Original Article

Yonsei Med J 2022 Jan;63 Suppl:S93-107
https://doi.org/10.3349/ymj.2022.63.S93

Yonsei Medical Journal
YMJ
pISSN: 0513-5796 · eISSN: 1976-2437

Artificial Intelligence for Detection of Cardiovascular-Related Diseases from Wearable Devices: A Systematic Review and Meta-Analysis

Solam Lee[1,2*], Yuseong Chu[3*], Jiseung Ryu[3], Young Jun Park[4], Sejung Yang[3], and Sang Baek Koh[1]

Departments of [1]Preventive Medicine and [2]Dermatology, Yonsei University Wonju College of Medicine, Wonju;
[3]Department of Biomedical Engineering, Yonsei University, Wonju;
[4]Division of Cardiology, Department of Internal Medicine, Wonju Severance Christian Hospital, Yonsei University Wonju College of Medicine, Wonju, Korea.

Purpose: Several artificial intelligence (AI) models for the detection and prediction of cardiovascular-related diseases, including arrhythmias, diabetes, and sleep apnea, have been reported. This systematic review and meta-analysis aimed to identify AI models developed for or applicable to wearable and mobile devices for diverse cardiovascular-related diseases.

Materials and Methods: The searched databases included Medline, Embase, and Cochrane Library. For AI models for atrial fibrillation (AF) detection, a meta-analysis of diagnostic accuracy was performed to summarize sensitivity and specificity.

Results: A total of 102 studies were included in the qualitative review. There were AI models for the detection of arrythmia (n=62), followed by sleep apnea (n=11), peripheral vascular diseases (n=6), diabetes mellitus (n=5), hyper/hypotension (n=5), valvular heart disease (n=4), heart failure (n=3), myocardial infarction and cardiac arrest (n=2), and others (n=4). For quantitative analysis of 26 studies reporting AI models for AF detection, meta-analyzed sensitivity was 94.80% and specificity was 96.96%. Deep neural networks showed superior performance [meta-analyzed area under receiver operating characteristics curve (AUROC) of 0.981] compared to conventional machine learning algorithms (meta-analyzed AUROC of 0.961). However, AI models tested with proprietary dataset (meta-analyzed AUROC of 0.972) or data acquired from wearable devices (meta-analyzed AUROC of 0.977) showed inferior performance than those with public dataset (meta-analyzed AUROC of 0.986) or data from in-hospital devices (meta-analyzed AUROC of 0.983).

Conclusion: This review found that AI models for diverse cardiovascular-related diseases are being developed, and that they are gradually developing into a form that is suitable for wearable and mobile devices.

Key Words: Electrocardiography, photoplethysmography, artificial intelligence, cardiovascular disease, machine learning, deep learning

Received: October 21, 2021 **Revised:** October 27, 2021
Accepted: October 31, 2021

Co-corresponding authors: Sang Baek Koh, MD, PhD, Department of Preventive Medicine, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju 26426, Korea.
Tel: 82-33-741-0345, Fax: 82-33-747-0409, E-mail: kohhj@yonsei.ac.kr and Sejung Yang, PhD, Department of Biomedical Engineering, Yonsei University, 1 Yonsei-daegil, Wonju 26493, Korea.
Tel: 82-33-760-2459, Fax: 82-33-760-2919, E-mail: syang@yonsei.ac.kr

*Solam Lee and Yuseong Chu contributed equally to this work.
•The authors have no potential conflicts of interest to disclose.

INTRODUCTION

Digital healthcare is a field that deals with personal health and medical information, devices, systems, and platforms, and encompasses the convergence of comprehensive medical services.[1] Although digital healthcare is an umbrella term used for various purposes, in essence, it refers to offering healthcare services for the prevention, diagnosis, treatment, and follow-up management of diseases regardless of time and location by connecting healthcare with information and communication technology.

Traditionally, the field of healthcare has been centered on healthcare institutions and physicians, and the medical infor-

mation they generated was exclusively handled by healthcare workers within the hospital. However, over time, the focus shifted to the prediction and prevention of chronic diseases, rather than treatment-oriented medicine, with the growing interest in healthcare outside the hospital.[2] Against this background, digital healthcare is also developing at a rapid pace due to the remarkable advances in computing power and technologies related to data acquisition and analysis. For instance, the development of various sensors, communication networks, and portable and wearable devices has enabled data acquisition during the day-to-day life of a patient.[3] Based on the acquired data as such, the development of algorithms for diagnosing or predicting specific diseases outside the hospital, so-called medical artificial intelligence (AI), has emerged as a major area of research.

Cardiovascular disease is the most common cause of death in the United States,[4] and it incurs annual medical costs of up to $363 billion.[5] Since most cardiovascular diseases follow a chronic course, lifestyle modifications and the periodic monitoring of diseases are very important even after acute treatment in hospital. Therefore, it is difficult to cope with the burden and death of cardiovascular disease by treatment in hospital alone.

Meanwhile, several AI models for the detection and prediction of arrhythmias, diabetes, and sleep apnea using biosignals, such as those used for electrocardiogram (ECG) or photoplethysmography (PPG), have been reported. A systematic literature review revealed that deep learning models for analyzing ECG showed better performance compared to existing methods, such as the hardware or rule-based algorithms used for the detection or prediction of cardiovascular disease.[6] However, it is difficult to apply any of them to a wearable device. A conventional supine 12-lead ECG or arterial pressure waveform, for instance, requires complicated or invasive procedures.

We postulated that a separate review of AI algorithms which could be applied to wearable devices was necessary to evaluate the applicability of AI for cardiovascular-related diseases in digital healthcare settings. This systematic literature review and meta-analysis aimed to identify the AI models developed for or applicable to wearable devices for cardiovascular-related diseases.

MATERIALS AND METHODS

Search strategy

This systematic review and meta-analysis was performed in accordance with the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020.[7] The searched databases included Medline, Embase, and Cochrane Library. The purpose of the search was to identify the studies related to medical AI used to detect or predict cardiovascular-related disease, which have been developed or applicable to portable and wearable devices. A wide range of keywords were set by referring to previous related reviews[8,9] to include the AI models for various cardiovascular-related diseases (cardi*, heart*, coronary*, angina*, ventric*, myocard*, pericard*, isch(a)em*, arrhythm*, atrial fibrillat*, tachycardi*, endocardi*, stroke, cerebro*, hypertensi*, blood pressure, hyperlipid*, hyperchole*, hyperlipo*, hypertrigly*, arterio*, arthero*, metabolic, deep learning, neural network, artificial intelligence, smart, wearable, portable, 1-lead, single lead, photoplethysmog*, and PPG). The search was performed on September 5, 2021, by one author (S.L), and all publications searched from January 1, 1970 to the present were searched without any limitation on the publication date.

Study selection

The abstracts and titles of the retrieved studies were first screened by two main reviewers (S.L and Y.C). The criteria for eligible studies were as follows: 1) studies whose main tasks related to the detection or prediction of cardiovascular-related disease; 2) studies that developed AI models by utilizing either deep learning or conventional machine learning algorithms; 3) studies that developed AI models for smartphones or using data acquired from mobile and wearable devices; and 4) studies that used biosignals that are considered relatively easy to obtain outside of healthcare institutions, such as 1-lead or 2-lead ECG and PPG, even if those studies did not directly present the keywords associated with wearable device. The exclusion criteria were as follows: 1) studies that simply measured or estimated cardiovascular parameters (e.g., heart rate and blood pressure), regardless of disease status; 2) studies that used invasive signals (e.g., arterial pressure waveform); 3) studies that reported only the models with biosignals that are difficult to apply to wearable devices, such as conventional supine 12-lead ECG; and 4) studies that did not report the quantitative performance of the model.

All studies whose eligibility could not be determined based on the abstract and title alone were included in the full-text evaluation. Studies written in languages other than English or Korean were excluded due to the language proficiency of the authors. Studies with only abstracts in which details on the study settings or design, the nature of the data used were unknown, and the details of the model were unknown were also excluded. In the case of disagreement between the two main reviewers, eligibility was finally decided through discussion with two additional reviewers (J.S.R and S.Y).

Data extraction and quality assessment of study

We extracted important variables using data extraction sheets from the final selected studies. The data sheets were defined prior to conducting the literature searches and study selection. The items extracted from each study included the publication year, author, target disease, task, data availability (public or proprietary), data source (in-hospital device or wearable device), input data domain (ECG, PPG, combined or others),

https://doi.org/10.3349/ymj.2022.63.S93

algorithm type (conventional machine learning or deep neural network), and model performance metrics.

So far, there has been no optimal assessment tool that can evaluate the applicability and risk of bias for AI studies. Therefore, we evaluated the risk of bias of individual studies using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool.[10] However, since the validity of using the scale has not been verified for AI studies, it was not used to determine whether to include or exclude a particular study.

Data synthesis

We intended to synthesize the performance of AI models that could be mounted on wearable devices or mobile devices, and that could be used to predict various cardiovascular-related diseases and present them as summary statistics. However, according to our pilot study, most of the identified models were developed for the detection of atrial fibrillation (AF). Therefore, the quantitative meta-analysis was limited to studies on AF. For those studies, the number of true positives, false positives, true negatives, and false negatives were extracted. If those parameters were not directly presented, they were estimated from the number of subjects (AF and non-AF) and the performance metrics, such as sensitivity and specificity. In order to resolve heterogeneity among studies in the meta-analysis, subgroup analysis was performed according to the algorithm type (conventional machine learning vs. deep neural network), data availability (public vs. proprietary), and data source (in-hospital device vs. wearable device). Only qualitative analysis was performed for studies on diseases other than AF.

