Computer Methods and Programs in Biomedicine 231 (2023) 107359

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Extensive deep learning model to enhance electrocardiogram application via latent cardiovascular feature extraction from identity identification

Yu-Sheng Lou[a,b], Chin-Sheng Lin[c,d], Wen-Hui Fang[e], Chia-Cheng Lee[f,g], Chin Lin[a,b,d,h,*]

[a] Graduate Institutes of Life Sciences, National Defense Medical Center, Taipei, Taiwan
[b] School of Public Health, National Defense Medical Center, Taipei, Taiwan
[c] Division of Cardiology, Department of Internal Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.
[d] Medical Technology Education Center, School of Medicine, National Defense Medical Center, Taipei, Taiwan, R.O.C.
[e] Department of Family and Community Medicine, Department of Internal Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.
[f] Department of Medical Informatics, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.
[g] Division of Colorectal Surgery, Department of Surgery, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.
[h] School of Medicine, National Defense Medical Center, Taipei, Taiwan, R.O.C.

ARTICLE INFO

Article history:
Received 21 March 2022
Revised 22 December 2022
Accepted 17 January 2023

Keywords:
Electrocardiogram
Deep learning
Cardiovascular disease
Transfer learning
Unsupervised learning
Identity identification

ABSTRACT

Background and Objective: Deep learning models (DLMs) have been successfully applied in biomedicine primarily using supervised learning with large, annotated databases. However, scarce training resources limit the potential of DLMs for electrocardiogram (ECG) analysis.
Methods: We have developed a novel pre-training strategy for unsupervised identity identification with an area under the receiver operating characteristic curve (AUC) >0.98. Accordingly, a DLM pre-trained with identity identification can be applied to 70 patient characteristic predictions using transfer learning (TL). These ECG-based patient characteristics were then used for cardiovascular disease (CVD) risk prediction. The DLMs were trained using 507,729 ECGs from 222,473 patients and validated using two independent validation sets (n = 27,824/31,925).
Results: The DLMs using our method exhibited better performance than directly trained DLMs. Additionally, our DLM performed better than those of previous studies in terms of gender (AUC [internal/external] = 0.982/0.968), age (correlation = 0.886/0.892), low ejection fraction (AUC = 0.942/0.951), and critical markers not addressed previously, including high B-type natriuretic peptide (AUC = 0.921/0.899). Additionally, approximately 50% of the ECG-based characteristics provided significantly more prediction information for cardiovascular risk than real characteristics.
Conclusions: This is the first study to use identity identification as a pre-training task for TL in ECG analysis. An extensive exploration of the relationship between ECG and 70 patient characteristics was conducted. Our DLM-enhanced ECG interpretation system extensively advanced ECG-related patient characteristic prediction and mortality risk management for cardiovascular diseases.
© 2023 Elsevier B.V. All rights reserved.

1. Introduction

Cardiovascular diseases (CVD) are a leading cause of death globally [1]. More than 19 million global deaths were caused by CVD in 2020, an increase of 18.7% from 2010 [2]. Symptoms, specific clinical history, and proper serial physical examinations such as electrocardiography (ECG), echocardiography, chest radiography, and laboratory testing, are still the cornerstone for the assessment and evaluation of CVD [3,4]. Although several studies have been conducted on CVD risk stratification [5–7], improvements can be made. Blood samples are the major component driving these risk stratification calculators [8], which are often lacking due to their intrusiveness. Moreover, these risk scores are available for less than 30% patients during electronic health record-based cardiovascular screening [9]. In addition to structural information, individual health information is largely unstructured [10]. Further, coronary artery calcium levels provide promising discrimination and risk reclassification for the prediction of incident CVD in intermediate-

* Correspondence at: Dr. Chin Lin, No.161, Min-Chun E. Rd., Sec. 6, Neihu, Taipei 114, Taiwan, ROC. Tel: 886–2-87923100#18574; Fax: 886–2-87923147
E-mail address: xup6fup0629@gmail.com (C. Lin).

https://doi.org/10.1016/j.cmpb.2023.107359
0169-2607/© 2023 Elsevier B.V. All rights reserved.

risk individuals [11]. Additionally, the application of free-text medical records to develop deep learning disease severity scores significantly enhances mortality prediction in adults with congenital heart disease [12]. Accordingly, these results highlight the critical role of unstructured data in the CVD risk stratification. Although clinical guidelines recommend CVD risk assessment, traditional methods such as CVD risk prediction equations and newly developed strategies are not widely conducted for people free of cardiovascular disease yet [13,14].

