Frameworks for Developing Machine Learning Models

Journal of Diabetes Science and Technology
2023, Vol. 17(3) 862–863

DOI: 10.1177/19322968231158399
journals.sagepub.com/home/dst
SAGE

Simon Lebech Cichosz, PhD[1]

Keywords
machine learning, good practice, best practice, transparent reporting, artificial intelligence, diabetes mellitus

Machine learning (ML) has the potential to revolutionize health care and diabetes care. One of the key benefits of using ML in diabetes care is its flexibility and scalability. This makes it suitable for a wide range of tasks, such as identifying patient risk levels, making diagnoses, and predicting outcomes. In addition, ML algorithms are capable of analyzing various types of data, including demographic information, laboratory results, imaging studies, and physician notes, and integrating that information to make predictions about disease risk, diagnosis, treatment options, and prognosis.[1,2]

However, its limited adoption in clinical care suggests that current strategies are deficient. Currently, ML solutions are often being developed in silos with a focus on insufficient studies of important aspects from development to validation, implementation, and clinical relevance of the models.[3] Good Machine Learning Practice (GMLP) includes aspects such as:

- Ensuring that the data used for training and testing the models are diverse, representative, and of high quality;
- Using appropriate evaluation metrics that are specific to the task and the population being studied;
- Being transparent about the limitations and uncertainty of the models;
- Considering ethical and legal implications, such as ensuring patient privacy and obtaining informed consent;
- Regularly monitoring and updating the models to ensure they remain accurate and current;
- Collaborating with domain experts, such as physicians and health care professionals, to ensure the models are clinically relevant and useful; and
- Regularly assessing the impact and value of the models in the real-world setting and making adjustments accordingly.

Over the last years, several frameworks for good practice and reporting have been proposed. It is important that these frameworks are adopted to a larger degree to ensure that more models are successfully implemented in clinical settings and can improve care, diagnostics, and workflow in a real-world world setting. In 2021, the US Food and Drug Administration (FDA), Health Canada, and the United Kingdom's Medicines and Healthcare Products Regulatory Agency (MHRA) published a framework with 10 guiding principles that can help inform the development of GMLP:

https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

In 2019, Wiens et al[3] published a framework titled "Do no harm: a roadmap for responsible machine learning for health care," which covers important aspects from choosing the right problem and data to ethical considerations and deploying the models responsibly. Recently, FDA has published a document of Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. The document presents an non-exhaustive list of 521 AI/ML-enabled devices across medical disciplines marketed in the United States as of January 31, 2023:

https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

Transparent and adequate reporting of published ML models is additionally very important to advance progress in the field and evaluate proposed models. The reporting quality of published ML models has been identified as insufficient in most studies, making it difficult to replicate, assess, and interpret study findings.[4] Numerous frameworks and checklists, such as the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD), TRIPOD-Artificial Intelligence (under development), and the *International Journal of Medical Informatics* (IJMEDI) checklist, exist and should be adopted by researchers and journals to ensure adequate reporting for ML models.[5,6]

[1]Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

Corresponding Author:
Simon Lebech Cichosz, PhD, Department of Health Science and Technology, Aalborg University, Fredrik Bajers Vej 7D2, Aalborg DK-9220, Denmark.
Email: simcich@hst.aau.dk

Abbreviations

FDA, US Food and Drug Administration; GMLP, good machine learning practice; IJMEDI, *International Journal of Medical Informatics*; MHRA, United Kingdom's Medicines and Healthcare products Regulatory Agency; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Simon Lebech Cichosz https://orcid.org/0000-0002-3484-7571

References

1. Kavakiotis I, Tsave O, Salifoglou A, et al. Machine learning and data mining methods in diabetes research. *Comput Struct Biotechnol J*. 2017;15:104-116. doi:10.1016/J.CSBJ.2016.12.005.
2. Cichosz SL, Johansen MD, Hejlesen O. Toward big data analytics: review of predictive models in management of diabetes and its complications. *J Diabetes Sci Technol*. 2016;10:27-34. doi:10.1177/1932296815611680.
3. Wiens J, Saria S, Goldenberg A. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med*. 2019;25:1337-1340. doi:10.1038/s41591-019-0548-6.
4. Yusuf M, Atal I, Li J, et al. Reporting quality of studies using machine learning models for medical diagnosis: a systematic review. *BMJ Open*. 2020;10:e034568. doi:10.1136/BMJOPEN-2019-034568.
5. Cabitza F, Campagner A. The IJMEDI checklist for assessment of medical AI [published online ahead of print May 28, 2021]. *Int J Med Inform*. doi:10.5281/ZENODO.4835800.
6. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162:55-63. doi:10.7326/M14-0697.