Transparency and human oversight regarding medical diagnostic systems

Edinburgh Law School
donitiu@ed.ac.uk

Daria Onitiu
Research Associate, Governance & Regulation Node

Engineering and Physical Sciences Research Council

UKRI Trustworthy Autonomous Systems Hub

Key Questions

Transparency

- What are the transparency goals in medical diagnostic systems?
- What is the value of transparency for risk communication & management?

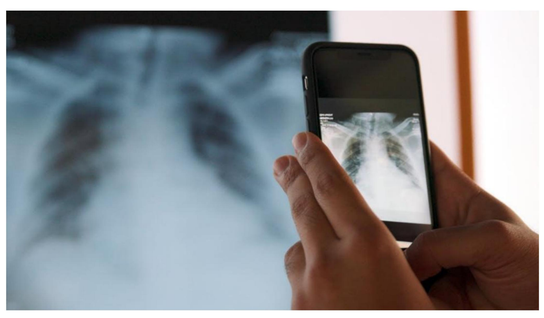
Accountability

- How are transparency goals aligned to patient autonomy & shared decision-making
- How do we ensure human agency with medical diagnostic systems operating on the ground?

Responsibility

- What is the role of human-in-the –loop concept considering the role of a healthcare professional's reliance on AI, rather than control?

Defining the Problem of Transparency



Picture taken from: Lynch 2020.

How AI *informs* diagnostic decisions is a key element to understand the role of safety & trust in medical diagnostic systems.

What needs to be addressed:

- How a system's real-life performance can be aligned to user interests and patient-centred outcomes
- How to ensure system verification is aligned to clinical judgement and patient values

Issues with Verification

Discrepancy regarding technical safeguards and how to ensure patient-centred outcomes:

- The EU Commission's proposal for the Regulation of Artificial Intelligence (AI Act proposal) gives manufacturers the opportunity to implement post hoc explainability methods in medical imaging and Uncertainty Estimates based on Articles 13 and 14 (AI Act proposal, arts 13-14). However, the proposal does not establish the parameters of how technical safeguards align with shared decision-making when the system interacts with the doctor and the patient.

Adaptive algorithms and intended uses

The role of verification within a social and organizational construct:

- The U.S. Food and Drug Administration (FDA) envisages a 'total product life-cycle approach' regarding the validation of AI /ML as medical device which allows for device modification when the system is operating on the ground, whilst ensuring a tool's continuous safety and use (FDA 2019, p. 3; FDA 2021, p.1). Nevertheless, further guidance needs to make the connection between a system's performance and reliability in individual circumstances more explicit (cf FDA, Health Canada, Medicines & UK Medicines and Healthcare products Regulatory Agency (MHRA) 2021, Principle 7).

Example:

- The [FDA] is reminding health care providers about the intended use of radiological computer-aided triage and notification (CADt) devices for intracranial large vessel occlusion (LVO)…Health care providers may not be fully aware that LVO CADt devices are intended for prioritization and triage only. These devices should *not be relied on when making any diagnostic decisions* (emphasis added) (FDA 2022).

Need for Cross-disciplinary engagement

Misalignment regard the role of technical safeguards to minimise impact on human safety and patient autonomy:

- Articles 13 and 14 AI Act proposal require further guidance on how technical safeguards align with patient perception of risk (AI Act proposal 2021, arts 13-14).
- Another important aspect of human oversight is regard measuring the degree of intervention regarding algorithmic changes. The FDA needs to consider the system's alignment with patient values, in addition to their 'Discussion Paper' and 'Action Plan' (FDA 2019, FDA 2021).

Lack of guidance on how a product safety approach aligns with human-centric regulation of AI (focus on AI Act proposal):

- '[a]n interdisciplinary approach to AI governance is tools testing a system's operation on the ground, considering user perspectives of the tool's reliability, a patient's perception of risk, as well as core ethical values in decision-making including patient-centred care. This outlook will eventually provide a more consistent approach to AI governance in healthcare, as well as legal certainty' (Onitiu 2022)

Accountability, Transaprency & Responsibility

Too many meanings?

- Further research needs to identify the relationships and gaps between transparency , accountability, and responsibility for coherence & effective multidisciplinary dialogue.
- Specific case studies on how multidisciplinary engagement could inform the entire product life-cycle and together with our industry partners.

References

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. COM/2021/206 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

FDA. (2019). 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. https://www.fda.gov/media/122535/download

FDA. (2021). Artificial Intelligence/Machine Learning (AI/ML)-Based: Software as a Medical Device (SaMD) Action Plan. https://www.fda.gov/media/145022/download

FDA (2022, April 11). *Intended Use of Imaging Software for Intracranial Large Vessel Occlusion - Letter to Health Care Providers*. U.S Food & Drug Administration. https://www.fda.gov/medical-devices/letters-health-care-providers/intended-use-imaging-software-intracranial-large-vessel-occlusion-letter-health-care-providers

FDA., Health Canada., MHRA. (2021). Good Machine Learning Practice for Medical Device Development: Guiding Principles. https://www.fda.gov/media/153486/download

Onitiu, D. (2022, April 4). The EU Commission's Proposal for the Regulation of Artificial Intelligence and "human-centric" Regulation for AI as Medical Device. KSLR EU Law Blog. https://blogs.kcl.ac.uk/kslreuropeanlawblog/?m=202204

Lynch, S. (2020, September 21). *The Geographic Bias in Medical AI Tools*. Stanford HAI. https://hai.stanford.edu/news/geographic-bias-medical-ai-tools.

THE UNIVERSITY of EDINBURGH
Edinburgh Law School