Statistical analysis

The number of true positives, false positives, true negatives, and false negatives were used to calculate meta-analyzed sensitivity and specificity. As high heterogeneity between studies was suspected, a random-effects model was used for synthesis. Forest plots for sensitivity and specificity for AF detection was created by using R package meta.[11] Hierarchical summary receiver operating characteristics curves and 95% confidence interval (CI) were estimated by the Reitsma bivariate model[12] using R package mada.[13] All statistical analyses were performed by using R version 4.1.0 (R Statistical Computing).

RESULTS

Study selection

The PRISMA 2020 flow diagram is shown in Fig. 1. After a full text review, 102 studies in total were included in the qualitative review.[14-115] Table 1 presents the summary statistics for the included studies. Table 2 and Supplementary Table 1 (only on-

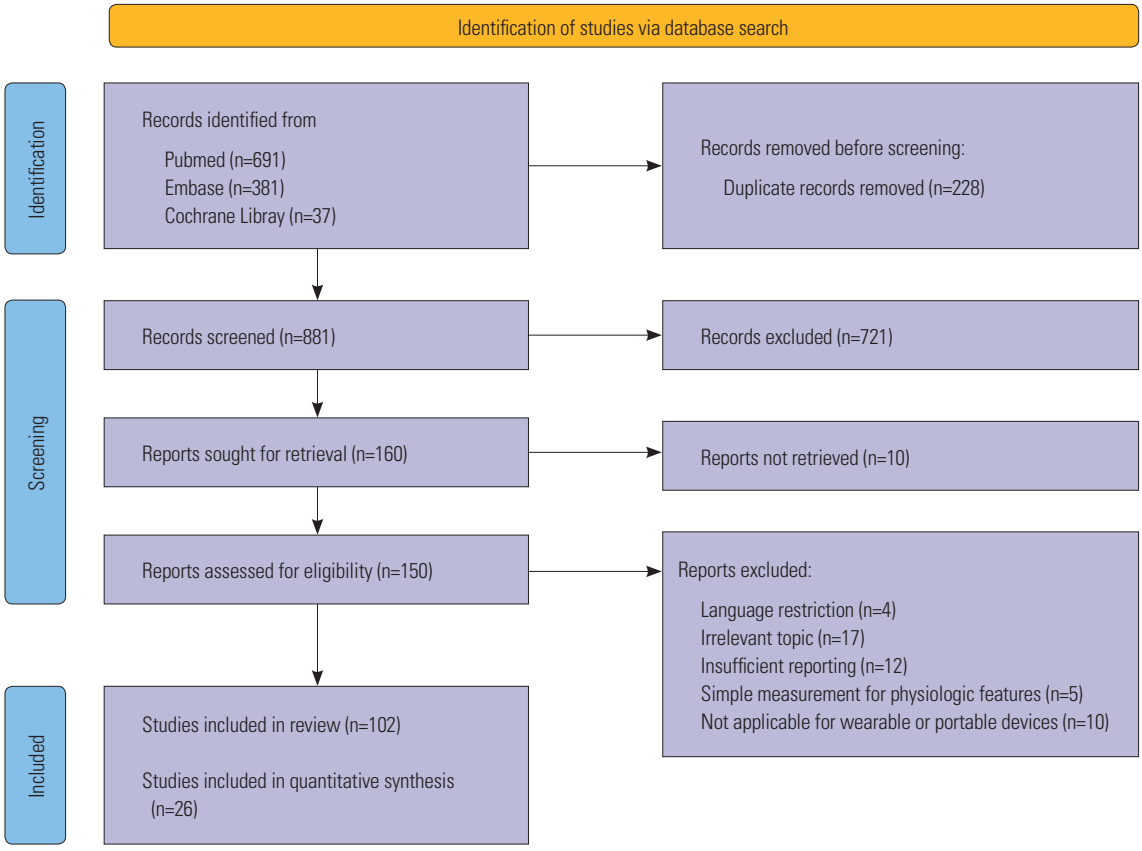

Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 flow diagram for study selection.

Table 1. Summary of the Included Studies

Characteristics	No. of studies
Target disease	
Arrhythmia	62
Sleep apnea	11
Peripheral vascular disease	6
Diabetes mellitus	5
Hyper/hypotension	5
Valvular heart disease	4
Heart failure	3
Critical care	2
Others	4
Algorithm type	
Conventional machine learning alone	34
Deep learning network (alone or combined)	68
Input domain	
Electrocardiography	70
Photoplethysmography	17
Multimodal data	10
Others	5
Data availability	
Public dataset	59
Proprietary dataset (alone or combined)	43
Data source	
In-hospital device alone	55
Wearable device (alone or combined)	47

line) shows the description and findings of the individual studies. Supplementary Fig. 1 (only online) shows the quality assessment for the included studies using the QUADAS-2 tool. As described above, as the validity of the tool used to evaluate the quality of research for development and validation of AI models has not yet been verified, all individual studies were included in the final analysis regardless of the quality assessment results.

Model and dataset characteristics

In 68 studies, deep learning methods, such as neural networks, were used alone or in combination with other machine learning methods (Table 1). Unlike conventional methods that extracted several handcrafted features from ECG or PPG and used them for training and validation of algorithms for disease prediction, methods that used deep learning tended to utilize the signal itself as an input (Fig. 2). For input domain, 1-lead and 2-lead ECG were the most commonly used biosignals, followed by PPG. They were often used as a single input source, but some studies used multimodal data.[23,28,98]

A total of 59 studies used an open dataset, most of which were related to the detection of arrhythmias or sleep apnea. Supplementary Table 2 (only online) summarizes the representative public datasets containing 1-lead or 2-lead ECG or PPG data.[116-137] However, the models for diabetes and other diseases mostly used proprietary datasets that were newly collected or constructed for the study. There were 47 studies that utilized data collected through wearable devices, such as smart watches.

Arrhythmia

The studies related to arrhythmias accounted for the largest proportion, at 62 studies[14-75] (Supplementary Table 1, only online). Most studies had AF detection as the main task. The arrhythmias other than AF, such as premature ventricular contraction, were often grouped and treated as "non-AF" or "other rhythms." A large number of models have been reported based on the rich public dataset, which allowed performance comparisons between the studies. Although many studies have utilized public datasets, such as the MIT-BIH database alone, there were also studies that performed external validation by constructing a proprietary dataset with multi-institutional data.[17,23,26,27]

In total, 26 studies included AF as a detection target, and thus were included in the quantitative meta-analysis. The diagnostic accuracy of AF detection models are presented in Figs. 3 and 4A. The results overall show very high performance with a meta-analyzed sensitivity of 94.80% (95% CI, 91.94%–96.68%) and specificity of 96.96% (95% CI, 94.99%–98.17%). In the subgroup analysis performed to resolve the heterogeneity between studies, however, there was a significant difference in the model performance between studies (Fig. 4B–D). The models developed with deep neural network alone or combined with other algorithms (AUROC of 0.981) showed a superior performance compared to the models developed with conventional machine learning alone (AUROC of 0.961). In contrast, the models tested on the proprietary data (AUROC of 0.972) showed an inferior performance compared to the models tested on the public data (AUROC of 0.986). In addition, the model tested on the data acquired from wearable devices (AUROC of 0.977) showed an inferior performance than the models tested on the data acquired from in-hospital devices (AUROC of 0.983).

With the advent of deep neural networks, the competition for predictive performance has virtually reached saturation, and the development of a lightweight model that can be mounted on a wearable device and operated in real time has recently become a major topic of research.[18,27] There have been studies to determine a critical point that does not compromise the diagnostic accuracy while reducing the data size through compression of the ECG signal itself, in addition to reducing the weight of the model.[32,38,43]

Meanwhile, research using the ECG data collected through a ring-type or patch-type device, rather than data collected in a hospital, has been reported.[40,44] These studies have also shown satisfactory performance in general, but reported slightly lower performance for detecting arrhythmias compared to the models trained with resting ECG-based data obtained in the hospital. One study, in particular, reported difficulty for an algorithm trained on resting ECG data to show proper prediction perfor-

https://doi.org/10.3349/ymj.2022.63.S93

Table 2. Artificial Intelligence Models for Cardiovascular-Related Diseases Other Than Arrhythmia