An ECG is an important biomedical engineering tool that contains whole-body anatomic features in addition to heart activities, and is rapid, inexpensive, and readily available. Although experienced clinicians can provide a comprehensive interpretation of ECG data, a significant amount of vital knowledge may remain hidden [15]. By learning the appropriate features based on data rather than manual engineering [16], deep learning models (DLMs) can extract features unrecognizable by humans, such as retinal fundus images, for cardiovascular risk factor estimation [17]. Moreover, ECG has been applied to extract individual information via DLM, such as gender and age [18]. Importantly, the performance of DLMs can reach human levels when large, annotated datasets are available [19–22]. Previous studies developed a series of ECG-based DLMs for arrhythmia [23], myocardial infarction [24], dyskalemia [25,26], left ventricular dysfunction [27,28], mortality [29], and anemia [30]. Moreover, ECGs can be used to predict the occurrence of atrial fibrillation (AF) in patients with sinus rhythm nearly one month earlier [31]. These achievements have significantly improved certain risk stratifications, prompting our development of a comprehensive ECG interpretation system equipped with maximal ECG potentials for CVD risk stratification.

A robust and efficient strategy for learning the appropriate features is critical for DLM development, particularly for transfer learning (TL) in small databases. The most widely used TL strategy in images classification and natural language processing were shown in Fig. 1 a. DLM pretrained via ImageNet has previously been applied to improve the accuracy of medical image analysis [32]. Additionally, bidirectional encoder representations from transformers using unsupervised learning are now a successful and basic architecture for most natural language interpretation tasks [33]. However, no widely recognized pretrained strategy for ECGs is available. A functional, sophisticated, and easily accessible database is crucial for designing pretrained unsupervised strategies. Individual identification via an ECG-based DLM has shown promising results [34]. Additionally, the geometrical aspects of the heart-lung torso available in the ECG [35] data may represent latent cardiovascular status. Importantly, the accessible identity information in hospital databases provides an opportunity to apply DLMs to recognize individual identities. We hypothesized that there is no direct relationship between identity and ECG exists, compelling the DLMs to extract cardiovascular-related features to constitute identity information. These high-order features may be associated with known cardiovascular biomarkers, and unknown information exists to predict the CVD outcomes. Our study provides a promising TL opportunity based on identity identification for future DLM studies to enhance the ECG interpretation, as shown in Fig. 1 a. To demonstrate the advantages of the proposed TL strategy, this study includes a complete experiment, as shown in Fig. 1 b, that predicted 70 patient characteristics (Table C.1) for thorough CVD risk stratification.

1.1. Literature review

Recently, studies have applied TL approaches to enhance the ability of ECG-based DLMs to diagnose CVDs [36–40]. Increasing training sample size may improve DLM performance. However, pretraining sources are limited since large, annotated ECG datasets with corresponding disease or laboratory data are often difficult to acquire [41].

One common approach is supervised TL, which pretrains the ECG-based DLM on a public dataset. Previous study pre-processed ECG data as a spectrogram and used these spectrograms for EfficientNet pre-trained on ImageNet [40]. Utilizing pre-trained EfficientNet and fine-tuning it resulted in a high F-measure of 86.13%, up from 74.36%, for the classification of AF on PhysioNet. However, another study using ECG spectrograms for GoogleNet also pretrained on ImageNet but obtained lower performance (F-measure:0.811) than random initialization (F-measure:0.843) for arrhythmia diagnosis [38]. These results imply that TL with a cross-domain needs further investigation to demonstrate consistently improving performance of ECG-based DLMs. In contrast, with the homogeneity of source data and target domain, a previous study applied TL for classifying AF using 8,528 ECGs from a premature atrial contractions database with over 20 million records to improve F-measure from 0.711 to 0.777 [39]. The highest accuracy improvement of DLM was noted for chest X-ray (CXR) analysis using TL of the same anatomical site compared with other anatomical sites using X-rays and ImageNet [42]. However, the lack of availability of large similar source data has restrained extensive CVD marker learning using the TL. Moreover, these studies only evaluated the impact of TL on a limited number of ECG tasks, and a comprehensive evaluation of TL performance improvement on wider ECG tasks should be conducted.

