Establishing Global Guidelines for Integrating Artificial Intelligence Systems into Health Care and Policy

Ayush Sangari[1], Aditya Sood[2]

1. Renaissance School of Medicine at Stony Brook University 2. Emory University School of Medicine

INTRODUCTION

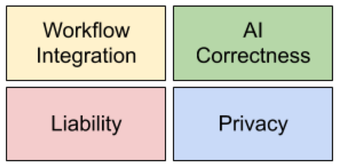
In an increasingly digitized world, artificial intelligence (AI) is poised to integrate itself into health care systems across the globe in order to improve clinical care by helping create personalized treatment plans, improving patient diagnoses, and increasing accessible care. Global health policies have already begun implementation of AI-based analysis in the areas of the tuberculosis and malaria epidemics. As AI increasingly enters the clinical sphere, specific guidelines will be critical in maintaining trust of clinical decision-making. The establishment of CONSORT, SPIRIT, and PRISMA guidelines were undeniably instrumental in ensuring transparency and minimum standards of accuracy in randomized clinical trials, interventional trials, and systematic reviews. A similar industry-wide framework must be created and adopted to facilitate the transition of AI into clinical decision-making.

OBJECTIVES

This study establishes a framework of system requirements necessary for the successful and ethical adoption of artificial intelligence systems with an emphasis on guidelines for global adoption.

METHODS

This study began by surveying existing regulations and guidelines for the use of artificial intelligence in clinical care from prominent regulatory agencies, non-profits, and think-tanks internationally. Four key topics of consideration were identified to guide this analysis:



These existing regulations and guidelines were then analyzed based upon their abilities to address the identified topics of consideration for diverse health care practices and models. New guidelines were also proposed to meet identified gaps between existing frameworks and current global needs.

REFERENCES

- 2018 AI Board Report Summary: Augmented Intelligence in Health Care. American Medical Association. Online. https://www.ama-assn.org/system/files/2019-08/ai-2018-board-policy-summary.pdf. Accessed Dec 31, 2020.

- Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ*. 2019;7:e7702. Published 2019 Oct 4. doi:10.7717/peerj.7702

- AI in Healthcare: Data Privacy and Ethics Concerns. Lexalytics. 12 Nov 2019. Online. https://www.lexalytics.com/lexablog/ai-healthcare-data-privacy-ethics-issues. Accessed Dec 31, 2020.

- Sheh, R., Monteath, I. Defining Explainable AI for Requirements Analysis. *Künstl Intell* **32,** 261–266 (2018). https://doi.org/10.1007/s13218-018-0559-3

- The emergence of artificial intelligence and machine learning algorithms in healthcare: Recommendations to support governance and regulation. BSI and AAMI. 2019. Online. https://www.bsigroup.com/globalassets/localfiles/en-gb/about-bsi/nsb/innovation/mhra-ai-paper-2019.pdf?_ga=2.123249483.1127287983.1574741777-2118595267.1574741777. Accessed Dec 31, 2020.

- Tjoa E, Guan C. A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. IEEE Trans Neural Netw Learn Syst. 2020 Oct 20;PP. doi: 10.1109/TNNLS.2020.3027314. Epub ahead of print. PMID: 33079674.

WORKFLOLW INTEGRATION

Proposed Guideline 1.1: Any decision made by a physician that impacts patient care should not solely be based on the recommendation of an artificial intelligence system.

Proposed Guideline 1.2: The terminology for artificial intelligence systems should be standardized, and classification of artificial intelligence systems in health care should be done in a risk-based manner.

Proposed Guideline 1.3: Physicians and other healthcare officials shall be required to take a standardized exam certifying competency in understating the best practices of how to use artificial intelligence systems in clinical practice.

WHY?