Study	Task	Dataset	Method	Results
Yu, et al., 2021[76]	Sleep apnea (Screening for sleep apnea)	Apnea-ECG DB	LSTM	Accuracy, 87.09%; Sensitivity, 77.96%; Specificity, 91.74%; F1 score: 0.8161
Chang, et al., 2020[77]	Sleep apnea (Detection of sleep apnea)	Apnea-ECG DB	CNN	Accuracy, 97.1%; Sensitivity, 95.7%; Specificity, 100%
Iwasaki, et al., 2021[78]	Sleep apnea (Screening of sleep apnea)	Proprietary (1-lead ECG of 24 patients)	LSTM	Sensitivity, 100%; Specificity, 100%
Papini, et al., 2020[79]	Sleep apnea (Estimation of the apnea-hypopnea index)	SOMNIA and HealthBed DB	CNN	AUC, 0.80; Accuracy, 85%; Sensitivity, 38%; Specificity, 94%
Wang, et al., 2019[80]	Sleep apnea (Detection of sleep apnea)	Apnea-ECG DB	Time window with a neural network	Per segment: AUC, 0.945; Accuracy, 87.3%; Sensitivity, 85.1%; Specificity, 88.7% Per recording: AUC, 1.000; Accuracy, 97.1%; Sensitivity, 100.0%; Specificity, 91.7%
Bozkurt, et al., 2019[81]	Sleep apnea (Determination of respiratory arrests)	Proprietary DB (2358 PPG)	SVM	Accuracy, 87.36%; Sensitivity, 86%; Specificity, 88%
Wang, et al., 2019[82]	Sleep apnea (Detection of sleep apnea)	Apnea-ECG DB and UCDODB	CNN	AUC, 0.950; Accuracy, 87.6%; Sensitivity, 83.1%; Specificity, 90.3%
Lin, et al., 2018[83]	Sleep apnea (Detection of obstructive sleep apnea)	Apnea-ECG DB	ANN	Accuracy, 79%; Sensitivity, 90%; Specificity, 73%
Urtnasan et al., 2018[84]	Sleep apnea (Detection of obstructive sleep apnea)	Proprietary DB (1-lead ECG of 82 persons)	CNN	Sensitivity, 96%; F1 score: 0.96
Sharma, et al., 2016[85]	Sleep apnea (Detection of sleep apnea)	Apnea-ECG DB	SVM	AUC, 0.978; Accuracy, 97.14%; Sensitivity 95.8%, Specificity, 100%
Babaeizadeh, et al., 2010[86]	Sleep apnea (Detection of sleep apnea)	Apnea-ECG DB	Quadratic classifier	Accuracy, 84.7%; Sensitivity, 76.7%; Specificity, 89.6%
Allen, et al., 2021[87]	Peripheral vascular disease (Detection of peripheral arterial disease)	Proprietary DB (214 PPG)	CNN	Accuracy, 88.9%; Sensitivity, 86.6%; Specificity, 90.2%
Lee, et al., 2020[88]	Peripheral vascular disease (Prediction of ankle brachial index)	MIMIC III	LSTM	Accuracy, 98.34%; Sensitivity, 97.14%; F1-score: 0.9743
Dall'Olio et al., 2020[89]	Peripheral vascular disease (Prediction of vascular aging)	Heart for Heart	CNN	AUC, 0.953
Alty, et al., 2007[90]	Peripheral vascular disease (Prediction of arterial stiffness)	Proprietary DB (461 PPG)	SVM	Accuracy, 86.1%; Sensitivity, 86.7%; Specificity, 85.3%
Allen and Murray 1996[91]	Peripheral vascular disease (Arterial pulse waveform classification)	Proprietary DB (366 PPG)	ANN	Accuracy, 80%; Sensitivity, 92%; Specificity, 63%
Allen and Murray 1993[92]	Peripheral vascular disease (Classification of peripeheral vascular disease of the lower limb artieries)	Proprietary DB (150 PPG)	ANN	Accuracy, 90%; Sensitivity, 93%; Specificity, 85%
Baig, et al., 2021[93]	Diabetes mellitus (Early detection of prediabetes and type 2 diabetes mellitus)	Proprietary DB (Demographics, vital signs, activity data, ECG, and others)	Fuzzy inference system	Accuracy, 91%; Sensitivity, 94%; Specificity, 90%
Avram, et al., 2020[94]	Diabetes mellitus (Detection of diabetes)	Proprietary DB (2589448 PPG)	CNN	Primary cohort: AUC, 0.766 (95% CI, 0.750–0.782); Sensitivity, 75%; Specificity, 65%, Contemporary cohort: AUC, 0.740 (95% CI, 0.723–0.758); Sensitivity, 81%; Specificity, 54%
Porumb, et al., 2020[95]	Diabetes mellitus (Detection of nocturnal low glucose)	Proprietary DB (1-lead ECG of 25 persons)	CNN	AUC, 0.907; Accuracy, 92.8%; Sensitivity, 91.6%; Specificity, 89.9%
Porumb, et al., 2020[96]	Diabetes mellitus (Detection of hypoglycemic events)	Proprietary DB (ECG of 4 persons)	CNN and RNN	5-min prediction: Accuracy 87.7%; Sensitivity, 88.3%; Specificity, 88.5% 10-min prediction: Accuracy, 90.0%; Sensitivity, 87.4%; Specificity, 92.2%

Table 2. Artificial Intelligence Models for Cardiovascular-Related Diseases Other Than Arrhythmia (continued)

Study	Task	Dataset	Method	Results
Faruqui, et al., 2019[97]	Diabetes mellitus (Forecasting daily glucose levels)	Proprietary DB (Daily monitoring of diet, physical activity, weight, and blood glucose over 6 months of 10 patients)	LSTM	Accuracy of 64.837% for ±10% range of the actual glucose level value
Lee, et al., 2021[98]	Hyper/hypotension (Prediction of intraoperative hypotension)	The VitalDB	CNN	AUC, 0.931 (95% CI, 0.929–0.934); Sensitivity, 85.6% (95% CI, 85.3%–86.0%); Specificity, 85.6% (95% CI, 85.3%–85.9%)
Kwon, et al., 2020[99]	Hyper/hypotension (Detection of pulmonary hypertension)	Proprietary DB (70709 1-lead ECG)	CNN	Internal validation: AUC, 0.859 (95% CI, 0.855–0.863); Accuracy, 76.4% (95% CI, 76.1%–76.8%); Sensitivity, 80.0% (95% CI, 79.6%–80.3%); Specificity, 74.7% (95% CI, 74.4%–75.0%) External validation: AUC, 0.902 (95% CI, 0.900–0.905); Accuracy, 84.0% (95% CI, 83.7%–84.3%); Sensitivity, 80.0% (95% CI, 79.7%–80.2%); Specificity, 84.3% (95% CI, 84.0%–84.6%)
Devaki, et al., 2020[100]	Hyper/hypotension (Diagnosis of hypertension)	Proprietary (PPG of 140 subjects)	CNN	Accuracy, 83.3%; Sensitivity, 100%; Specificity, 75%
Naifisi, et al., 2018[101]	Hyper/hypotension (Identification of hypotension-related episodes)	Proprietary DB (781 PPG of 10 patients)	AdaBoost	Accuracy, 94.5%; Sensitivity, 91.7%; Specificity, 95.8%
Liang, et al., 2018[102]	Hyper/hypotension (Hypertension risk stratification)	MIMIC II and MIMIC III	CNN	F1 score of Normal vs. prehypertension: 0.8052; F1 score of Normal vs. hypertension: 0.9255; F1 score of Normal+prehypertension vs. hypertension: 0.8295
Kwon, et al., 2020[103]	Valvular heart disease (Detection of mitral regurgitation)	Proprietary DB (70529 1-lead ECG)	CNN	Internal validation: AUC, 0.758 (95% CI, 0.753–0.762); Accuracy, 52.6% (95% CI, 51.2%–53.7%); Sensitivity, 90.0% (95% CI, 89.6%–90.3%); Specificity, 40.8% (95% CI, 39.6%–41.9%) External validation: AUC, 0.850 (95% CI, 0.842–0.857); Accuracy, 57.3% (95% CI, 56.1%–59.2%); Sensitivity, 90.1% (95% CI, 89.5%–90.5%); Specificity, 56.0% (95% CI, 54.9%–57.2%)
Yang, et al., 2020[104]	Valvular heart disease (Detection of aortic stenosis)	Proprietary DB (Seismocardiogram and gyrocardiogram of 21 patients)	CNN	Accuracy, 95%; Sensitivity, 94%
Yang, et al., 2020[105]	Valvular heart disease (Detection of aortic stenosis)	Proprietary DB (Seismocardiogram and gyrocardiogram of 21 patients)	Random forest	Accuracy, 98.96%; Sensitivity, 98.33%; Specificity, 99.58%
Kwon, et al., 2020[106]	Valvular heart disease (Detection of aortic stenosis)	Proprietary DB (56689 1-lead ECG)	CNN	Interval validation: AUC, 0.845 (95% CI, 0.841–0.848) External validation: AUC, 0.821 (95% CI, 0.816–0.825)
Cho, et al., 2021[109]	Heart failure (Detection of heart failure with reduced ejection fraction)	Proprietary DB (47203 1-lead ECG)	CNN	Internal validation: AUC, 0.874 (95% CI, 0.859–0.890); Accuracy, 67.1% (95% CI, 65.5%–68.6%); Sensitivity, 93.2% (95% CI, 90.9%–95.6%); Specificity, 63.2% (95% CI, 61.5%–65.0%)

https://doi.org/10.3349/ymj.2022.63.S93

Table 2. Artificial Intelligence Models for Cardiovascular-Related Diseases Other Than Arrhythmia (continued)

Study	Task	Dataset	Method	Results
				External validation: AUC, 0.929 (95% CI, 0.911–0.946); Accuracy, 82.5% (95% CI, 81.4%–83.7%); Sensitivity, 92.1% (95% CI, 88.0%–95.8%); Specificity, 82.1% (95% CI, 80.9%–83.2%)
Ahmedov, and Amirjanov, 2021[107]	Heart failure (Measurement of a cardiac stroke volume)	Proprietary DB (Blood pressure, heart performance measured by ballistocardiographic sensor, skin warming time of 92 persons)	Fuzzy model	Correlation r: 0.803; Mean square error: 8.185
Wang and Zhou, 2019[108]	Heart failure (Detection of congestive heart failure)	BIDMC-CHF, CHF-RR, MITNSRDB, FD, and NSR-RR	LSTM	Accuracy, 82.51%–99.22%
Rashid and Al Faruque, 2020[110]	Critical care (Detection of myocardial infarction)	PTB diagnostic ECG DB	Binarized neural network	Accuracy, 90.29%; Sensitivity, 90.41%; Specificity, 90.16%
Kwon, et al. 2020[111]	Critical care (Detection of cardiac arrest)	Proprietary DB (47505 1-lead ECG)	CNN	Internal validation: AUC, 0.887 (95% CI, 0.846–0.929); Sensitivity, 85.7% (95% CI, 75.9%–92.6%); Specificity, 78.1% (95% CI, 76.9%–79.4%) External validation: AUC, 0.921 (95% CI, 0.899–0.998); Sensitivity, 82.2% (95% CI, 81.5%–83.0%); Specificity, 82.2% (95% CI, 81.5%–83.0%)
Grogan, et al., 2021[112]	Others (Detection of cardiac amyloidosis)	Proprietary DB (4995 1-lead ECG)	CNN	AUC, 0.86
Kwon, et al., 2020[113]	Others (Detection of anemia)	Proprietary DB (70074 1-lead ECG)	CNN	Internal validation: AUC, 0.870 (95% CI, 0.853–0.887); Sensitivity, 87.8% (95% CI, 84.1%–90.8%); Specificity 68.0% (95% CI, 67.0%–69.1%) External validation: AUC 0.841 (95% CI, 0.815–0.866); Sensitivity, 88.7% (95% CI, 83.3%–92.8%); Specificity 65.4% (95% CI, 64.0%–66.8%)
Chiarelli, et al., 2019[114]	Others (Prediction of cardiovascular age)	Proprietary DB (2400 1-lead ECG + PPG)	CNN	Correlation r, 0.92; Mean square error, 7 years
Fan, et al., 2019[115]	Others (Prediction of 1-day-forward self-reported wellness)	Proprietary DB (1-lead ECG of 11 persons)	Bidirectional LSTM	Accuracy, 93.21%; Sensitivity, 92.51%; F1 score: 91.98%

