



Application of Neural Networks to 12-Lead Electrocardiography

— Current Status and Future Directions —

Shinichi Goto, MD, PhD; Shinya Goto, MD, PhD

The 12-lead electrocardiogram (ECG) is a fast, non-invasive, powerful tool to diagnose or to evaluate the risk of various cardiac diseases. The vast majority of arrhythmias are diagnosed solely on 12-lead ECG. Initial detection of myocardial ischemia such as myocardial infarction (MI), acute coronary syndrome (ACS) and effort angina is also dependent upon 12-lead ECG. ECG reflects the electrophysiological state of the heart through body mass, and thus contains important information on the electricity-dependent function of the human heart. Indeed, 12-lead ECG data are complex. Therefore, the clinical interpretation of 12-lead ECG requires intense training, but still is prone to interobserver variability. Even with rich clinically relevant data, non-trained physicians cannot efficiently use this powerful tool. Furthermore, recent studies have shown that 12-lead ECG may contain information that is not recognized even by well-trained experts but which can be extracted by computer. Artificial intelligence (AI) based on neural networks (NN) has emerged as a strong tool to extract valuable information from ECG for clinical decision making. This article reviews the current status of the application of NN-based AI to the interpretation of 12-lead ECG and also discusses the current problems and future directions.

Key Words: Artificial intelligence; Convolutional neural network (CNN); Electrocardiography; Machine learning; Neural network; Recurrent neural network (RNN)

The 12-lead electrocardiogram (ECG) is a non-invasive and easy to conduct, but potentially powerful tool for diagnosing cardiac disease or evaluating the risk of future cardiac events. Surface lead ECG was first developed by Willem Einthoven in 1901.¹ It provides important diagnostic and treatment indications for conditions such as arrhythmia,² cardiac hypertrophy,³ and myocardial ischemia,⁴ and for those that require immediate attention, such as acute coronary syndrome (ACS).⁵

While ECG contains rich information regarding cardiac disorders such as myocardial ischemia, interpretation of the results is not easy. Intensive training for physicians is necessary and, even with this, there is still huge interobserver variability.

Artificial intelligence (AI) has emerged as a powerful tool to automatically interpret medical data. It has been applied to medical imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI),⁶ and echocardiography,⁷ as well as to ECG. AI technology has been successful in multiple tasks of automatic interpretation of ECG. These studies have further shown that ECG may contain more profound pathophysiological information that is not understood even by well-trained cardiologists.⁸ This article will review the current state of the application of neural network (NN) AI to the interpretation of 12-lead ECG, and discuss the current problems

