Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240506

Monitoring of Artificial Intelligence in Hospitals

Arian RANJBAR[a,1] and Jesper RAVN[a]
[a] *Medical Technology and E-health, Akershus University Hospital, Norway*
ORCiD ID: Arian Ranjbar https://orcid.org/0000-0002-0422-2255

Abstract. Monitoring of artificial intelligence (AI)-based algorithms is necessary for safe implementation and will be required in upcoming regulations. This study investigates the potential for monitoring of AI in hospitals. First, by reviewing regulatory requirements and state of the art of monitoring. Then, by conducting a gap analysis of ISO42001, containing industry agnostic requirements harmonized with the EU AI Act. The analysis illustrates the need for comprehensive monitoring capable of capturing deviations in input, performance drifts and unintended interactions. However, hospitals often suffer from a technical debt, and the gap analysis provides qualitative indications on implementation challenges, including data quality, infrastructure and limitations in continuous improvement.

Keywords. Artificial intelligence, Monitoring, Implementation, Quality assurance

1. Introduction and Methods

Artificial intelligence (AI) has been heralded to significantly improve healthcare, and the desire to implement the technology is widespread. Contemporary AI systems are data driven, exploiting big data to find correlations and build non-linear instruction sets, often addressed as the black box problem. Even when showing great overall performance, models may suffer from variability in accuracy due to the data reliance. Errors may arise from either epistemic uncertainty, i.e. lack of knowledge, where the training data insufficiently covers the inference domain; or aleatoric uncertainty, i.e. by inherently stochastic events [1]. The former is addressed by defining an intended operational domain, enabling statistical validation. The latter is challenging to quantitatively evaluate since specific adversarial events could either have very low probability or not be known beforehand, while still posing a risk in aggregate. Similarly, the actual domain properties may drift over time, resulting in operation outside of the intended domain. For these purposes, monitoring have been suggested, both within clinical settings [2] and other safety critical domains [3]. This study investigates the potential for monitoring of AI in hospitals, in accordance with upcoming regulation and current state of the art.

A review of the EU AI act was made, together with the current state of quality assurance and monitoring of AI. Furthermore, a gap analysis of ISO42001 was conducted on a large tertiary acute care hospital. ISO42001 is a standard harmonized with the EU AI act, ensuring compliance, thereby giving an indication of monitoring requirements.

[1] Corresponding Author: Arian Ranjbar, Akershus Universitetssykehus HF, 1478 Lørenskog, Norway; E-mail: arian.ranjbar@medisin.uio.no.

2. Results, Discussion and Conclusions

Monitoring needs: The EU AI act explicitly demands monitoring of the systems, from article 14: "(…) be able to duly monitor its operation, also in view of detecting and addressing anomalies, dysfunctions and unexpected performance." [4]. Although monitoring may be supplied by algorithm providers as part of the solutions, as required by article 61, article 29 puts explicit responsibility on the deployer [4]. This is further covered by ISO42001, where monitoring is a core part of the management system [5].

The regulation aligns with the quality assurance literature and focuses on both sustained and isolated variations; where both input, output and relationships in-between needs to be monitored [2]. Local continuous evaluation of performance may not be sufficient, since a global perspective is necessary to capture unintended feedback loops or interactions [6], indicating a need for centralized monitoring systems within hospitals.
Implementation challenges: The gap analysis of ISO42001 qualitatively indicates challenges for healthcare organizations to adhere to the industry agnostic requirements, and hospitals already tend to suffer from a large technical debt [7]. One significant challenge relates to data quality. Monitoring output requires ground truth, often compared to outcome variables such as final diagnosis. However, as an example, diagnosis codes are known to lack in validity and timeliness [8]. Another challenge is infrastructure, where most hospitals run on distributed systems consisting of IT silos. This may limit the access to data other than algorithm input data [9]. Even if monitoring is in place, continuous improvement is currently prohibited by medical device regulation, as it would require re-certification of safety before use [10].
Conclusions: Extensive monitoring may be necessary to safely implement AI. A larger focus needs to be put on building a proper foundation for AI before implementation, including infrastructure and data quality. In parallel, legislators need to revise interactions between regulations for AI, to facilitate continuous improvement.

References

[1] Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach Learn* 2021; 110: 457–506.
[2] Feng J, Phillips RV, Malenica I, et al. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. *Npj Digit Med* 2022; 5: 1–9.
[3] Ranjbar A, Hornauer S, Fredriksson J, et al. Safety Monitoring of Neural Networks Using Unsupervised Feature Learning and Novelty Estimation. *IEEE Trans Intell Veh*.
[4] The European Union Artificial Intelligence Act - Regulation (EU) 2021/0106.
[5] ISO/IEC FDIS 42001. *ISO*, https://www.iso.org/standard/81230.html.
[6] Adam GA, Chang C-HK, Haibe-Kains B, et al. Hidden Risks of Machine Learning Applied to Healthcare: Unintended Feedback Loops Between Models and Future Data Causing Model Degradation. In: *Proceedings of the 5th Machine Learning for Healthcare Conference*. PMLR.
[7] Kraus S, Schiavone F, Pluzhnikova A, et al. Digital transformation in healthcare: Analyzing the current state-of-research. *J Bus Res* 2021; 123: 557–567.
[8] Quan H, Li B, Duncan Saunders L, et al. Assessing Validity of ICD-9-CM and ICD-10 Administrative Data in Recording Clinical Conditions in a Unique Dually Coded Database. *Health Serv Res* 2008.
[9] Ranjbar A, Ravn J, Ronningen E, et al. Enabling Clinical Trials of Artificial Intelligence: Infrastructure for Heart Failure Predictions. In: *Proceedings of Medical Informatics Europe*. 2023.
[10] The European Union Medical Device Regulation – Regulation (EU) 2017/745 (EU MDR).