In addition, using unsupervised TL with an autoencoder for classifying arrhythmia on ECG data successfully enhanced the F-measure from 0.843 to 0.857 [38]. An autoencoder is designed with a constrained number of latent features to extract a meaningful representation [43], which is only available in networks containing latent variables with small dimensions [44], thereby negating the strength of a large network on a DLM [16]. Another study applied an unsupervised GAN-based framework with outlier detection to detect abnormal ECGs [45]. However, this method may be unstable during training [46] or suffer from overfitting [47] when used to deal with extensive CVD classifications. In this study, we propose unlimited pre-extraction of CVD-related features via unsupervised identity recognition that is easily accessible, has sophisticated strengths, and can be fine-tuned using large networks.

Research groups currently focus on diagnosing a single type of CVD through ECG-based DLM, including paroxysmal AF [31], arrhythmias [48,49], pulmonary hypertension [50], valvular heart disease [51,52], hypertrophic cardiomyopathy [53], low ejection fraction (EF) [28], myocardial ischaemia [24,36], pericarditis [54], anemia [30], and dyskalemia [26,55]. Some of these algorithms have already been applied clinically, such as the AI–ECG dashboard at Mayo Clinic [56] and Cardiologs® in emergency department ECGs [49]. The AI–ECG dashboard at the Mayo Clinic provides diagnoses of several cardiac diseases, including left ventricular systolic dysfunction, silent atrial fibrillation, and hypertrophic cardiomyopathy. Cardiologs® has been reported to conduct multi-label prediction in ECG analysis, including abnormal ECG features, different types of arrhythmias, and myocardial ischemia. These studies focused on a single cardiac disease, and its applications were restricted to a limited type of CVD. However, ECG can provide broad diagnostics of various cardiovascular diseases, such as myocarditis, myocardial fibrosis, pulmonary embolism, and dextrocardia, as well as systemic conditions such as hypothermia, cardiovascular side effects of drugs, and electrolyte imbalances including dskalemia and hyprthyroidism [57]. Nevertheless, an automatic solution for extensive diagnostics in ECG interpretation is lacking. Previous studies have chosen the disease outcome of ECG-based DLM based on the understanding of electrophysiological knowledge or relying on the experience of assessing patients in clinical care [56]. However, ECG-based DLM may be able to identify unknown systemic condi-