ECG, electrocardiography; PPG, photoplethysmography; ANN, artificial neural network; AUC, area under the curve; CNN, convolutional neural network; LSTM, long short-term memory; SVM, support vector machine; RNN, recurrent neural network.
Refer to Supplementary Table 1 (only online) for artificial intelligence models for arrhythmia detection. Refer to Supplementary Table 2 (only online) for dataset abbreviations and description.

mance when applied to ambulatory ECG data.[64] In addition, a model that could detect AF in a large number of patients using a video of the facial region rather than ECG monitoring was reported, confirming that diverse types of input data could be utilized.[36]

Other conditions

There were few studies on the screening and detection of diseases such as sleep apnea,[76-86] peripheral vascular diseases,[87-92] diabetes,[93-97] hyper/hypotensive disease,[98-102] valvular disease,[103-106] heart failure,[107-109] myocardial infarction,[110] cardiac arrest,[111] and other conditions, including cardiac amyloidosis and anemia (Table 2).[112-115]

The most common study following arrhythmia was the detection of sleep apnea. Sleep apnea itself is not a cardiovascular disease, but obstructive sleep apnea is a major detection target for prediction models for cardiovascular-related disease that use wearable devices, as it is associated with or increases the risk of major cardiovascular diseases. The public dataset PhysioNet Apnea-ECG[116] has been used in numerous studies. There have been several studies that used raw ECG with deep learning algorithms, but machine learning models using hand-

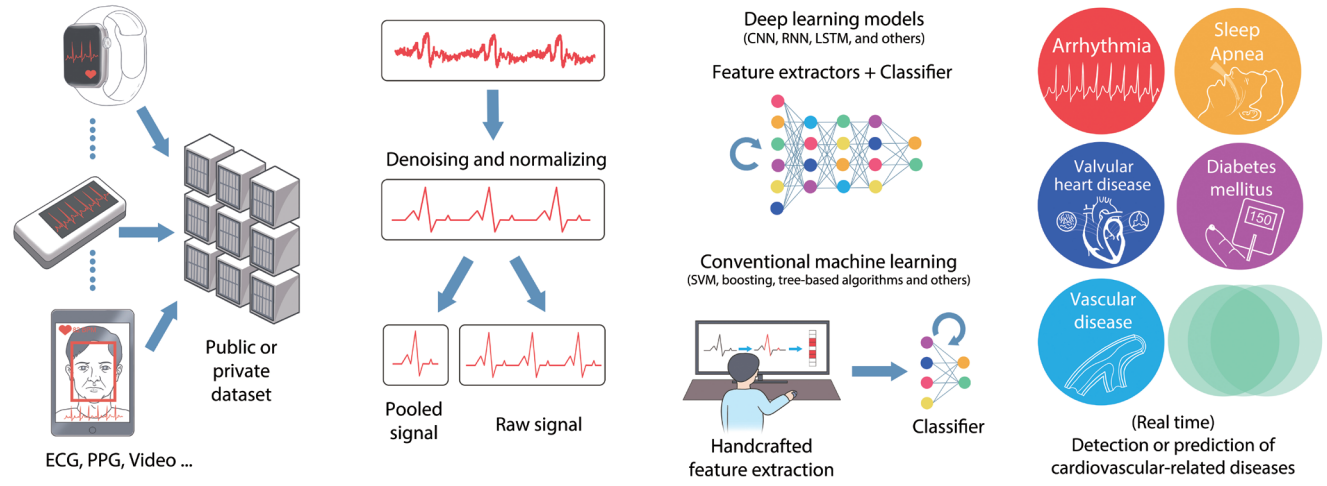

Fig. 2. Schematic illustration for wearable device-based artificial intelligence for cardiovascular-related diseases. ECG, electrocardiography; PPG, photoplethysmography; CNN, convolutional neural network; RNN, recurrent neural network; LSTM, long short-term memory.

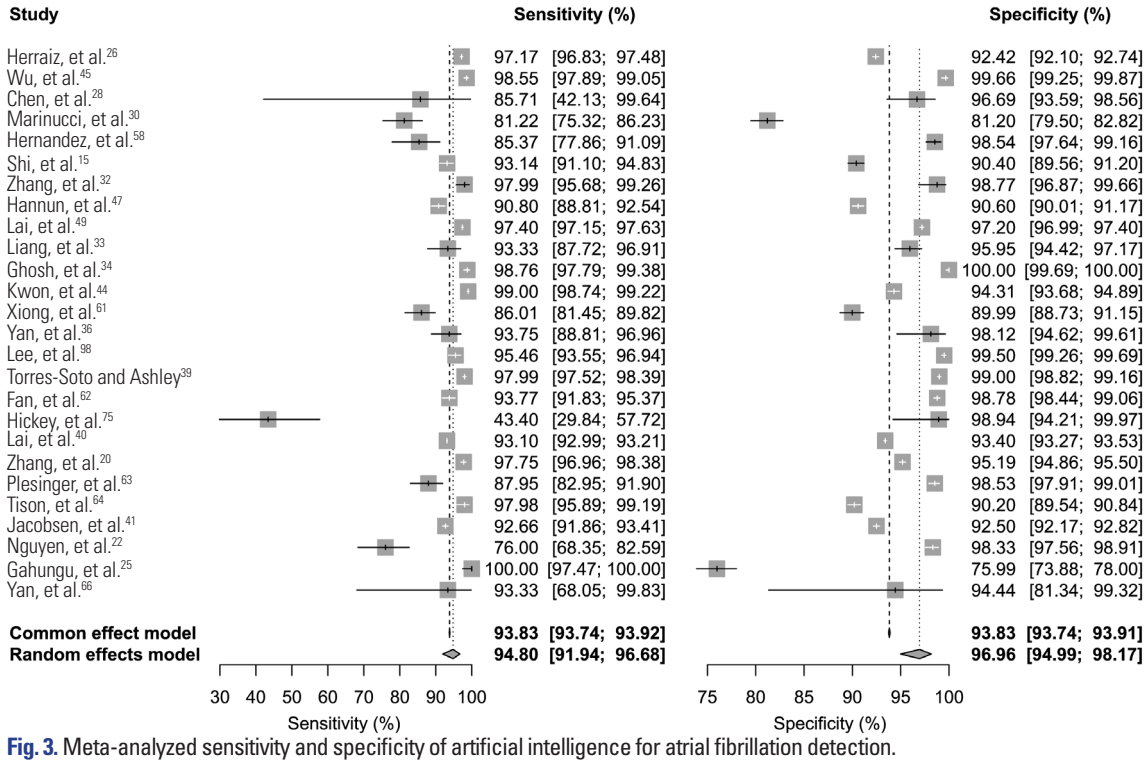

Fig. 3. Meta-analyzed sensitivity and specificity of artificial intelligence for atrial fibrillation detection.

crafted features, such as heart rate variability, have also been reported. In addition to ECG, there have also been studies that use PPG, many of which were studies that used proprietary datasets. There was a study to predict the apnea-associated index using wrist-worn reflective PPG obtained directly using a wearable device.[79] This particular study investigated the changes in model performance due to limbic movement, and reported that false positive detection had an adverse effect on the specificity and positive predictive value of the model.

These studies mainly used proprietary datasets, as there were very few public datasets available. As expected, they showed very heterogeneous performance, depending on the task. For instance, the models showed very high predictive power for valvular diseases, such as aortic stenosis, which is known to accompany some changes in the ECG. However, the models did not show an equally high performance for the detection of diabetes or anemia that were not accompanied by substantial changes in the ECG. In a study that developed an AI model for heart failure, the algorithm was tested on several populations with a varying prevalence of heart failure, confirming a very