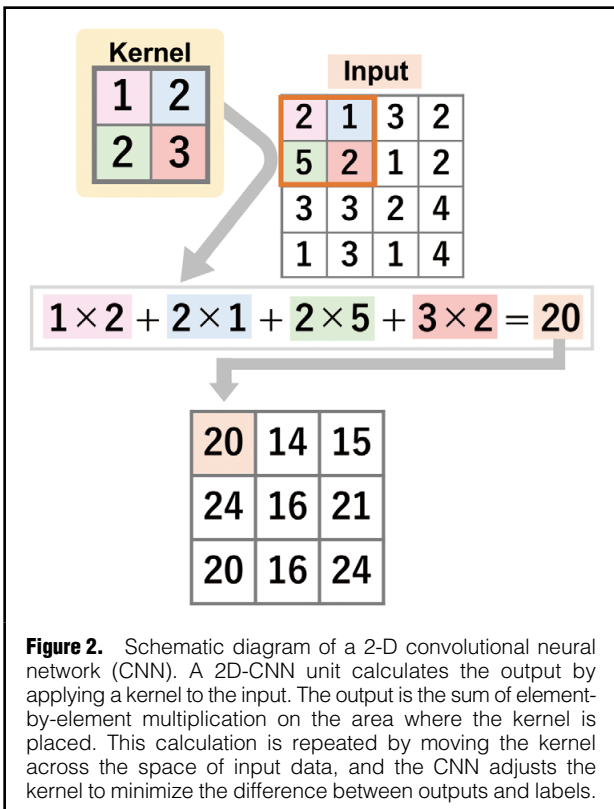
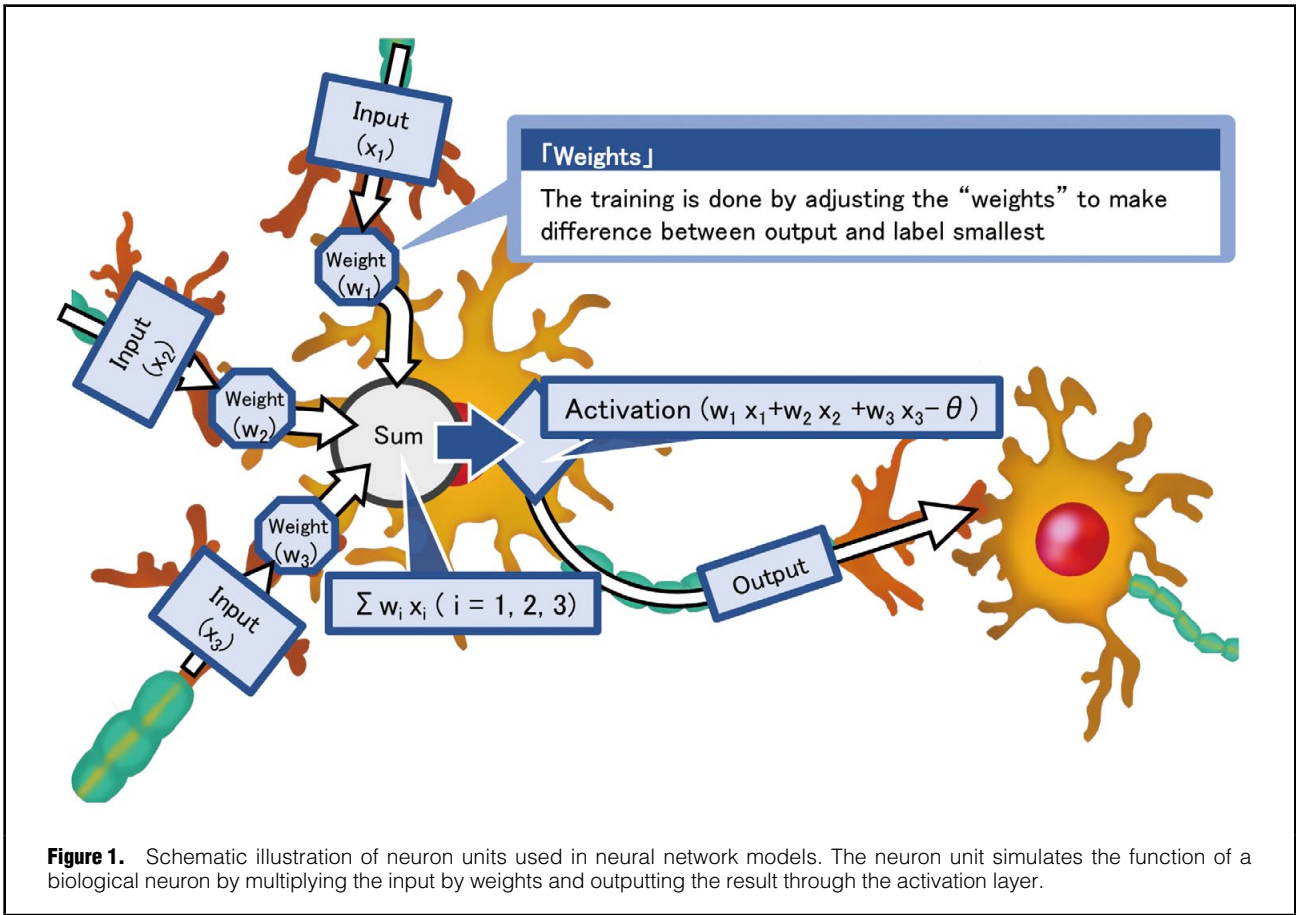
and future directions of this technology in the current medical field.