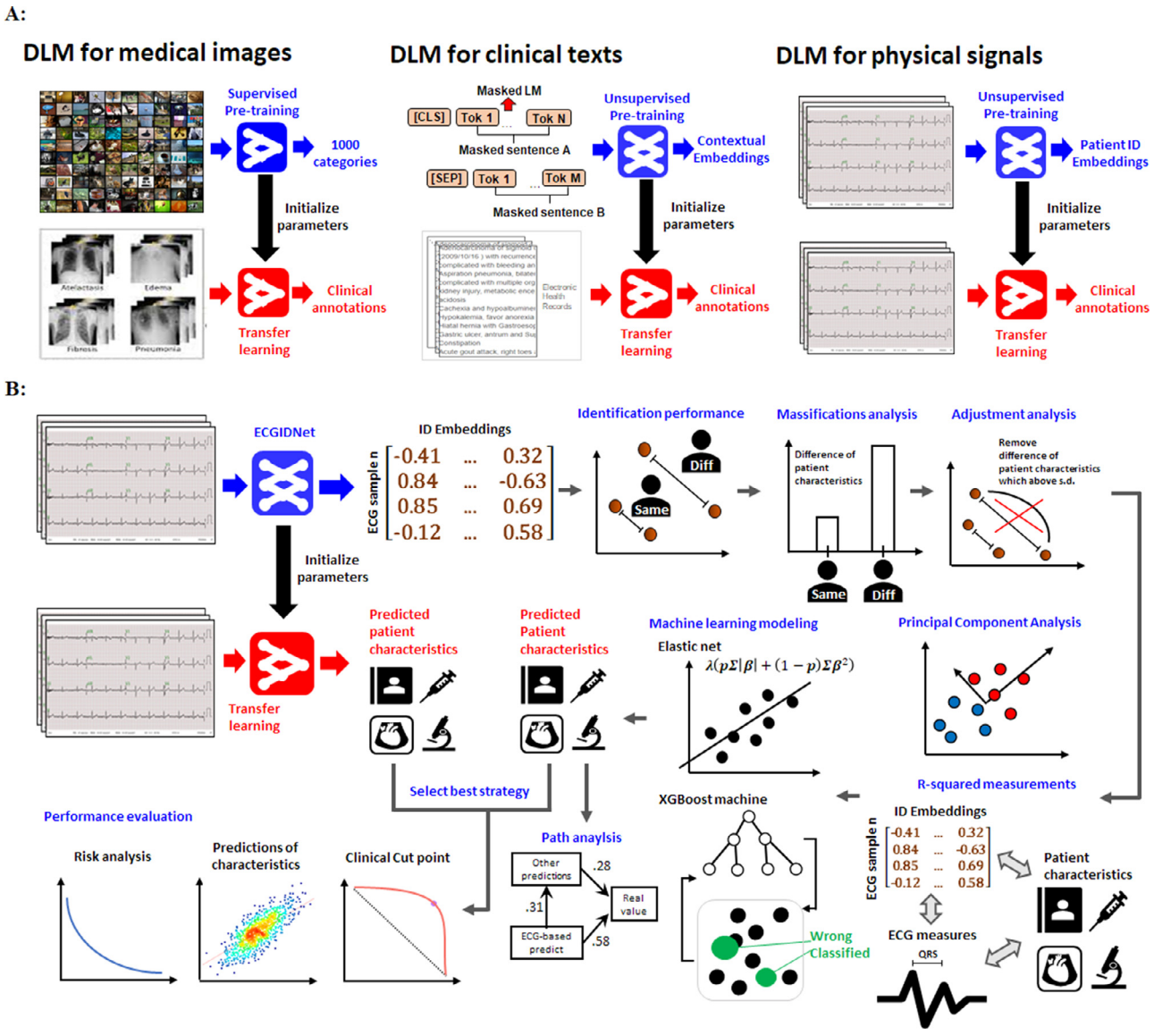

Fig. 1. Overview of experimental design for unsupervised identity identification. a. Contradistinction of training strategies for image recognition, natural language processing, and ECG analysis. ECGIDNet provides unsupervised pretrained parameters to make up for the deficiency in DLM studies on ECG. b. Summary of the study design, starting with ID embedding extraction and exploration.

tions of diseases since it has already demonstrated the ability to extract ECGs features beyond the capacity of an expert, such as predicting low EF [28] and mortality [29]. Since the private ECG dataset retrospectively collected by institutions is seldom openly published [58], the availability of datasets that may influence the disease was selected to extensive studying [59]. To extend and fully explore the capability of ECG-based DLM, a comprehensive strategy for the systematic extraction of cardiac-related features in the absence of large, well-annotated ECG datasets should be developed. However, some conditions might not be accurately predicted by the ECG-based DLM due to a lack of proper techniques or large-dataset availability. Therefore, in this study, we propose a method of pre-training DLM using identity recognition to enhance the DLM prediction of known cardiovascular biomarkers and explore unknown information to predict CVDs.

Importantly, DLM can identify patients with no initial disease having a higher risk of developing CVD. In a previous study, patients with normal EF were identified as having low EF by the ECG-based DLM; they had a 4-fold risk of developing ventricular dysfunction compared to patients initially predicted as normal [28]. Another study implemented an ECG-based DLM to predict patient age and reported that patients may have a higher mortality rate when their ECG-estimated age is greater than their actual age by 8 years [60]. In addition, the association between cardiovascular risk factors and prediction gap between ECG-estimated and actual age remained even in patients with normal ECG. These studies revealed the important prognostic ability of ECG-based DLM in analyzing false-positive cases. Moreover, this ability has also been reported in other studies using different medical data as model inputs. For CXR-based DLMs, patients positively predicted to have pulmonary hypertension had 2-fold risk of heart failure (HF) compared to patients predicted as negative [61]. Another study demonstrated that CXR-based age predicted by DLM had better prognostic performance for cardiovascular and all-cause long-term mortality compared to the traditional method [62] with an r^2 of 0.25∼0.37 between CXR-estimated age and actual age. In another study, contrast-enhanced computed tomography-based DLM achieved an AUC of 0.786 for diagnosing signet-ring cell carcinoma of gastric cancer [63], and these probability scores can be used to significantly predict patient prognosis, including overall survival

