EDITORIAL

Twenty-six years of machine learning for ECG: and we are not there yet

John W. Pickering[1] · Laura Joyce[1] · Martin Than[1]

Machine learning (ML) and artificial intelligence (AI) have promised a lot for the clinician but delivered little to date, maybe in part because like all new tools it is taking time to learn to master. Yet, such is the strength of the hope to automate tasks that are subject to human error, or to find subtle relationships between independent and dependent variables, we keep attempting to apply machine learning. Automated identification of acute coronary syndrome (ACS) from the ECG is one such hope that still appears to be alive, simply because it could have a profound impact on busy emergency departments. The hope would be that early diagnosis may improve outcomes via earlier reperfusion [1].

In this issue, Zworth and colleagues present the evidence for adoption of machine learning algorithms to diagnose ACS with an ECG [2]. By including only studies with 12-lead ECG in the emergency department (ED) or pre-hospital setting, they focussed on studies that are of most clinical relevance, where an ECG prior to troponin measurements could affect patient disposition. Importantly, they limited the studies to those which compared ML to clinicians or non-ML-based software. Interestingly, of the ten studies identified, the earliest was 1997—considered within the AI winter when the hype had worn off and the computing power was insufficient. Four of the ten studies were pre-hospital and six in the ED; three were STEMI only, and four were externally validated. All but one used a form of neural networks as the training method.

In the four studies where areas under receiver operating characteristic curves (AUC) were provided, ML models had higher AUCs than clinicians. However, the validation data sets for some of the studies had the proportion of MI enhanced. This tends to artificially increase AUCs which makes their interpretation difficult. Zworth et al. also report the comparative sensitivities and specificities of the algorithms. The sensitivity was greater for the ML models than the clinicans, but at the expense of specificity. While most of the sensitivities were inadequate to safely exclude MI, two studies had sensitivities > 95% which is getting into the realm of diagnostic usefulness as a tool for stratifying patients to low risk of MI. One study had very high sensitivity and specificity for STEMI [3]. However, it is yet to be externally validated.

Also noted in the review was the possibility of bias in some studies and unclear definitions of ACS in half the studies (notably, some were conducted before the introduction of high-sensitivity troponin assays which has affected the diagnosis and definition of NSTEMI).

This paper has uncovered some hints that ML may aid diagnosis, but there is still a lack of high-quality studies that clearly demonstrate utility in a relevant clinical context. Zworth and colleagues have recommended future ECG databases record accompanying clinical context. Other technologies—high-sensitivity point-of-care (POC) troponin assays in particular—will change the ED and perhaps the pre-hospital context. Whereas previously the ECG was the most rapid method to identify some MI, new POC troponin devices mean that troponin results are available during the initial clinical exam somewhat rendering the "competitive advantage" of ECG mute. While early identification of STEMI remains a goal, and studies focussed on STEMI alone may result in utility, a paradigm shift to identification of occlusion myocardial infarction (OMI) is possible [4]. This would mean a change of primary outcome for future machine learning studies. We would also encourage those studies to have as outcomes risk stratification into those very unlikely to have MI, highly likely to have MI, and an intermediate risk group that requires further work-up. This paradigm is familiar to most who utilise accelerated diagnostic pathways (ADPs) for assessment of patients with possible ACS.

Zworth and colleagues deliberately excluded studies in which ECG was embedded within a diagnostic pathway which had other components including cardiac biomarkers. For

✉ Martin Than
martinthan@xtra.co.nz

1   Department of Emergency Medicine, Christchurch Hospital, University of Otago Christchurch, Christchurch, New Zealand

cohorts where STEMI has been excluded troponin is by far the strongest diagnostic component of risk scores like HEART or accelerated diagnostic pathways like EDACS-ADP. In this context, machine learning has created safe and effective models without ECG [5] or with a simple clinical judgment of whether the ECG displays evidence or not of ischaemia [6, 7]. It may be a fruitful area of study to establish if in statistical or machine learning algorithms the considerable data from an ECG actually contributes to the risk stratification and diagnosis of patients who are non-STEMI (or non-OMI). If not, then it may be time to to accept we are not there yet and wait for some considerable improvement in technology or AI.

Declaration

Conflict of interest All authors have no conflicts of interest to declare.

References

1. Lambert L, Brown K, Segal E, Brophy J, Rodes-Cabau J, Bogaty P. Association between timeliness of reperfusion therapy and clinical outcomes in ST-elevation myocardial infarction. JAMA. 2010;303(21):2148–55.
2. Zworth M, Kareemi H, Boroumand S, Sikora L, Stiell IG, Yadav K. Machine learning for the diagnosis fo Acute Coronary Syndrome using a 12-lead ECG: a systematic review. Can J Emergency Med. (2023). https://doi.org/10.1007/s43678-023-00572-5
3. Liu WC, Lin CS, Tsai CS, Tsao TP, Cheng CC, Liou JT, et al. A deep learning algorithm for detecting acute myocardial infarction. EuroIntervention. 2021;17(9):765–73.
4. McLaren JTT, Meyers HP, Smith SW, Chartier LB. From STEMI to occlusion MI: paradigm shift and ED quality improvement. Can J Emerg Med. 2022 Apr;24(3):250–5.
5. Than MP, Pickering JW, Sandoval Y, Shah ASV, Tsanas A, Apple FS, et al. Machine learning to predict the likelihood of acute myocardial infarction. Circulation. 2019;140(11):899–909.
6. Doudesis D, Lee KK, Boeddinghaus J, Bularga A, Ferry AV, Tuck C, et al. Machine learning for diagnosis of myocardial infarction using cardiac troponin concentrations. Nat Med. 2023;11(29):1201–10.
7. Neumann JT, Twerenbold R, Ojeda F, Aldous SJ, Allen BR, Apple FS, et al. Personalized diagnosis in suspected myocardial infarction. Clin Res Cardiol [Internet]. 2023 May 2 [cited 2023 May 3]. https://doi.org/10.1007/s00392-023-02206-3