https://doi.org/10.3349/ymj.2022.63.S93

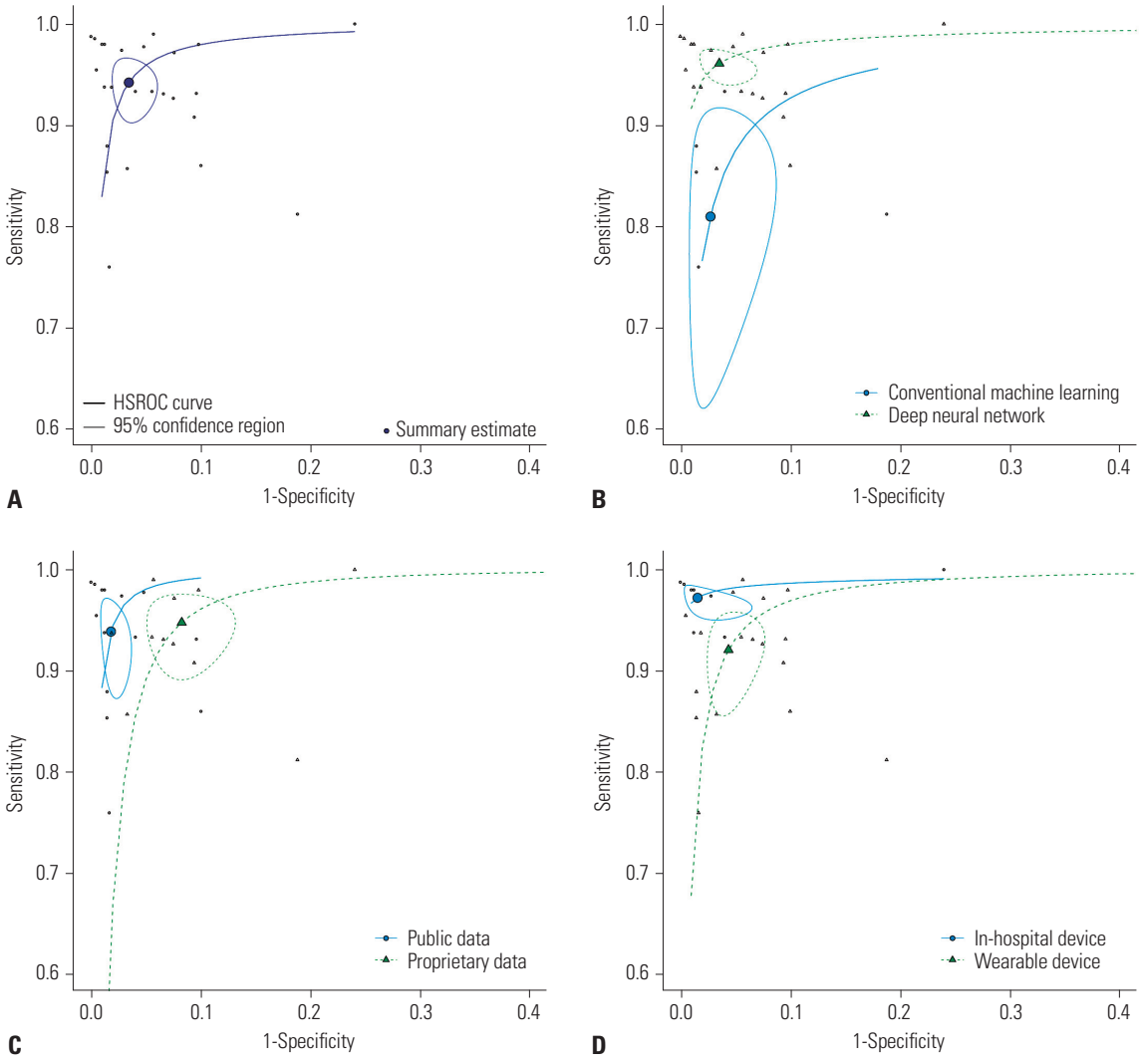

Fig. 4. Hierarchical summary of receiver operating characteristics curves of artificial intelligence for atrial fibrillation detection. (A) All studies. (B) Studies with conventional machine learning vs. studies with deep neural networks. (C) Studies tested with public dataset vs. studies tested with proprietary dataset. (D) Studies tested with data acquired from in-hospital devices vs. studies tested with data acquired from wearable devices. HSROC, hierarchical summary receiver operating characteristics.

large difference in the accuracy of the algorithm from 82.51% to 99.22%, depending on the prevalence of heart failure in the study population.[108]

For diseases other than arrhythmia, it was difficult to compare the performance of models due to the small number of studies for each disease, as well as the use of data from heterogeneous settings. Considering the low availability of public data and the characteristics of studies with individual datasets, external validation was rarely performed.

DISCUSSION

This systematic review and meta-analysis summarized the AI models that were developed for or that are potentially useful for wearable devices, and that could be used to detect or predict cardiovascular-related diseases. In addition to the models for arrhythmia detection, including AF, studies have reported on models to predict various cardiovascular-related diseases such as sleep apnea, diabetes, valvular disease, and anemia.

Fig. 2 summarizes the development and application of an AI model for the detection and prediction of cardiovascular-related diseases. The first step is data collection. The most commonly used biosignal for monitoring using a wearable device is the 1-lead or 2-lead ECG signal, followed by PPG. Owing to a large database, such as the MIT-BIH arrhythmia database,[122,123] various detection models for arrhythmias, including AF, have been reported. As these were hospital-based data, there was a limit to using the data in implementing an algorithm for wearable devices. Fortunately, the availability of biosignals that were collected using mobile devices has increased, as databases such as the PhysioNet Computing in Cardiology Challenge[117] have been made public. However, there is a clear need for ambulatory biosignals, since the performance of an algorithm devel-

oped from the data of a resting setting is quite poor in an ambulatory setting.[64] Recently, in addition to Holter monitoring, which has been traditionally used, the use of patch-type, watch-type, ring-type, and clothing-type monitoring systems have been approved by the FDA.[138] Accordingly, the availability of ambulatory biosignals is expected to increase significantly in the future.

One thing to note in relation to the data is the difference in systematic performance between the models that use public datasets, proprietary datasets, data acquired from in-hospital devices, and data acquired from wearable devices. Data characteristics may vary significantly depending on the methods (resting vs. ambulatory) and settings (intensive care unit, inpatient, outpatient, and general population) of data acquisition. For example, the prevalence of cardiovascular diseases varies greatly depending on the population, and the class imbalance between normal data and disease data may have a significant impact on the model performance.[108] Therefore, it is very important in model development to use the collected from a population group that has characteristics similar to the target group to which the developed AI model will be applied. In order to develop a more universal model from the perspective of digital healthcare, it is preferable to use data collected from large general populations rather than data collected from hospitals.

The collected data are transformed into a form more suitable for model learning through preprocessing, such as denoising or normalization. For example, the length of the input signal could be changed or converted to an average signal, depending on the type or design of the AI model. Various machine learning methods have been used to predict cardiovascular disease using handmade features such as heart rate variability, R-R interval, and QRS amplitude extracted from ECG or PPG as general input values. However, the popularization of deep neural networks has enabled raw signals themselves to be used as input data. Since biosignal data has a time-series characteristic and is not limited in the length of an input signal, recurrent neural network (RNN) or long short-term memory is logically the most preferred model. However, there have been a considerable number of models that used convolutional neural network (CNN) alone or CNN combined with RNN-based algorithms. In addition to signals, facial imaging videos have been used in a model for detecting arrhythmias.[36] Those developed with deep neural networks generally perform better than traditional models, suggesting the possibility that AI can extract features from imaging as well as signal data that can be used for predicting disease states that humans cannot visualize. Although current studies have often used a single biosignal as an analysis domain, it would be possible to develop a multimodal model that utilizes various signals, imaging data, and other clinical information together.

Nevertheless, the most significant problem with deep learning is that it has a black box-like characteristic that makes understanding its operation and judgment principle difficult. Despite the high accuracy of deep learning models, the decisions cannot be accepted by healthcare workers without proper interpretation. While a small number of studies have adopted a strategy that uses salient maps to improve interpretability,[98,113] such attempts are still difficult to find compared to the models in the field of medical imaging. Extracting interpretable handcrafted features from clinical information or biosignals that are familiar to human experts and using them together in deep learning models could be a feasible alternative to improving the interpretability of AI.

Unlike medical imaging, AI models for cardiovascular-related disease are preferred to be used with wearable and portable devices carried by individuals in their day-to-day lives. It is difficult to operate a large-scale model on portable devices due to hardware limitations. For the analysis of lifelogs collected in real time, the operating speed is also a major consideration. Therefore, it is essential to develop a lightweight and fast model, even if there is a slight loss in performance.[18,27,32,38,43] Second, the algorithm is operated in an environment that is not well-controlled, unlike a hospital environment. As mentioned earlier, a model that is developed using resting biosignals as a source rarely works well for ambulatory biosignals.[64] In addition, even resting ECG signals may be more prone to artifacts or noise if they are acquired in daily life. If too many false positive alarms for these abnormal signals occur, the user may prefer to stop using the wearable device. Therefore, proper data preprocessing and thresholding are required to detect a life-threatening condition without tiring the user.

Wearable device-enabled detection of cardiovascular-related diseases are likely to become more common as healthcare technology expands.[139] It is estimated that more than 50 million people in the United States use wearable device to record their daily activities.[140] These lifelogs could lead to the early detection of diverse cardiovascular-related diseases and potentially life-threatening conditions. However, although patient-activated daily monitoring has the potential benefit to improve the detection of subclinical or occult diseases, there are limitations of their widespread use. For example, poor signal quality and false alarms can lead to inappropriate interpretation, resulting in unnecessary medical referrals and testing.[141] It is essential to understand the limitations of consumer-based technologies to avoid improper dependence on the diagnosis and treatment of cardiovascular diseases.[142] Even while these devices may help diagnose cardiovascular-related diseases, it remains to be seen how they can be optimally incorporated in current healthcare practices to improve patient outcomes.[139]

One limitation of this systematic review and meta-analysis lies in the high heterogeneity of the studies included in the analysis. For instance, some studies used ECG data obtained from an intensive care unit, whereas some studies directly utilized data obtained with a smart watch. Although a subgroup analysis was performed to resolve the heterogeneity, nevertheless, inappropriate statistical estimates may have been produced in the

https://doi.org/10.3349/ymj.2022.63.S93

process of quantitatively synthesizing studies that were conducted with data that were collected using different modalities from different population groups. In addition, for diseases and tasks other than AF, meta-analysis could not be performed, as there were too few studies to be quantitatively synthesized.