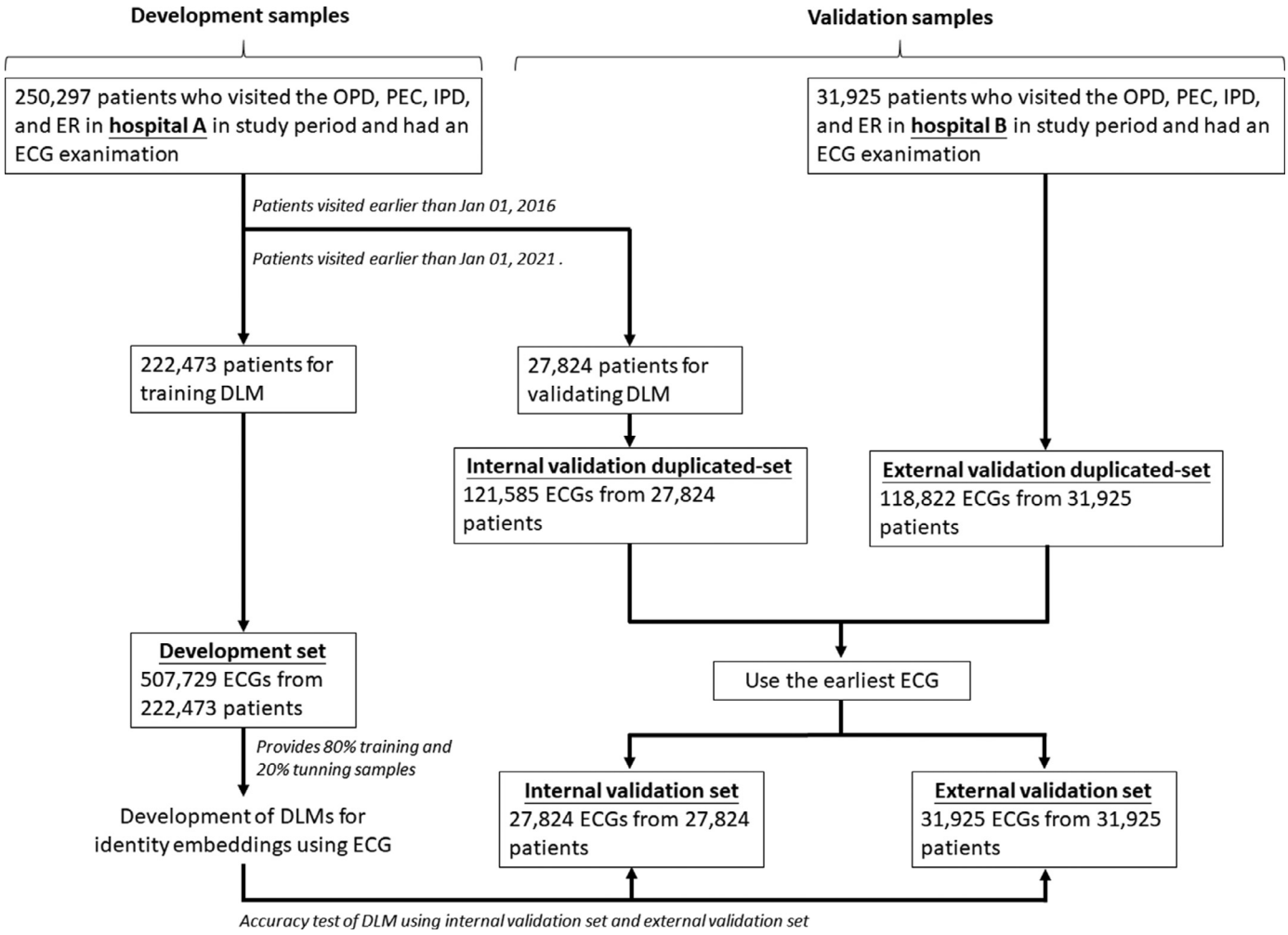

Fig. 2. The generation of development sets, internal validation sets, and external validation sets.

rate and chemotherapy resistance. Similar to the use of TL for DLM in medical image study [61–63], our TL strategy may empower ECG-based DLM as a more reliable tool to identify healthy patients with a higher risk of developing CVD via comprehensive CVD risk assessment.

