European Heart Journal - Digital Health (2021) **2**, 125–126
doi:10.1093/ehjdh/ztab024

EDITORIAL

Artificial intelligence for automated ECG analysis: an experimental study revealing knowns and mysteries: still a long pathway ahead?

Peter P.T. de Jaegere*

Department of Cardiology, Thoraxcenter, Erasmus Medical Center, Dr Molewaterplein 40, 3015 GD Rotterdam, The Netherlands

Received 3 February 2021; accepted 5 February 2021; online publish-ahead-of-print 9 February 2021

This editorial refers to 'The effect of confounding data features on a deep learning algorithm to predict complete coronary occlusion in a retrospective observational setting', by R. Brisk *et al.*, on page 127.

In this issue of the *European heart Journal-Digital Health* (official abbreviation: *Eur Heart J Digit Health*), Brisk *et al.* report the performance of a Deep Learning (DL) algorithm using a deep convolutional neural network for automated ECG analysis to detect complete coronary occlusion.[1] The study is of an experimental nature in design including the origin and size of the population (ECG's). The authors used previously collected baseline ECG's and ECG's after 60 s of balloon-induced coronary occlusion in 104 patients (STAFF III database). Typically, the algorithm was developed by dividing the ECG dataset into training and testing sets. Of note, the performance of the algorithm was assessed in two situations, one in which the baseline ECG was recorded outside the operating theatre and the other in which the baseline ECG was recorded in the theatre. The algorithm performed well in the 1st but not in the 2nd situation due to background electrical activity in the theatres. Noteworthy, the algorithm used in the 1st setting (i.e. baseline ECG taken outside the operating theatre) outperformed clinical (cardiologist) ECG reading using STEMI criteria.

The technical and preclinical nature of the paper may fail to attract the attention or interest of the clinical reader and, as such, it is exemplar of the task of the *European Heart Journal-Digital Health*, namely building bridges and reaching out to both preclinical and clinical researchers and/or professionals, thereby, responding to the mission of the European Society of Cardiology, namely the reduction of the burden of cardiovascular disease through—among others—education and publishing.

This is not a simple task neither for the Editorial Board nor the preclinical researcher, nor the general reader. There is a large disparity between the world in which the preclinical and clinical expert/professional are trained and work in. They differ not only in background and mission but also language. To convey his/her message and, thus, relevance, the preclinical researcher must use a verbal account of the experiment that is understandable for the general reader (e.g. clinician). The latter must do an effort to try to understand that language, the concept of thinking and, hence, the methods used so to benefit of the information and get a feel of relevance and possible role in clinical practice.

The design of the study should instantly awake the clinical reader about generalizability of the findings and the complexity of algorithm development. The first does not need any further explanation, the second is more complex. The authors properly clarify the process of algorithm development by explaining the division of the dataset into training, validation, and testing sets. More specifically, 5-fold cross-validation (CV) was used for model evaluation, while the experiment itself was based upon a 10-fold CV process. What it means is that the model was trained on data from 80% of the patients and tested on data from the remaining 20% while for the experiment itself the data were split into 80% training, 10% validation, and 10% test sets to avoid overfitting of the five-fold CV. Overfitting was not defined but means that the algorithm is too complex for a dataset and, therefore, unreliable outside the training setting.[2] It calls for a critical selection of the relevant variables for a given problem.[3] A clinician may react to such a rather complex design and breaking down of the data by asking whether this affects the power (clinical utility) of the algorithm. Yet, it is mandatory for model training similar to an athlete who trains for an upcoming contest during which he/she has to prove his/her performance.

The opinions expressed in this article are not necessarily those of the Editors of the *European Heart Journal – Digital Health* or of the European Society of Cardiology.

Corresponding author.Tel: +31107035656, Email: p.dejaegere@erasmusmc.nl

The number of patients may raise a concern. For instance, the study was too small to assess performance of the 2nd model. Similar to clinical research, preclinical research starts with the raising of the appropriate question based upon what is known of the pathophysiology of the disease (mechanisms and expression) followed by the selection of the relevant target population (i.e. source of input variables), outcome measures and, henceforth, sample size. Given the design and absence of more detailed information from which stems this study population, the reader may remain suspicious about validity in this and more importantly the regular STEMI patient where automated analysis is important. Also, one may wonder how this DL-derived algorithm developed in this particular model of coronary occlusion compares to the less sophisticated ones currently used in patients with acute coronary syndrome of which detailed large high-quality datasets exist. These datasets can be used for external validation in populations in whom improved performance of automated ECG analysis is welcome.

The authors mention flaws in algorithm performance due to the well-known imposter of confounding that curses preclinical research as well. In this case, electrical background noise (e.g. fluoroscopy) leads to spurious results when the baseline ECG was recorded within the theatre. The authors do not elucidate the nature of the possible confounding variables to be considered for algorithm improvement. This obviously is not relevant for the setting used in this study (i.e. prediction of coronary occlusion during percutaneous coronary intervention) but raises the question what confounders are to be accounted for in the typical clinical setting and whether different ECG signals can or should be used other than the ones described in this study of which the rationale was not explained.

As the authors correctly point out in the Discussion, one needs to be aware of and possibly understand the pitfalls of whatever model of disease detection or prediction (AI-based or not). It is the responsibility of the preclinical researcher to report, instruct, and inform. It is the responsibility of the clinician to seek information and try to understand. It is the responsibility of the Editorial Board of the *European heart Journal-Digital Health* to build bridges and close the gap between pre- and clinical professionals since one cannot ignore the future of Digital Health while conceding that artificial intelligence stems from human intelligence. Artificial Intelligence (AI) cannot and may not operate in isolation. Human supervision and control permanently remain mandatory.[4]

Conflict of interest: none declared.

References

1. Brisk R, Bond R, Finlay D, McLaughlin D, Piadlo A, Leslie SJ, Gossman DE, Menown IB, McEneaney DJ, Warren S. The effect of confounding data features on a deep learning algorithm to predict complete coronary occlusion in a retrospective observational setting. *Eur Heart J Digit Health* 2021;**2**:127–134.
2. Hawkins DM. The problem of overfitting. *J Chem Inf Comput Sci*.2004:**44**:1–12.
3. Ribeiro JM, Astudillo P, de Backer O, Rojani R, Budde R, Van Mieghem N, Lumens J, Mortier P, Cummins P, Bruining N, de Jaegere PPT. Artificial intelligence and transcatheter interventions for structural heart disease: a glance at the (near) future. *Trends Cardiovasc Med* 2021;doi: 10.1016/j.tcm.2021.02.002.
4. Ribeiro JM, Cummins P, Bruining N, de Jaegere PPT. Patient-specific computer simulation in TAVR: is artificial intelligence superior to human experience in interventional cardiology? *JACC Interv* 2020;**13**:2581–2582.