcare

Artificial intelligence in healthcare holds great promise to improve the quality of care. Algorithms are already being used as cutting-edge tools for healthcare professionals (detection of cancers and complex pathologies) and the rapid evolution of research suggests even greater future benefits for medicine (generative AI, surgical assistance). AI can also empower patients and communities to assume control of their own health care and better understand their evolving needs.

While there are key advantages to using AI in healthcare, they often go together with important ethical concerns in relation to surveillance and privacy concerns, health and social inequity and the conditions necessary for trust and legitimate uses of data-intensive applications. In healthcare, AI raises two important challenges: **reducing the risk of bias and preventing privacy violations**.

When AI systems are not properly programmed, they can contain biases which, in the case of healthcare, can lead to serious inequalities and discrimination in access to care. Machine learning models often reproduce race and gender bias already present in current research and datasets. It is essential to test the quality of algorithms by making sure the database used for their training is relevant, diverse and reliable and by detecting the presence of bias right from the conception phase.

Since algorithms need to be trained on massive amounts of data, health data has become a goldmine[1] for AI conceptors. Their high value makes them even more vulnerable to leakage,

[1] Bochud M, Le Pogam MA, Thabard J, Monod S. Données de santé : le nouvel or numérique, mais pour qui ? [Health data: the new digital gold, but for whom?]. *Rev Med Suisse*. 2022. https://pubmed.ncbi.nlm.nih.gov/35822752/

theft, or malicious use, and heightens the risk of privacy violations. This can lead to identity theft, blackmail, or misuse by private companies (f.i. impeding access to loans, increasing cost of insurance). As cyber-attacks on healthcare establishments multiply, the risk of privacy violation or even harming the functioning of certain hospital departments increases. It is therefore crucial to ensure data safety by choosing the right storage and hosting organism.

II - New EU regulations impacting AI in healthcare

To regulate the use of AI, the European Union (EU) recently voted two binding regulations which will impact the healthcare sector: the EU AI Act, and the European Health Data Space (EHDS).

The **EU AI Act** regulates artificial intelligence technologies on the basis of their level of risk - limited risk, high risk and unacceptable risk. AI contained in medical devices is considered "high risk", and will therefore be subject to reinforced oversight and extensive assessment (transparency, good data governance and human oversight). Once EU institutions agree on the text, the AI Act will need to be implemented at national level. Member States will have to appoint and give proper resources to national authorities in charge of ensuring compliance and market surveillance.

The recently voted **European Health Data Space (EHDS)** establishes a regulation for health data storage and sharing in Europe, to facilitate **health data's primary use** (sharing of patients' medical records) and **secondary use** (access to anonymized, pooled health data for research). While the tech industry is enthusiastic about the EHDS, patient rights' advocates fear the regulation is not protective enough. To date, the negotiations' main sticking point lies with the patients' right to opt-out of secondary use of data.

III - AI in healthcare in France: a worrying lack of regulatory frameworks

In the battle to shape AI and digital technologies policies in France, few voices challenge the mainstream "start-up nation" narrative and France is reluctant to regulate the field. Even more so in healthcare where technologies are often seen as the future of medicine.

The "**French digital health acceleration strategy**" led to the creation of a digital ethic unit that published recommendations for the development of ethical AI systems in health[2] in 2022, as well as a repository of ethical criteria in 2023[3]. However, these criteria are not binding, and their use and scope have yet to be defined. The **French Public Health Code** regulates the use of medical devices with AI. This law notably helped improve transparency, patient information but remains vague when it comes to liability in case of medical errors and provides no measures to prevent bias, or respect for fundamental rights.

[2] Ministère des solidarités et de la santé, op. cit.

[3] Agence du Numérique en Santé, Référentiel thématique « Ethique de l'IA en santé ». https://esante.gouv.fr/sites/default/files/media_entity/documents/cens_referentiel-ethique-des-services-numeriques-de-sante-integrant-lia.pdf

The biggest issue is the lack of regulatory framework for the control and conception of AI systems and technologies for healthcare. Before a medical device using AI can be marketed in France, it must obtain a "CE" label, which affirms the goods' conformity with European health, safety, and environmental protection standards. However, the "CE" label does not attest to the clinical relevance, quality or efficiency of these tools, nor to their impact on human rights and discrimination. We currently can't be sure that an artificial intelligence tool is efficient, regardless of gender, age or skin color. There are also big question marks around other medical tools not strictly considered as medical devices (e-health app, etc.), falling out of the scope of current regulations.

In France, medical devices go through separate assessments, whether they are reimbursed or not by the French health insurance system. When they are reimbursed, they are evaluated by the French health authority (Haute Autorité de Santé). All other AI medical devices (those designed for doctors to help with diagnosis or analysis of medical data) and this includes the majority of them, are not currently evaluated. Yet, when they contain biases, these devices can also have a negative impact on patients' therapeutic pathways.

The lack of clear regulations governing technology hides in reality a desire to let private players regulate themselves. Private players have started to create a variety of labels certifying that their technologies meet certain ethical and/or quality criteria. However, this approach is inadequate considering the wide sets of criterias and variables being used. The proliferation of private labels raises questions about the willingness and ability of public institutions to play their regulator role and ensure a reliable and harmonized application of criteria.

IV - The Health Data issue

Since data is essential to the conception of AI systems, AI in healthcare is intimately linked to the issue of health data management and human rights. **The European General Data Protection Regulation (GDPR)** is a key regulatory tool to ensure that stakeholders processing personal data in Europe comply with certain standards[4]. However, when health data are anonymized - when processed for scientific research for example - the GDPR no longer applies. Since anonymised data can never be fully anonymised, it is crucial for governments to provide solutions to enhance data protection in this case. **The simplification of data access procedures[5] cannot work to the detriment of people's security and privacy**. Public authorities must conciliate simpler access to data with guaranteeing health data governance and control bodies' independence.

[4] "Intelligence artificielle : la CNIL dévoile ses premières réponses pour un eIA innovante et respectueuse de la vie privée", cnil.fr, 2023

[5] Ministère du Travail, de la santé et des solidarités, Rapport de la mission sur l'utilisation secondaires des données de santé, 2024, https://sante.gouv.fr/IMG/pdf/rapport_donnees_de_sante.pdf

Recommendations

In order to meet the challenges posed by the risks of AI in healthcare, and to follow the lead of the European Union in terms of AI regulation, France must guarantee the development of reliable and trustworthy AI in healthcare by:

- **Supporting a harmonized certification process for AI in healthcare by issuing a public label:**
    - This label must be built on EU's principles for an ethical AI and include the different provisions issued by the French Ethical Framework for Digital Health.
    - Issuing the label is conditional upon the algorithms being tested in real-life circumstances.
    - Compliance with this label should be a requirement for private players to benefit from public funding for AI.

- **Enhancing the assessment of AI systems used in healthcare, for which there are currently no control**
    - Reinforce the French Health Authority's (Haute Autorité de Santé) role in evaluating systems that are currently not controlled, such as digital medical devices for professional use and non-medical e-health applications.

- **Increasing resources for relevant national agencies and authorities to take part in the supervision and deployment of AI technologies in healthcare**
    - These resources are essential to ensure the enforcement of the European AI act and its specific provisions for high-risk systems: transparency, discrimination impact assessment, pre-market testing, validation of training datasets, etc.

- **Ensuring health data protection as part of AI technologies' development**
    - Data processing and data collection for research and innovation purposes must be handled ethically and with respect for humans' rights, particularly in the context of the implementation of the future European Health Data Space (EHDS).
    - Solutions must be found for sovereign storage of health data.
    - Independence of data access control bodies must be preserved.