AI in the Medical Field

AI is a broad term representing the replication of human “intelligence” with machines that include but are not limited to computers. In the current medical field, the term “AI” mainly refers to a complex machine-learning model that automatically extracts information from data and which is, in most cases, based on an NN. Therefore, in this article, the term “AI” will be used to refer to such statistical models and the methods to create them.

The use of AI is in rapid development in many areas such as image recognition,⁹ language recognition¹⁰ and automatic driving.¹¹ In particular, in the image recognition field, AI has achieved results exceeding human ability in speed and accuracy. The technology has received attention from the medical field because it may extend the utilization of complex medical data beyond the limit of human brain function that is unable to handle high-dimension data.^{12,13} As expected, the application of AI to medical still-images has been greatly successful,^{7,14,15} but AI has also produced striking results in other tasks such as electronic digitalized health record analysis,^{16,17} and prediction of clinical outcome from ECG.^{1,18–33} Of note, AI applied to the interpretation

Received September 25, 2019; accepted September 26, 2019; J-STAGE Advance Publication released online November 2, 2019
Department of Cardiology, Keio University School of Medicine, Tokyo (Shinichi G.); Department of Medicine (Cardiology), Tokai University School of Medicine, Isehara (Shinichi G., Shinya G.), Japan
Mailing address: Shinya Goto, MD, PhD, Department of Medicine (Cardiology), Metabolic Disease Research Center, Tokai University Graduate School of Medicine, 143 Shimokasuya, Isehara 259-1193, Japan. E-mail: shinichi@is.icc.u-tokai.ac.jp
ISSN-2434-0790 All rights are reserved to the Japanese Circulation Society. For permissions, please e-mail: cr@j-circ.or.jp



of 12-lead ECG has shown a potential beyond the ability of even a well-trained cardiologist, and will be the main focus of this article.

Ability and Limitations of NN

Before going into detailed specific discussion of NN application to 12-lead ECG, this section will explain the general strength and limitation of current NN model. NN is a network constructed by multiple units of statistical model called “neuron” that simulates the function of neuron cells (Figure 1). A single unit of neuron takes multiple inputs and calculate an output by multiplying the input with its internal parameters called “weights” and summing them up. In most cases, the output of the unit will go through a non-linear activation function and passed on to the next neuron unit. NN constructed from these units are capable of replicating a complex non-linear function if the network layers are deep enough. What makes this structure called an AI is the fact that these networks can be trained with pairs of multi-dimensional input and mono/multi-dimensional outputs to produce a model that allows calculation/prediction of the output from the input without prior knowledge of the underlining mechanism or important features. This means that NN can automatically extract important features and automatically create classifiers or predictions models only from data. This characteristics of NN may allow tasks that could not be done by human. Human brains can hardly deal with data with more than