In conclusion, this systematic review and meta-analysis revealed that AI models for the diagnosis and prediction of various cardiovascular-related diseases as well as arrhythmias are being developed, and that they are gradually developing into a form that is suitable for wearable and mobile devices. Numerous studies have demonstrated that the deep learning algorithm shows very high performance compared to the existing analysis methods that use human visualization or the extraction of handmade features for biosignals, such as ECG or PPG signals. However, there must still be sufficient consideration of various aspects, such as the data acquisition process, characteristics of the acquired data, characteristics of the population to which the algorithm is applied, weight reduction of the algorithm, working principle, and interpretability of the model, to develop a practical medical AI model that can be used in the real world.

ACKNOWLEDGEMENTS

This research was supported by a grant from the MD-PhD/Medical Scientist Training Program through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Korea.

AUTHOR CONTRIBUTIONS

Conceptualization: Sejung Yang and Sang Baek Koh. Data curation: Solam Lee, Yuseong Chu, and Jiseung Ryu. Formal analysis: Solam Lee and Yuseong Chu. Funding acquisition: Solam Lee and Sang Baek Koh. Investigation: Solam Lee and Yuseong Chu. Methodology: Solam Lee and Yuseong Chu. Project administration: Sejung Yang and Sang Baek Koh. Resources: Young Jun Park, Sejung Yang, and Sang Baek Koh. Software: Solam Lee and Yuseong Chu. Supervision: Young Jun Park, Sejung Yang, and Sang Baek Koh. Validation: Young Jun Park, Sejung Yang, and Sang Baek Koh. Visualization: Solam Lee and Yuseong Chu. Writing—original draft: Solam Lee and Yuseong Chu. Writing—review & editing: all authors. Approval of final manuscript: all authors.

ORCID iDs

Solam Lee https://orcid.org/0000-0001-6458-9449
Yuseong Chu https://orcid.org/0000-0003-0930-4628
Jiseung Ryu https://orcid.org/0000-0001-6421-0160
Young Jun Park https://orcid.org/0000-0002-1461-5597
Sejung Yang https://orcid.org/0000-0002-5841-851X
Sang Baek Koh https://orcid.org/0000-0001-5609-6521

REFERENCES

1. Turakhia MP, Desai SA, Harrington RA. The outlook of digital health for cardiovascular medicine: challenges but also extraordinary opportunities. JAMA Cardiol 2016;1:743-4.
2. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017;3:38.
3. Duncker D, Ding WY, Etheridge S, Noseworthy PA, Veltmann C, Yao X, et al. Smart wearables for cardiac monitoring-real-world use beyond atrial fibrillation. Sensors (Basel) 2021;21:2539.
4. Centers for Disease Control and Prevention. Underlying Cause of Death, 1999-2019. [Internet] [accessed on 2021 October 1]. Available at: https://wonder.cdc.gov/ucd-icd10.html.
5. Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, et al. Heart disease and stroke statistics-2021 update: a report from the American Heart Association. Circulation 2021;143:e254-743.
6. Krittanawong C, Johnson KW, Rosenson RS, Wang Z, Aydar M, Baber U, et al. Deep learning for cardiovascular medicine: a practical primer. Eur Heart J 2019;40:2058-73.
7. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
8. Horjus DL, Oudman I, van Montfrans GA, Brewster LM. Creatine and creatine analogues in hypertension and cardiovascular disease. Cochrane Database Syst Rev 2011;2011:CD005184.
9. Taylor F, Huffman MD, Macedo AF, Moore TH, Burke M, Davey Smith G, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013;2013:CD004816.
10. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.
11. Schwarzer G. Package 'meta.' [Internet] [accessed on 2021 October 1]. Available at: https://mirror-hk.koddos.net/CRAN/web/packages/meta/meta.pdf.
12. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58:982-90.
13. Doebler P. Package 'mada.' [Internet] [accessed on 2021 October 1]. Available at: http://r.meteo.uni.wroc.pl/web/packages/mada/mada.pdf.
14. Huang Y, Li H, Yu X. A multiview feature fusion model for heartbeat classification. Physiol Meas 2021;42:065003.
15. Shi J, Chen C, Liu H, Wang Y, Shu M, Zhu Q. Automated atrial fibrillation detection based on feature fusion using discriminant canonical correlation analysis. Comput Math Methods Med 2021;2021:6691177.
16. Yu J, Wang X, Chen X, Guo J. Automatic premature ventricular contraction detection using deep metric learning and KNN. Biosensors (Basel) 2021;11:69.
17. Jo YY, Cho Y, Lee SY, Kwon JM, Kim KH, Jeon KH, et al. Explainable artificial intelligence to detect atrial fibrillation using electrocardiogram. Int J Cardiol 2021;328:104-10.
18. Lu P, Gao Y, Xi H, Zhang Y, Gao C, Zhou B, et al. KecNet: a light neural network for arrhythmia classification based on knowledge reinforcement. J Healthc Eng 2021;2021:6684954.
19. Lee H, Shin M. Learning explainable time-morphology patterns for automatic arrhythmia classification from short single-lead ECGs. Sensors (Basel) 2021;21:4331.
20. Zhang X, Li J, Cai Z, Zhang L, Chen Z, Liu C. Over-fitting suppression training strategies for deep learning-based atrial fibrillation detection. Med Biol Eng Comput 2021;59:165-73.
21. Yang J, Cai W, Wang M. Premature beats detection based on a novel convolutional neural network. Physiol Meas 2021;42:075003.

22. Nguyen QH, Nguyen BP, Nguyen TB, Do TT, Mbinta JF, Simpson CR. Stacking segment-based CNN with SVM for recognition of atrial fibrillation from single-lead ECG recordings. Biomed Signal Process Control 2021;68:102672.
23. Liu X, Liu T, Zhang Z, Kuo PC, Xu H, Yang Z, et al. TOP-Net prediction model using bidirectional long short-term memory and medical-grade wearable multisensor system for tachycardia onset: algorithm development study. JMIR Med Inform 2021;9:e18803.
24. Tan L, Yu K, Bashir AK, Cheng X, Ming F, Zhao L, et al. Toward real-time and efficient cardiovascular monitoring for COVID-19 patients by 5G-enabled wearable medical devices: a deep learning approach. Neural Comput Appl 2021 Jul 4. [Epub]. Available at: https://doi.org/10.1007/s00521-021-06219-9.
25. Gahungu N, Shariar A, Playford D, Judkins C, Gabbay E. Transfer learning artificial intelligence for automated detection of atrial fibrillation in patients undergoing evaluation for suspected obstructive sleep apnoea: a feasibility study. Sleep Med 2021;85:166-71.
26. Herraiz ÁH, Martínez-Rodrigo A, Bertomeu-González V, Quesada A, Rieta JJ, Alcaraz R. A deep learning approach for featureless robust quality assessment of intermittent atrial fibrillation recordings from portable and wearable devices. Entropy (Basel) 2020;22:733.
27. Jeon E, Oh K, Kwon S, Son H, Yun Y, Jung ES, et al. A lightweight deep learning model for fast electrocardiographic beats classification with a wearable cardiac monitor: development and validation study. JMIR Med Inform 2020;8:e17037.
28. Chen E, Jiang J, Su R, Gao M, Zhu S, Zhou J, et al. A new smart wristband equipped with an artificial intelligence algorithm to detect atrial fibrillation. Heart Rhythm 2020;17:847-53.
29. Gao Y, Wang H, Liu Z. A novel approach for atrial fibrillation signal identification based on temporal attention mechanism. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:316-9.
30. Marinucci D, Sbrollini A, Marcantoni I, Morettini M, Swenne CA, Burattini L. Artificial neural network for atrial fibrillation identification in portable devices. Sensors (Basel) 2020;20:3570.
31. Aschbacher K, Yilmaz D, Kerem Y, Crawford S, Benaron D, Liu J, et al. Atrial fibrillation detection from raw photoplethysmography waveforms: a deep learning application. Heart Rhythm O2 2020;1:3-9.
32. Zhang H, Dong Z, Gao J, Lu P, Wang Z. Automatic screening method for atrial fibrillation based on lossy compression of the electrocardiogram signal. Physiol Meas 2020;41:075005.
33. Liang Y, Yin S, Tang Q, Zheng Z, Elgendi M, Chen Z. Deep learning algorithm classifies heartbeat events based on electrocardiogram signals. Front Physiol 2020;11:569050.
34. Ghosh SK, Tripathy RK, Paternina MRA, Arrieta JJ, Zamora-Mendez A, Naik GR. Detection of atrial fibrillation from single lead ECG signal using multirate cosine filter bank and deep neural network. J Med Syst 2020;44:114.
35. Li Z, Zhou D, Wan L, Li J, Mou W. Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram. J Electrocardiol 2020;58:105-12.
36. Yan BP, Lai WHS, Chan CKY, Au ACK, Freedman B, Poh YC, et al. High-throughput, contact-free detection of atrial fibrillation from video with deep learning. JAMA Cardiol 2020;5:105-7.
37. Vijayarangan S, Murugesan B, Vignesh R, Preejith SP, Joseph J, Sivaprakasam M. Interpreting deep neural networks for single-lead ECG arrhythmia classification. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:300-3.
38. Cheng Y, Ye Y, Hou M, He W, Pan T. Multi-label arrhythmia classification from fixed-length compressed ECG segments in real-time wearable ECG monitoring. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:580-3.
39. Torres-Soto J, Ashley EA. Multi-task deep learning for cardiac rhythm detection in wearable devices. NPJ Digit Med 2020;3:116.
40. Lai D, Bu Y, Su Y, Zhang X, Ma CS. Non-standardized patch-based ECG lead together with deep learning based algorithm for automatic screening of atrial fibrillation. IEEE J Biomed Health Inform 2020;24:1569-78.
41. Jacobsen M, Dembek TA, Ziakos AP, Gholamipoor R, Kobbe G, Kollmann M, et al. Reliable detection of atrial fibrillation with a medical wearable during inpatient conditions. Sensors (Basel) 2020;20:5517.
42. Lennox C, Mahmud MS. Robust classification of cardiac arrhythmia using a deep neural network. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:288-91.
43. Abdelazez M, Rajan S, Chan ADC. Transfer learning for detection of atrial fibrillation in deterministic compressive sensed ECG. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5398-401.
44. Kwon S, Hong J, Choi EK, Jeong ER, Lee B, Lee E, et al. Diagnostic performance of a wearable ring-type device with deep learning analysis of photoplethysmography for detecting atrial fibrillation (Abstract). Arrhythmia 2019;35:104 AP19-00184.
45. Wu Z, Feng X, Yang C. A deep learning method to detect atrial fibrillation based on continuous wavelet transform. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:1908-12.
46. Kwon S, Hong J, Choi EK, Yi Y, Jeong ER, Lee B, et al. A novel ring-type werable device equipped with deep learning algorithm using photoplesthymographic signals for detecting atrial fibrillation (Abstract). Heart Rhythm 2019;16(5 Suppl):S578 S-IA01-013.
47. 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;25:65-9.
48. Hong S, Zhou Y, Wu M, Shang J, Wang Q, Li H, et al. Combining deep neural networks and engineered features for cardiac arrhythmia detection from ECG recordings. Physiol Meas 2019;40:054009.
49. Lai D, Zhang X, Zhang Y, Bin Heyat MB. Convolutional neural network based detection of atrial fibrillation combing R-R intervals and F-wave frequency spectrum. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:4897-900.
50. Kwon S, Hong J, Choi EK, Lee E, Hostallero DE, Kang WJ, et al. Deep learning approaches to detect atrial fibrillation using photoplethysmographic signals: algorithms development study. JMIR Mhealth Uhealth 2019;7:e12770.
51. Amirshahi A, Hashemi M. ECG classification algorithm based on STDP and R-STDP neural networks for real-time monitoring on ultra low-power personal wearable devices. IEEE Trans Biomed Circuits Syst 2019;13:1483-93.
52. Wang N, Zhou J, Dai G, Huang J, Xie Y. Energy-efficient intelligent ECG monitoring for wearable devices. IEEE Trans Biomed Circuits Syst 2019;13:1112-21.
53. Wasserlauf J, You C, Patel R, Valys A, Albert D, Passman R. Smartwatch performance for the detection and quantification of atrial fibrillation. Circ Arrhythm Electrophysiol 2019;12:e006834.
54. Sadr N, Jayawardhana M, Pham TT, Tang R, Balaei AT, de Chazal P. A low-complexity algorithm for detection of atrial fibrillation using an ECG. Physiol Meas 2018;39:064003.
55. Mathews SM, Kambhamettu C, Barner KE. A novel application of deep learning for single-lead ECG classification. Comput Biol Med 2018;99:53-62.
56. Kamaleswaran R, Mahajan R, Akbilgic O. A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using single lead electrocardiograms of variable length.

