Digital medicine

What's lurking in your electrocardiogram?

For decades one of my favourite tasks in medicine has been reading 12-lead electrocardiograms (ECGs). I've always thought the wealth of information provided was impressive—eg, conduction and heart rhythm abnormalities, lack of blood supply or damage to the heart, chamber enlargement or hypertrophy, and inflammation of the pericardium. In the 1980s, when I did emergency coronary angiograms for patients with acute myocardial infarction, I marvelled at how the ECG accurately predicted the infarct-related artery and whether the occlusion was proximal or distal. But today I realise that whatever I could discern from a close review of an ECG was only rudimentary—it was just human.

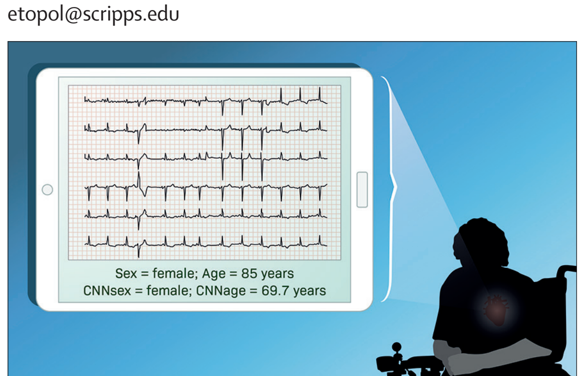
Now deep learning of ECGs is underway. Several reports have used hundreds of thousands of 12-lead ECGs to train deep neural networks to identify things that humans can't detect when reviewing ECGs. For example, ECGs from over 770 000 patients were used to determine sex and age, within 7 years, with high levels of accuracy. In another study, the ECGs from almost 50 000 patients were used to detect anaemia with an accuracy of over 90% when integrated with age and sex demographics. ECG data were also used to estimate the likelihood of a low ejection fraction (<35%) in over 100 000 patients. And there has been automation of detection of amyloid heart, hypertrophic cardiomyopathy, or mitral valve prolapse. These are all the result of extensive training from large annotated datasets to enable detection of clinical features that were not readily obtainable with any level of accuracy via expert ECG interpretation. There are also preliminary studies that suggest diagnosis of type 2 diabetes or COVID-19 might be possible via deep learning of ECGs. Additionally, two large datasets suggest the potential for predicting atrial fibrillation in people who are in normal sinus rhythm. More far-reaching was the use of 2·3 million ECGs from over 530 000 patients older than 34 years to predict 1-year mortality with a model performance area under the curve of more than 0·85.

The potential of machine algorithmic ECG detection and prediction is encouraging, but there are major limitations and concerns that have to be addressed. All of these studies are retrospective and have yet to be confirmed through prospective assessment. These models are still missing clinical validation for performing at the level needed when used in clinical practice. But even if these machine ECG interpretative capabilities hold up to prospective assessment, another problem lies with utility. We don't need an algorithm to determine a person's gender. The ECG estimate of age is intriguing since it may be possible that the machine determination is accurate for physiological rather than chronological age, such as telomeres or the epigenetic clock; however, although of academic interest, it is not a practical or proven tool relevant for understanding ageing. By contrast, accuracy in predicting atrial fibrillation for a given individual could incentivise a person to reduce modifiable risk factors, such as reduction of excessive alcohol intake or bodyweight. But the accuracy or potential favourable impact of this approach has yet to be prospectively validated.

More troubling is the use of machine algorithmic ECG to predict death at 1-year. Such information has no clear clinical usefulness but might be used in a perverse way by life insurance companies to raise a person's rates or deny coverage. Life insurance companies could exploit an insured person's ECG in this way, even though it could have an untoward effect for the individual. This use of people's data without their knowledge or consent conceptually represents an extension of the recent finding that facial recognition of each person in a crowd could be algorithmically processed to identify people with atrial fibrillation. The very subtle machine-detected changes in facial colour with each heartbeat potentially enable identification of people without symptoms. Once again, this technology might be used in either a helpful way to reduce risk of stroke or, adversely, for discrimination, or some form of penalty.

Deep learning of ECGs exemplifies the double-edged sword of artificial intelligence in medicine: it has the potential to ameliorate diagnoses while at the same time promote misuse. Since we are still in the early days of applying deep neural networks to ECGs, and most other medical scans and images, it is worth considering all the things we could detect with machine help but haven't yet attempted. However, we also need to question the value of such data. Some of this information will not be useful at all. More importantly, vigilance is needed to anticipate how such automation could be used the wrong way.

Eric J Topol
Scripps Research Translational Institute and the Scripps Research Institute, La Jolla, CA 92037, USA
etopol@scripps.edu



Faith Hark/Scripps Research Translational Institute

For more on **Digital medicine** see **Comment** *Lancet* 2016; **388:** 740 and **Perspectives** *Lancet* 2020; **396:** 1874

Further reading

Attia ZI, Friedman PA, Noseworthy PA, et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. *Circ Arrhythm Electrophysiol* 2019; **12:** e007284

Kwon JM, Cho Y, Jeon KH, et al. A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. *Lancet Digit Health* 2020; **2:** e358–67

Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med* 2019; **25:** 70–74

Tison GH, Zhang J, Delling FN, Deo RC. Automated and interpretable patient ECG profiles for disease detection, tracking, and discovery. *Circ Cardiovasc Qual Outcomes* 2019; **12:** e005289

Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet* 2019; **394:** 861–67

Raghunath S, Ulloa Cerna AE, Jing L, et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nat Med* 2020; **26:** 886–91

EJT is supported by the US National Institutes of Health/National Center for Advancing Translational Sciences grant UL1TR001114