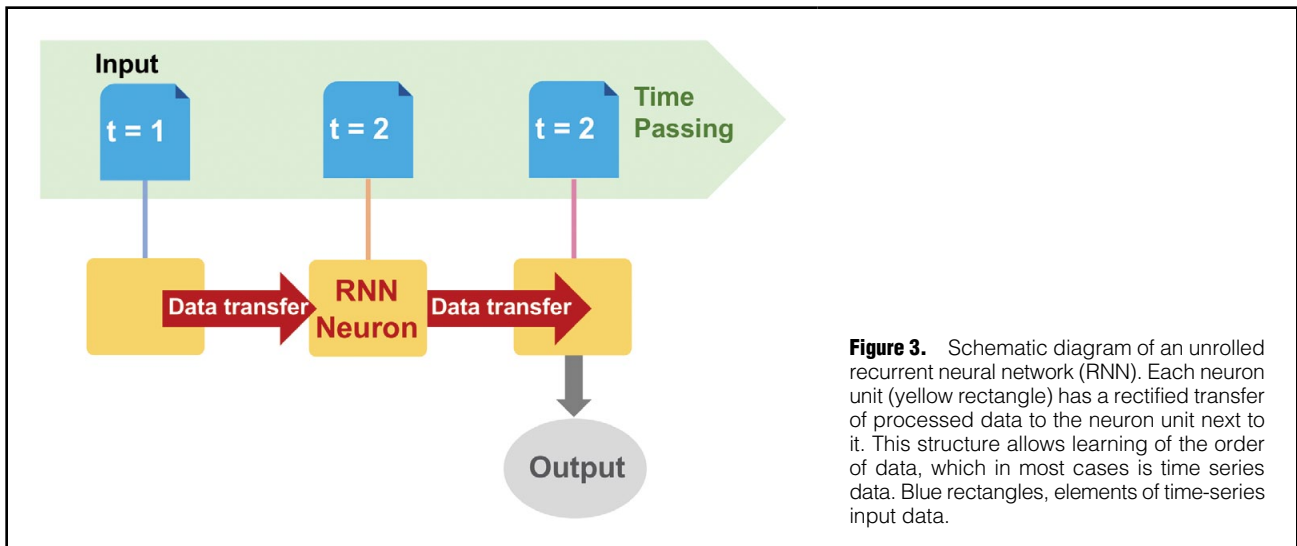


Figure 3. Schematic diagram of an unrolled recurrent neural network (RNN). Each neuron unit (yellow rectangle) has a rectified transfer of processed data to the neuron unit next to it. This structure allows learning of the order of data, which in most cases is time series data. Blue rectangles, elements of time-series input data.

3-dimensions while NN can take an arbitrary dimensional input, which means that it can process extremely complex input and automatically produce meaningful models. Medical data contain large numbers of multiple-dimensional data and is a field of interest on AI application.

However, the strength of NN that it doesn't require human interpretation of data translates into its weakness. Although NN automatically extract important features and allows us to produce meaningful models, this does not mean that we can directly understand the features learned in the NN. In most cases, the features used by NN are difficult to extract in human understandable formats and it is also difficult to combine the human knowledge of mechanism into the model. Therefore, we are currently forced in the situation that we need to use the model without understanding what the AI are doing. This makes the AI model a black box and means that it is difficult to tell if the model is not suitable for the specific dataset that user wants to input. It is known that there are ways to engineer inputs that completely fakes AI models without being noticed by human.³⁴ These results suggest that NN uses different features from human and that the human perception of difficulty may not be useful for selecting suitable dataset for a network. It is necessary to share this common sense regarding the limitation of AI.

Types of NN for Special Tasks

The basic structure of NN as described in the previous section is called a "fully connected neural network" or a "multilayer-perceptron" (MLP). This structure is capable of learning features from multi-dimensional data, but they are not good at explicitly handling the location of data such as spatial or temporal position. Therefore, various special neuron units were developed to handle this problem. Of these special units, convolutional NN (CNN) and recurrent NN (RNN) are the most commonly used neuron units in the medical field.

CNN calculates outputs by performing "convolutions" using kernels. The method of calculation is shown in **Figure 2**. Instead of multiplying each input element by a weight, the 2-D CNN unit applies a kernel that is moved