- Clinical decisions regarding care should be augmented by artificial intelligence systems, not driven by artificial intelligence systems in order to ensure that recommendations are scientifically sound.
- Standardization of the terminology for and classification of artificial intelligence systems is important for the development, use, and regulation of artificial intelligence systems.
- A risk-based classification scheme should be used instead of a complexity-based classification scheme since the risk-based version is easier to use, easier to interpret, considerate of patients, and applicable across a wide variety of clinical uses.
- Passing a standardized exam will provide assurance that physicians are successfully trained on how to safely and effectively incorporated artificial intelligence systems into their clinical practice, while allowing different nations to integrate the educational component in an ad hoc manner. Currently, the American Medical Association recommends that physicians receive education and training regarding using augmented intelligence in clinical care.

AI CORRECTNESS

Proposed Guideline 2.1: Artificial intelligence systems used in clinical settings should be free of bias.

Proposed Guideline 2.2: Artificial intelligence systems used in clinical settings should be approved by regulatory agencies.

Proposed Guideline 2.3: Artificial intelligence systems used in clinical settings should be frozen.

Proposed Guideline 2.4: Artificial intelligence systems should be able to accommodate monotonic constraints when necessary.

Proposed Guideline 2.5: Artificial intelligence systems should be able to meet the criteria for consent form comprehensibility.

WHY?

- As proposed by the British Standards Institution group, the datasets used to train artificial intelligence systems should be representative of the data that the system will be used on in practice, so that the system is able to generalize well and does not adversely affect certain groups of individuals due to biased predictions or outputs.
- Peer-reviewed literature is not sufficiently rigorous by itself to ensure the safety of artificial intelligence systems and algorithms designed for health care use. As such, these systems should be thoroughly reviewed by the regulatory agencies, which have the expertise, capacity, and interest to do so.
- Freezing AI systems/models prevents a decrease in quality of output caused by continual retraining.
- For a regulated industry such as health care, monotonic constraints should be imposed to ensure a predictive model aligns with common scientific belief.
- In order for artificial intelligence systems to be useful, they must be interpretable by health care providers and patients alike. Adapting the pre-existing criteria for consent form comprehensibility for artificial intelligence systems allows the use of various mathematical explainability techniques to meet the same criteria for explainability.

LIABILITY

Proposed Guideline 3.1: Physicians are liable for the decisions they make about patient care even if they consulted the assistance of honest artificial intelligence systems.

Proposed Guideline 3.2: Creators of artificial intelligence systems are liable if these systems are misleading or dishonest due to gross negligence, directly resulting in adverse patient care.

Proposed Guideline 3.3: Regulators are liable for adverse care resulting from adverse artificial intelligence systems mistakenly approved for use due to regulatory negligence.

WHY?

- In this proposed integration of artificial intelligence systems into clinical care, there is considerable burden on physicians to use the predictions and explanations of artificial intelligence systems responsibly and correctly. Therefore, physicians will face liability for care-related decisions made when using properly working, honest artificial intelligence systems.
- While physicians are primarily responsible for care-related decisions, this does not absolve the creators of artificial intelligence systems of negligence.
- Similarly, regulators are expected to thoroughly examine and approve each artificial intelligence system before it is used in clinical care.

PRIVACY

Proposed Guideline 4.1: Artificial intelligence systems should be trained anonymized training datasets.

Proposed Guideline 4.2: Patients are the owners of their own health data and can consent for their data to be used by clinical artificial intelligence systems.

Proposed Guideline 4.3: Artificial intelligence systems must disclose how they will use a patient's sensitive health data.

WHY?

- In order to protect the privacy of patients, all collected health data should be anonymized when used to train artificial intelligence systems for clinical uses, and the burden falls on those handling sensitive health data.
- Ensuring that patients are the owners of their own health data will enable them to legally control exactly how their health data can and cannot be used.
- Proposed Guideline 4.3 is taken from a recommendation from Lexalytics, and it is necessary in ensuring that patients are aware of the possible consequences of giving consent to another entity to use their sensitive health data.