Physiol Meas 2018;39:035006.
57. Teijeiro T, García CA, Castro D, Félix P. Abductive reasoning as a basis to reproduce expert criteria in ECG atrial fibrillation identification. Physiol Meas 2018;39:084006.
58. Hernandez F, Mendez D, Amado L, Altuve M. Atrial fibrillation detection in short single lead ECG recordings using wavelet transform and artificial neural networks. Annu Int Conf IEEE Eng Med Biol Soc 2018;2018:5982-5.
59. Rubin J, Parvaneh S, Rahman A, Conroy B, Babaeizadeh S. Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings. J Electrocardiol 2018;51:S18-21.
60. Khamis H, Chen J, Stephen Redmond J, Lovell NH. Detection of atrial fibrillation from RR intervals and PQRST morphology using a neural network ensemble. Annu Int Conf IEEE Eng Med Biol Soc 2018;2018:5998-6001.
61. Xiong Z, Nash MP, Cheng E, Fedorov VV, Stiles MK, Zhao J. ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. Physiol Meas 2018;39:094006.
62. Fan X, Yao Q, Cai Y, Miao F, Sun F, Li Y. Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings. IEEE J Biomed Health Inform 2018;22:1744-53.
63. Plesinger F, Nejedly P, Viscor I, Halamek J, Jurak P. Parallel use of a convolutional neural network and bagged tree ensemble for the classification of Holter ECG. Physiol Meas 2018;39:094002.
64. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. JAMA Cardiol 2018;3:409-16.
65. Zhou X, Zhu X, Nakamura K, Mahito N. Premature ventricular contraction detection from ambulatory ECG using recurrent neural networks. Annu Int Conf IEEE Eng Med Biol Soc 2018;2018:2551-4.
66. Yan BP, Lai WHS, Fong AHT, Lai PS, Cheng OW, Chan CKY, et al. Validation of a deep convolutional network for detecting atrial fibrillation with a wrist-worn wearable device (Abstract). Heart Rhythm 2018;15(5 Suppl):S362 B-PO03-185.
67. Clifford GD, Liu C, Moody B, Lehman LH, Silva I, Li Q, et al. AF classification from a short single lead ECG recording: the PhysioNet/computing in cardiology challenge 2017. Comput Cardiol (2010) 2017 Sep. [Epub]. Available at: https://doi.org/10.22489/CinC.2017.065-469.
68. Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Trans Biomed Eng 2016;63:664-75.
69. Liu X, Du H, Wang G, Zhou S, Zhang H. Automatic diagnosis of premature ventricular contraction based on Lyapunov exponents and LVQ neural network. Comput Methods Programs Biomed 2015;122:47-55.
70. Solosenko A, Petrenas A, Marozas V. Photoplethysmography-based method for automatic detection of premature ventricular contractions. IEEE Trans Biomed Circuits Syst 2015;9:662-9.
71. Liu SH, Cheng DC, Lin CM. Arrhythmia identification with two-lead electrocardiograms using artificial neural networks and support vector machines for a portable ECG monitor system. Sensors (Basel) 2013;13:813-28.
72. Oresko JJ, Duschl H, Cheng AC. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Trans Inf Technol Biomed 2010;14:734-40.
73. Jin Z, Sun Y, Cheng AC. Predicting cardiovascular disease from real-time electrocardiographic monitoring: an adaptive machine learning approach on a cell phone. Annu Int Conf IEEE Eng Med Biol Soc 2009;2009:6889-92.
74. Rodríguez J, Goñi A, Illarramendi A. Real-time classification of ECGs on a PDA. IEEE Trans Inf Technol Biomed 2005;9:23-34.
75. Hickey B, Heneghan C, de Chazal P. Non-episode-dependent assessment of paroxysmal atrial fibrillation through measurement of RR interval dynamics and atrial premature contractions. Ann Biomed Eng 2004;32:677-87.
76. Yu Y, Yang Z, You Y, Shan W. FASSNet: fast apnea syndrome screening neural network based on single-lead electrocardiogram for wearable devices. Physiol Meas 2021;42:085005.
77. Chang HY, Yeh CY, Lee CT, Lin CC. A sleep apnea detection system based on a one-dimensional deep convolution neural network model using single-lead electrocardiogram. Sensors (Basel) 2020;20:4157.
78. Iwasaki A, Nakayama C, Fujiwara K, Sumi Y, Matsuo M, Kano M, et al. Screening of sleep apnea based on heart rate variability and long short-term memory. Sleep Breath 2021 Jan 10. [Epub]. Available at: https://doi.org/10.1007/s11325-020-02249-0.
79. Papini GB, Fonseca P, van Gilst MM, Bergmans JWM, Vullings R, Overeem S. Wearable monitoring of sleep-disordered breathing: estimation of the apnea-hypopnea index using wrist-worn reflective photoplethysmography. Sci Rep 2020;10:13512.
80. Wang T, Lu C, Shen G. Detection of sleep apnea from single-lead ECG signal using a time window artificial neural network. Biomed Res Int 2019;2019:9768072.
81. Bozkurt MR, Uçar MK, Bozkurt F, Bilgin C. In obstructive sleep apnea patients, automatic determination of respiratory arrests by photoplethysmography signal and heart rate variability. Australas Phys Eng Sci Med 2019;42:959-79.
82. Wang T, Lu C, Shen G, Hong F. Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network. PeerJ 2019;7:e7731.
83. Lin Y, Wang P, Lin C, Sadrawi M, Lin C, Hsieh Y, et al. Apnea recognition in wearable device using the intensive evaluation of the ECG power spectral density. Sleep 2018;41:A127.
84. Urtnasan E, Park JU, Joo EY, Lee KJ. Automated detection of obstructive sleep apnea events from a single-lead electrocardiogram using a convolutional neural network. J Med Syst 2018;42:104.
85. Sharma H, Sharma KK. An algorithm for sleep apnea detection from single-lead ECG using Hermite basis functions. Comput Biol Med 2016;77:116-24.
86. Babaeizadeh S, White DP, Pittman SD, Zhou SH. Automatic detection and quantification of sleep apnea using heart rate variability. J Electrocardiol 2010;43:535-41.
87. Allen J, Liu H, Iqbal S, Zheng D, Stansby G. Deep learning-based photoplethysmography classification for peripheral arterial disease detection: a proof-of-concept study. Physiol Meas 2021;42:054002.
88. Lee JJ, Heo JH, Han JH, Kim BR, Gwon HY, Yoon YR. Prediction of ankle brachial index with photoplethysmography using convolutional long short term memory. J Med Biol Eng 2020;40:282-91.
89. Dall'Olio L, Curti N, Remondini D, Safi Harb Y, Asselbergs FW, Castellani G, et al. Prediction of vascular aging based on smartphone acquired PPG signals. Sci Rep 2020;10:19756.
90. Alty SR, Angarita-Jaimes N, Millasseau SC, Chowienczyk PJ. Predicting arterial stiffness from the digital volume pulse waveform. IEEE Trans Biomed Eng 2007;54:2268-75.
91. Allen J, Murray A. Comparison of three arterial pulse waveform classification techniques. J Med Eng Technol 1996;20:109-14.
92. Allen J, Murray A. Development of a neural network screening aid for diagnosing lower limb peripheral vascular disease from pho-