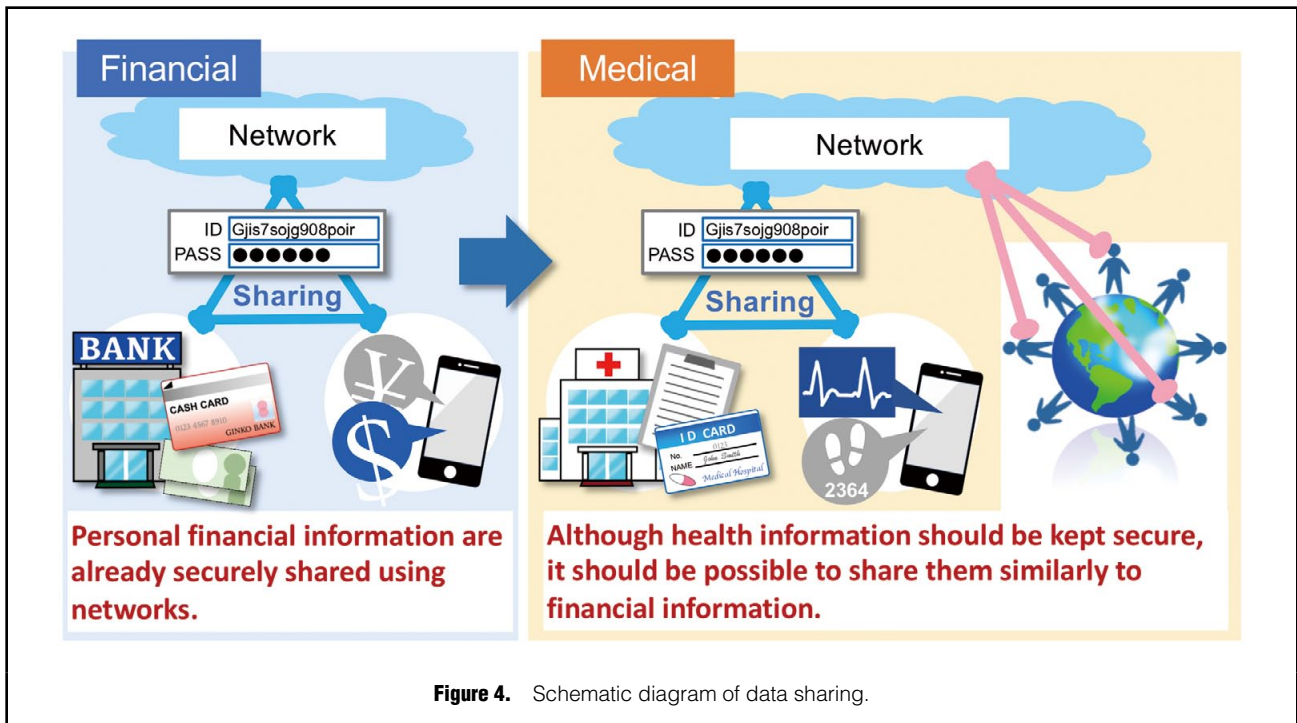
across the space of input data. The output is the sum of element-by-element multiplication on the area where the kernel is placed. Through training, the 2-D CNN unit then adjusts the kernel to minimize the difference between the outputs and the labels (ground truth). Kernels in the field of computer vision are used to extract specific features from images. Manually constructed kernels have been used for processes such as edge detection. CNN can automatically develop kernels that are suitable for extracting features that are important for the task it is used for. Thus, 2-D CNN emerged as a powerful tool to automatically extract features and classify images.

RNN, in contrast, returns the output of the neuron unit to an auxiliary input of the same unit. This can be understood as multiple units with a rectified data transfer between the units next to them (unrolling) as shown in **Figure 3**. This structure is sensitive to changes in data order in the input. Therefore, the neuron unit is useful in learning the order of data and can be used for time-series data and languages.

Application of AI to 12-Lead ECG

The early AI applied to 12-lead ECG used NN structures other than CNN or RNN (in most cases MLP with or without slight modification). They used manual methods or technologies other than NN to obtain the features to train the model. This approach was successful on lead reversal detection,²³ incomplete right bundle branch block pattern detection,²⁸ classification of ST-T segment,³⁵ classification of ECG into disease patterns (such as myocardial infarction and left ventricular hypertrophy) and detection of acute^{24,27,32} or healed^{25,33} myocardial infarction. These approaches, however, obviously did not fully use the ability of AI automatic feature extraction. The use of these approaches was limited by the selected features, which were usually based on human knowledge of ECG features. Thus, the applications were mainly focused on automating human tasks.

Improvements in computer and NN technology have allowed the development of deeper network patterns, which are capable of analyzing more complex data. These



improvements, along with the improvement in neuron units themselves, such as the introduction of CNN, RNN and so on, have allowed direct input of ECG data into the network. This cleared the way for AI models to perform tasks that are beyond human ability.

