European Heart Journal - Digital Health (2021) 2, 356–357
doi:10.1093/ehjdh/ztab066

EDITORIAL

The 12-lead surface electrocardiogram: a sheet of paper or a realm of concealed information asking for deep learning analysis

Peter de Jaegere[1]*, Joost Lumens [2], and Nico Bruining [1]

[1]Department of Cardiology, Erasmus MC, Dr Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands; and [2]CARIM School for Cardiovascular Diseases, Maastricht University Medical Center, Maastricht, The Netherlands

Received 19 July 2021 online publish-ahead-of-print 9 August 2021

It has often been said that one sees further by standing on the shoulders of giants. In a recent editorial, it is questioned whether this should be changed into 'standing on the shoulders of crowds'.[1] This appears to be justified in an era in which the analogue world is complemented by a digital one leading to a multitude of undertakings in which machine- and in particular deep learning (DL) techniques are proposed and evaluated to facilitate and improve the management of health and disease of our personal lives and planet.[2,3]

For that, we need data that represent and define the physical reality within and around us. The latter expresses itself by a continuous flow of variable physical quantities of different nature (e.g. voltage, photons, pressure, . . .) that are collected by instruments exploiting the specific physical properties of the naturally occurring phenomena. For instance, the heart does not only produce sound waves as a result of mechanical processes within the heart but also electric ones as result of a constant change in membrane potential induced by a continuous motion of charged particles (ions) across the cell membrane. Willem Einthoven (1860–1927, Leiden, The Netherlands) was the first to construct a machine capable of capturing small changes in voltage at the surface of the human body with a reasonable degree of sensitivity and speed, reflecting the electrical activity of the heart.[4]

The ability to convert an analogue (e.g. electrical current, light, temperature, force, . . .) into a digital signal (i.e. digitalization) that can be read by a computer has opened a vast array of novel data engineering and analysis techniques that help to facilitate diagnostic assessment by automation and to unveil 'hidden' patterns unseen or undetected by the human eye or brain.

The European Heart Journal – Digital Health (*EHJ-DH*) focusses on digital health (DH) in cardiovascular medicine and strives to connect physicians, researchers, and allied health professionals via publications and education, thereby, stimulating communication and knowledge transfer between all interested in DH. The EHJ-DH editorial board acknowledges the efforts all may need to undertake to understand each other's language and often esoteric jargon. Two elegantly written papers in this issue are an illustration of how the Journal's objectives can be achieved.[5,6]

Al Hinai *et al.* summarize the current status of end-to-end DL analysis of the resting electrocardiogram (ECG) for the detection of structural heart disease defined by left ventricular dysfunction, hypertrophy, and ischaemic heart disease. From a literature research conducted in 2019, they found 12 out of 1271 publications that fulfilled the search criteria. As mentioned, they focused on 'end-to-end DL' ECG analysis, which means DL for feature extraction (i.e. data engineering) and classification (i.e. data analysis). Papers only reporting performance of machine learning (ML) algorithms were excluded as these studies required an analysis or non-DL technique to extract features of interest from the ECGs. Based upon accuracy, sensitivity, and specificity of the DL-models, the authors found that DL algorithms and, in particular, those using Convolutional Neural Network (CNN) achieve a high degree of accuracy in detecting structural heart disease and outperform traditional computer-aided interpretation and non-DL ML algorithms based on support vector machines, random forest, or logistic regression. This conclusion needs to be seen in the context of the limitations of this review paper, acknowledged by the authors. For instance, does the conclusion hold when the investigated DL models (algorithm, architecture) are applied in different populations and whether data are extracted from ECGs of different vendors, etc.

But there is more. Machines and algorithms are like humans; they are not infallible. They may miss and/or misclassify features and, hence, may provide an erroneous output when exposed to the data of a diagnostic test. After all, humans are the architects of the software designed to perform specific tasks. While the output of expert reading of an ECG (i.e. human intelligence, HI) may include 'I am not

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

* Corresponding author. Tel: +31107036969, Email: p.dejaegere@erasmusmc.nl

(so) sure of my classification', uncertainty of predictions by artificial intelligence (AI) technologies is hardly considered. This is the rationale behind the work presented by Vranken *et al.* who investigated the performance of various uncertainty quantification methods applied to DL-based ECG analysis. For that purpose, the ECGs of three data sets (two local, one public) were exposed to a Deep Neural Network (i.e. Inception Residual Network, IRN). Then, six methods of uncertainty quantification were applied, accounting for epistemic uncertainty related to lack of knowledge of the algorithm and that can be addressed by additional training and for aleatoric uncertainty caused by noise in the ECG signal and that can be addressed by adding more data.

The main findings—stylishly illustrated by the wealth of Tables and Figures—are that regular Deep Neural Network (DNN) ECG analysis (i.e. IRN) is roughly 30% over- or underconfident with a higher prevalence of overconfidence (e.g. overfitting) in small data sets and under-confidence (e.g. underfitting) in large data sets and, that implementing uncertainty measures improves performance of automated ECG analysis using a DNN architecture. Interestingly, playing with the uncertainty thresholds, thereby rejecting uncertain ECGs, improves accuracy in the remaining ECGs also in a data set that the algorithm has not seen before (out-of-distribution data, sort of validation). Besides this succinct summary, the authors discuss many other technical aspects and limitations of their proposal plus automated ECG analysis stimulating future research and signalling a word of caution in case one would take an output for granted, whatever its source.

As mentioned before, the Editorial Board of the *European Heart Journal-Digital Health* wants to build bridges between disciplines and to close the gap between pre-clinical and clinical professionals. After all, one cannot ignore the future of Digital Health while conceding that AI cannot and may not operate in isolation.[7,8] Human intelligence and supervision, and, henceforth, responsibility will always remain obligatory of which the two herein discussed papers—that are a pleasure to read—are a testimony thereof.

Conflict of interest: none declared.

References

1. Anonymous. Standing on the shoulders of giants. *Nature* 2021;**594**:301–302.
2. Dignum V. AI—the people and places that make, use and manage it. *Nature* 2021; **593**:499–500.
3. Hohn N, Fleming O, Zhang R. *This AI Technique Could Use a Digital Version of Earth to Help Fight Climate Change*. https://www.weforum.org/agenda/2021/06/10.
4. Einthoven W. Un nouveau galvanomètre. *Arch Néerland Sci Exact Nature* 1901;**6**: 625–633.
5. Al Hinai G, Jammoul S, Vajihi Z, Afilalo J. Deep learning analysis of resting electrocardiogram for the detection of myocardial dysfunction, hypertrophy and ischemia: a systematic review. *Eur Heart J Digit Health* 2021.
6. Vranken J, van de Leur R, Gupta D, Orozco L, Hassink R, van der Harst P, Doevedans P, Gulshad S, van Es R. Uncertainty estimation for deep learning-based automated analysis of 12-lead electrocardiograms. *Eur Heart J Digit Health* 2021.
7. 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.
8. de Jaegere PPT. Artificial intelligence for automated ECG analysis: an experimental study revealing knowns and mysteries: still a long pathway ahead? *Eur Heart J Digit Health* 2021;**2**:125–126.