toelectric plethysmography pulse waveforms. Physiol Meas 1993;14:13-22.
93. Baig MM, GholamHosseini H, Gutierrez J, Ullah E, Lindén M. Early detection of prediabetes and T2DM using wearable sensors and internet-of-things-based monitoring applications. Appl Clin Inform 2021;12:1-9.
94. Avram R, Olgin JE, Kuhar P, Hughes JW, Marcus GM, Pletcher MJ, et al. A digital biomarker of diabetes from smartphone-based vascular signals. Nat Med 2020;26:1576-82.
95. Porumb M, Griffen C, Hattersley J, Pecchia L. Nocturnal low glucose detection in healthy elderly from one-lead ECG using convolutional denoising autoencoders. Biomed Signal Process Control 2020;62:102054.
96. Porumb M, Stranges S, Pescapè A, Pecchia L. Precision medicine and artificial intelligence: a pilot study on deep learning for hypoglycemic events detection based on ECG. Sci Rep 2020;10:170.
97. Faruqui SHA, Du Y, Meka R, Alaeddini A, Li C, Shirinkam S, et al. Development of a deep learning model for dynamic forecasting of blood glucose level for type 2 diabetes mellitus: secondary analysis of a randomized controlled trial. JMIR Mhealth Uhealth 2019;7:e14452.
98. Lee S, Lee HC, Chu YS, Song SW, Ahn GJ, Lee H, et al. Deep learning models for the prediction of intraoperative hypotension. Br J Anaesth 2021;126:808-17.
99. Kwon JM, Kim KH, Medina-Inojosa J, Jeon KH, Park J, Oh BH. Artificial intelligence for early prediction of pulmonary hypertension using electrocardiography. J Heart Lung Transplant 2020;39:805-14.
100. Devaki V, Jayanthi T. Smartphone-based diagnostic model for hypertension using features from photoplethysmogram. Biomed Eng-Appl Basis Commun 2020;32:2050027.
101. Nafisi VR, Shahabi M. Intradialytic hypotension related episodes identification based on the most effective features of photoplethysmography signal. Comput Methods Programs Biomed 2018;157:1-9.
102. Liang Y, Chen Z, Ward R, Elgendi M. Photoplethysmography and deep learning: enhancing hypertension risk stratification. Biosensors (Basel) 2018;8:101.
103. Kwon JM, Kim KH, Akkus Z, Jeon KH, Park J, Oh BH. Artificial intelligence for detecting mitral regurgitation using electrocardiography. J Electrocardiol 2020;59:151-7.
104. Yang C, Ojha BD, Aranoff ND, Green P, Tavassolian N. Classification of aortic stenosis using conventional machine learning and deep learning methods based on multi-dimensional cardio-mechanical signals. Sci Rep 2020;10:17521.
105. Yang C, Aranoff ND, Green P, Tavassolian N. Classification of aortic stenosis using time-frequency features from chest cardio-mechanical signals. IEEE Trans Biomed Eng 2020;67:1672-83.
106. Kwon JM, Lee SY, Jeon KH, Lee Y, Kim KH, Park J, et al. Deep learning–based algorithm for detecting aortic stenosis using electrocardiography. J Am Heart Assoc 2020;9:e014717.
107. Ahmedov S, Amirjanov A. Genetic-fuzzy logic model for a non-invasive measurement of a stroke volume. Comput Methods Programs Biomed 2021;203:106046.
108. Wang L, Zhou X. Detection of congestive heart failure based on LSTM-based deep network via short-term RR intervals. Sensors (Basel) 2019;19:1502.
109. Cho J, Lee B, Kwon JM, Lee Y, Park H, Oh BH, et al. Artificial intelligence algorithm for screening heart failure with reduced ejection fraction using electrocardiography. ASAIO J 2021;67:314-21.
110. Rashid N, Al Faruque MA. Energy-efficient real-time myocardial infarction detection on wearable devices. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:4648-51.
111. Kwon JM, Kim KH, Jeon KH, Lee SY, Park J, Oh BH. Artificial intelligence algorithm for predicting cardiac arrest using electrocardiography. Scand J Trauma Resusc Emerg Med 2020;28:98.
112. Grogan M, Lopez-Jimenez F, Cohen-Shelly M, Dispenzieri A, Attia ZI, Abou Ezzedine OF, et al. Artificial intelligence-enhanced electrocardiogram for the early detection of cardiac amyloidosis. Mayo Clin Proc 2021;96:2768-78.
113. Kwon JM, Cho Y, Jeon KH, Cho S, Kim KH, Baek SD, et al. A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. Lancet Digit Health 2020;2:e358-67.
114. Chiarelli AM, Bianco F, Perpetuini D, Bucciarelli V, Filippini C, Cardone D, et al. Data-driven assessment of cardiovascular ageing through multisite photoplethysmography and electrocardiography. Med Eng Phys 2019;73:39-50.
115. Fan X, Zhao Y, Wang H, Tsui KL. Forecasting one-day-forward wellness conditions for community-dwelling elderly with single lead short electrocardiogram signals. BMC Med Inform Decis Mak 2019;19:285.
116. Penzel T, Moody GB, Mark RG, Goldberger AL, Peter JH. The apnea-ECG database. Computers in Cardiology 2000;27:255-8.
117. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 2000;101:E215-20.
118. Taddei A, Distante G, Emdin M, Pisani P, Moody GB, Zeelenberg C, et al. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. Eur Heart J 1992;13:1164-72.
119. Hennessey S. Heart for heart [Internet] [accessed on 2021 October 1]. Available at: http://www.heartrateapp.com.
120. Petrutiu S, Sahakian AV, Swiryn S. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. Europace 2007;9:466-70.
121. Jager F, Taddei A, Moody GB, Emdin M, Antolic G, Dorn R, et al. Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. Med Biol Eng Comput 2003;41:172-82.
122. Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Eng Med Biol Mag 2001;20:45-50.
123. Moody G. A new method for detecting atrial fibrillation using RR intervals. Comput Cardiol 1983;10:227-30.
124. Moody GB, Muldrow W, Mark RG. A noise stress test for arrhythmia detectors. Comput Cardiol 1984;11:381-4.
125. Saeed M, Villarroel M, Reisner AT, Clifford G, Lehman LW, Moody G, et al. Multiparameter Intelligent Monitoring in Intensive Care II: a public-access intensive care unit database. Crit Care Med 2011;39:952-60.
126. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016;3:160035.
127. van Gilst MM, van Dijk JP, Krijn R, Hoondert B, Fonseca P, van Sloun RJG, et al. Protocol of the SOMNIA project: an observational study to create a neurophysiological database for advanced clinical sleep monitoring. BMJ Open 2019;9:e030996.
128. Baim DS, Colucci WS, Monrad ES, Smith HS, Wright RF, Lanoue A, et al. Survival of patients with severe congestive heart failure treated with oral milrinone. J Am Coll Cardiol 1986;7:661-70.
129. Cai Z, Liu C, Gao H, Wang X, Zhao L, Shen Q, et al. An open-access long-term wearable ECG database for premature ventricular contractions and supraventricular premature beat detection. J Med Imaging & Health Infor 2020;10:2663-7.
130. Iyengar N, Peng CK, Morin R, Goldberger AL, Lipsitz LA. Age-re-

https://doi.org/10.3349/ymj.2022.63.S93

lated alterations in the fractal scaling of cardiac interbeat interval dynamics. Am J Physiol 1996;271:R1078-84.
131. Lee HC, Jung CW. Vital recorder-a free research tool for automatic recording of high-resolution time-synchronised physiological data from multiple anaesthesia devices. Sci Rep 2018;8:1527.
132. Moody G, Goldberger A, McClennen S, Swiryn S. Predicting the onset of paroxysmal atrial fibrillation: the computers in cardiology 2001. Computers in Cardiology 2001;28:113-6.
133. Redmond SJ, Xie Y, Chang D, Basilakis J, Lovell NH. Electrocardiogram signal quality measures for unsupervised telehealth environments. Physiol Meas 2012;33:1517-33.
134. Zhang Z, Pi Z, Liu B. TROIKA: a general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Trans Biomed Eng 2015;62:522-31.
135. Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG-signaldatenbank CARDIODAT der PTB über das internet. Biomed Eng 1995;40:317-8.
136. Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, et al. PTB-XL, a large publicly available electrocardiography dataset. Sci Data 2020;7:154.
137. Zheng J, Zhang J, Danioko S, Yao H, Guo H, Rakovski C. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Sci Data 2020;7:48.
138. Periyaswamy T, Balasubramanian M. Ambulatory cardiac biosignals: from mirage to clinical reality through a decade of progress. Int J Med Inform 2019;130:103928.
139. Bayoumy K, Gaber M, Elshafeey A, Mhaimeed O, Dineen EH, Marvel FA, et al. Smart wearable devices in cardiovascular care: where we are and how to move forward. Nat Rev Cardiol 2021;18:581-99.
140. Steinberg JS, Varma N, Cygankiewicz I, Aziz P, Balsam P, Baranchuk A, et al. 2017 ISHNE-HRS expert consensus statement on ambulatory ECG and external cardiac monitoring/telemetry. Heart Rhythm 2017;14:e55-96.
141. Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, et al. Smartwatch algorithm for automated detection of atrial fibrillation. J Am Coll Cardiol 2018;71:2381-8.
142. Ip JE. Wearable devices for cardiac rhythm diagnosis and management. JAMA 2019;321:337-8.