Using 2D-CNN on raw ECG data, Attia et al have predicted age and sex,¹ and detected cardiac contractile dysfunction²⁰ and atrial fibrillation from sinus rhythm ECG;¹⁸ Yang et al have detected the origin of premature ventricular contractions;²¹ and Tison et al have predicted the values of echocardiographic parameters and related heart diseases such as pulmonary artery hypertension, hypertrophic cardiomyopathy and cardiac amyloidosis.²⁶ Baloglu et al utilized a 1-D CNN network to extract lead features and combined all the features in the fully connected neuron. They successfully classified the location of myocardial infarction from ECG.³⁰ Our group used a combination of 1D-CNN with RNN and constructed a model that successfully identified patients who need urgent revascularization in the emergency room only from the 12-lead ECG.⁸ These new approaches allowed tasks to be carried out that were beyond the ability of even a well-trained expert cardiologist.

CNN or RNN for 12-Lead ECG

When applying AI technology to 12-lead ECG, the use of CNN and RNN can be seen as reflecting differences in how the ECG data structure is recognized. Using CNN is similar to the physician's way of dealing with ECG, treating the data as an image. Even though the input is structured as raw voltage data, this approach is similar to interpreting images because the CNN approach extracts features with kernels. CNN kernels are pattern recognizers and will be activated by specific wave patterns.

In contrast, RNN explicitly deals with the ordering of

data. Thus, the RNN approach treats the ECG data as a time-series. The effect on model performance of selecting CNN or RNN, however, is unclear, and the choice is currently made empirically.

There is no consensus on what tasks are suitable for CNN and what tasks are suitable for RNN. One way of choosing which structure to use involves understanding the difference in computational resource requirements for these units, especially in the ECG field. Generally, CNN are computationally cheaper than RNN. ECG is a field of rapid development of wearable devices, including implantable ones.³⁶⁻³⁹ Some devices such as implantable cardioverter defibrillators require real-time classification of ECG data, which is a possible field of AI application.⁴⁰ These devices usually possess limited computational resources. In these situations, CNN would be an attractive choice, and in some cases the only choice because of the limited resources, even when RNN can achieve better accuracy. If there is no limitation on computational resources, the networks should be compared head to head. If there is no merit in using RNN over CNN, CNN should be preferred. Theoretically, however, RNN can learn more precisely the time-series voltage data, which is the rawest form of ECG data. It is not unreasonable to assume that RNN can extract more data from ECG than CNN. Thus, some complex tasks may still require RNN. Identifying the tasks suitable for CNN and RNN is an interesting topic for future AI research in this field.

Current Problems and Future Directions

There are some limitations of AI usage that should be specifically noted with regard to the medical field. The first problem is the need for large-scale data with which to train NN. This is a general issue in NN usage but is especially important in the medical field, where we model diseases

that are rare. Ways to reduce the required input are being actively developed such as reinforcement learning. From the medical side, however, we should develop a system for data sharing in order to maximize the utilization of available data. We have already created a way to digitize and share the information in the field of financial data. This information is almost as important as health information and requires similar security. Thus, digitizing and sharing health information should also be able to be achieved (Figure 4).

The second problem is lack of human understanding of the ECG features. NN are capable of automatically extracting features from complex inputs such as raw ECG data. This is powerful when the objective is to create an automatic classifier or automatic prediction tool. The features identified by AI models, however, are difficult to extract in a human-understandable format. This makes the model a black box and limits the clinical use of these models because it is difficult to tell if the model works for a specific population or not. Another shortcoming is that the model is not able to enhance the understanding of the disease. Contemporary models such as logistic regression modeling provide information about which features are independently associated with outcomes and so on, which give mechanistic insights into the disease. There are ongoing efforts to extract human-understandable features from AI models and this is an important topic for future AI research.

Conclusions

AI models based on various NN complexes have a huge potential to assist in clinical decision making. Limitation in the understanding of the mechanism on how the AI model makes predictions, however, should also be noted.

Acknowledgments

The authors thank Mai Goto for the illustration of Figure 4. Shinya Goto acknowledges financial support from MEXT/JSPS KAKENHI 17K19669, partly by 18H01726 and 19H03661. Shinya Goto acknowledges grant support from The Vehicle Racing Commemorative Foundation and Nakatani Foundation for Advancement of Measuring Technologies in Biomedical Engineering.

Disclosures

Shinya Goto acknowledges financial support from Bristol-Myers Squibb from their independent research support project (33999603). Shinya Goto received research funding from Sanofi, Pfizer, and Ono. Shinichi Goto declares no conflicts of interest.

References

- Attia ZI, Friedman PA, Noseworthy PA, Lopez-Jimenez F, Ladewig DJ, Satam G, et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. *Circ Arrhythm Electrophysiol* 2019; **12**: e007284.
- Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB, et al. 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. *Circulation* 2018; **138**: e272–e391.
- Hancock EW, Deal BJ, Mirvis DM, Okin P, Kligfield P, Gettes LS, et al. AHA/ACCF/HRS Recommendations for the standardization and interpretation of the electrocardiogram. Part V: Electrocardiogram changes associated with cardiac chamber hypertrophy: A scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society: Endorsed by the International Society for Computerized Electrocardiology. *Circulation* 2009; **119**: e251–e261.
- Rautaharju PM, Surawicz B, Gettes LS, Bailey JJ, Childers R, Deal BJ, et al. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram. Part IV: The ST segment, T and U waves, and the QT interval: A scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society endorsed by the International Society for Computerized Electrocardiology. *J Am Coll Cardiol* 2009; **53**: 982–991.
- Amsterdam EA, Wenger NK, Brindis RG, Casey DE, Ganiats TG, Holmes DR, et al. 2014 AHA/ACC guideline for the management of patients with non-ST-elevation acute coronary syndromes. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2014; **64**: e139–e228.
- Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. *Nat Med* 2019; **25**: 44–56.
- Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, et al. Fully automated echocardiogram interpretation in clinical practice. *Circulation* 2018; **138**: 1623–1635.
- Goto S, Kimura M, Katsumata Y, Goto S, Kamatani T, Ichihara G, et al. Artificial intelligence to predict needs for urgent revascularization from 12-lead electrocardiography in emergency patients. *PLoS One* 2019; **14**: e0210103.
- He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016: 770–778.
- Kim JH, Kwon HS, Seo HW. Evaluating a pivot-based approach for bilingual lexicon extraction. *Comput Intell Neurosci* 2015; **2015**: 13.
- Pasquier M, Quek C, Toh M. Fuzzylot: A novel self-organising fuzzy-neural rule-based pilot system for automated vehicles. *Neural Netw* 2001; **14**: 1099–1112.
- Ribeiro AL, de Oliveira GMM. Toward a patient-centered, data-driven cardiology. *Arq Bras Cardiol* 2019; **112**: 371–373.
- Goto S, Goto S. Letter by Goto and Goto regarding article, “Fully automated echocardiogram interpretation in clinical practice: Feasibility and diagnostic accuracy”. *Circulation* 2019; **139**: 1646–1647.
- Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115.
- Yoshida H, Shimazu T, Kiyuna T, Marugame A, Yamashita Y, Cosatto E, et al. Automated histological classification of whole-slide images of gastric biopsy specimens. *Gastric Cancer* 2018; **21**: 249–257.
- Rajkumar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018; **1**: 18.
- Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019; **572**: 116–119.
- Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, 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**: 861–867.
- Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med* 2019; **25**: 65–69.
- Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med* 2019; **25**: 70–74.
- Yang T, Yu L, Jin Q, Wu L, He B. Localization of origins of premature ventricular contraction by means of convolutional neural network from 12-lead ECG. *IEEE Trans Biomed Eng* 2018; **65**: 1662–1671.
- Nugent CD, Webb JA, Black ND, Wright GT, McIntyre M. An intelligent framework for the classification of the 12-lead ECG. *Artif Intell Med* 1999; **16**: 205–222.
- Heden B, Ohlsson M, Holst H, Mjoman M, Rittner R, Pahlm O, et al. Detection of frequently overlooked electrocardiographic lead reversals using artificial neural networks. *Am J Cardiol* 1996; **78**: 600–604.
- Haraldsson H, Edenbrandt L, Ohlsson M. Detecting acute myocardial infarction in the 12-lead ECG using Hermite expansions and neural networks. *Artif Intell Med* 2004; **32**:

- 127–136.
25. Heden B, Ohlsson M, Rittner R, Pahlm O, Haisty WK Jr, Peterson C, et al. Agreement between artificial neural networks and experienced electrocardiographer on electrocardiographic diagnosis of healed myocardial infarction. *J Am Coll Cardiol* 1996; **28**: 1012–1016.
 26. 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.
 27. Ohlsson M, Ohlin H, Wallerstedt SM, Edenbrandt L. Usefulness of serial electrocardiograms for diagnosis of acute myocardial infarction. *Am J Cardiol* 2001; **88**: 478–481.
 28. Yang S, Yamauchi K, Nonokawa M, Ikeda M. Use of an artificial neural network to differentiate between ECGs with IRBBB patterns of atrial septal defect and healthy subjects. *Med Inform Internet Med* 2002; **27**: 49–58.
 29. Nugent CD, Webb JA, Black ND. Feature and classifier fusion for 12-lead ECG classification. *Med Inform Internet Med* 2000; **25**: 225–235.
 30. Baloglu U, Talo M, Yildirim Ö, Tan RS, Acharya UR. Classification of myocardial infarction with multi-lead ECG signals and deep CNN. *Pattern Recogn Lett* 2019; **122**: 22–30.
 31. Javadi M, Ebrahimpour R, Sajedin A, Faridi S, Zakernejad S. Improving ECG classification accuracy using an ensemble of neural network modules. *PLoS One* 2011; **6**: e24386.
 32. Heden B, Ohlin H, Rittner R, Edenbrandt L. Acute myocardial infarction detected in the 12-lead ECG by artificial neural networks. *Circulation* 1997; **96**: 1798–1802.
 33. Heden B, Edenbrandt L, Haisty WK Jr, Pahlm O. Artificial neural networks for the electrocardiographic diagnosis of healed myocardial infarction. *Am J Cardiol* 1994; **74**: 5–8.
 34. Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. 2014. arXiv:1412.6572.
 35. Edenbrandt L, Devine B, Macfarlane PW. Classification of electrocardiographic ST-T segments: Human expert vs artificial neural network. *Eur Heart J* 1993; **14**: 464–468.
 36. Reek S, Burri H, Roberts PR, Perings C, Epstein AE, Klein HU, et al. The wearable cardioverter-defibrillator: Current technology and evolving indications. *Europace* 2017; **19**: 335–345.
 37. Lauschke J, Busch M, Haverkamp W, Bulava A, Schneider R, Andresen D, et al. New implantable cardiac monitor with three-lead ECG and active noise detection. *Herz* 2017; **42**: 585–592.
 38. Friedman DJ, Parzynski CS, Heist EK, Russo AM, Akar JG, Freeman JV, et al. Ventricular fibrillation conversion testing after implantation of a subcutaneous implantable cardioverter defibrillator: Report from the National Cardiovascular Data Registry. *Circulation* 2018; **137**: 2463–2477.
 39. Kober L, Thune JJ, Nielsen JC, Haarlo J, Videbaek L, Korup E, et al. Defibrillator implantation in patients with nonischemic systolic heart failure. *N Engl J Med* 2016; **375**: 1221–1230.
 40. Rojo-Alvarez JL, Arenal A, Garcia-Alberola A, Ortiz M, Valdes M, Artes-Rodriguez A. A new algorithm for rhythm discrimination in cardioverter defibrillators based on the initial voltage changes of the ventricular electrogram. *Europace* 2003; **5**: 77–82.