2. Method

2.1. Population and dataset

This study was approved by the Institutional Review Board of Tri-Service General Hospital, Taipei, Taiwan (IRB NO. C202105049). We performed a retrospective multisite study at two hospitals in the Tri-Service General System from January 2011 to December 2020. The first site was an academic medical center (hospital A, NeiHu General Hospital in NeiHu District), with 1,800 beds, ~100,000 annual emergency room (ER) visits, and ~1,500,000 annual outpatient department (OPD) visits. The second site was a community hospital (hospital B, Tingzhou Branch Hospital at Zhongzheng District) with 100 beds, ~15,000 annual ER visits, and ~400,000 annual OPD visits. Although the two hospitals belong to the same hospital group, they opened separately in 1999 and 1946, respectively.

Fig. 2 illustrates the dataset generation process. Data from hospital A consisted of 250,297 patients aged >20 years who underwent at least one ECG examination. Among them, 27,824 patients with scans prior to January 1, 2016, were used for validating the DLMs and following CVD-related outcomes. We selected the earliest dataset as the validation set to maximize follow-up time. The remaining 222,473 patients were used for developing DLMs and 507,729 ECGs with corresponding patient characteristics were used to construct a development set. This dataset simultaneously provided training and tuning samples for gradient descent and hyperparameter decision, respectively. An internal validation duplicated set including 121,585 ECGs from 27,824 independent patients was used to evaluate DLM performance for identity identification, and the earliest ECG of each patient was used to generate an internal validation set to conduct an accuracy test on patient characteristic predictions and CVD outcome analyses. The remaining 31,925 non-overlapping patients aged >20 years at Hospital B who met the same criteria were also included in this study. Following the same criteria mentioned above, 118,822 ECGs and 31,925 earliest ECGs were used to generate an external validation duplicated set and external validation set, respectively.

2.2. ECG data

The ECGs were recorded in a standard 12-lead format and collected at 500 Hz for 10 s, resulting in 5,000 sequence signals from each lead. The baseline numeric was 0 and the unit was 0.01 millivolt. This is called the ECG voltage–time trace for DLM usage. Quantitative measurements and findings within the final ECG clinical reports were extracted to identify 31 diagnostic pattern classes and eight continuous ECG measurements. The eight ECG measure-

ments included heart rate, PR interval, QRS duration, QT interval, correct QT interval, P-wave axis, RS-wave axis, and T-wave axis. Data for these variables were 93–100% complete, and the missing values were input using multiple imputations [64]. The patterns included abnormal T wave, AF, atrial flutter, atrial premature complex, complete AV block, complete left bundle branch block, complete right bundle branch block, first-degree AV block, incomplete left bundle branch block, incomplete right bundle branch block, ischemia/infarction, junctional rhythm, left anterior fascicular block, left atrial enlargement, left axis deviation, left posterior fascicular block, left ventricular hypertrophy, low QRS voltage, pacemaker rhythm, prolonged QT interval, right atrial enlargement, right ventricular hypertrophy, second-degree degree AV block, sinus bradycardia, sinus pause, sinus rhythm, sinus tachycardia, supraventricular tachycardia, ventricular premature complex, ventricular tachycardia, and Wolff–Parkinson–White syndrome. These 31 clinical diagnosis patterns were parsed from the structured findings statements based on key phrases that are standard within the Philips system. These parameters are called ECG measures for machine learning model usage.

2.3. Patient characteristic data

This study included 70 patient characteristics, as shown in Table C.1, collected from the electronic medical records (EMRs) of both hospitals. These patient characteristics were classified into 10 demographic data, 16 echocardiography results, and 44 laboratory markers (2 for diabetes mellitus, 4 for hyperlipidemia, 5 for chronic kidney disease, 4 for thyroid function, 6 for hepatitis, 1 for cholangitis, 2 for anemia, 4 for blood profile, 7 for electrolyte, 3 for cardiac enzyme, 4 for gas analysis, and 2 for others). Each ECG was annotated by the nearest record of the corresponding patient characteristics, and records without ECG examination within the corresponding time limit were excluded, as shown in Table C.1. Continuous variables were also limited by the corresponding value range, and some variables with a positive skew distribution were log-transformed. All analyses in this study are based on the values obtained after processing. All continuous variables were pre-decided based on clinically significant cutoff points to simulate real clinical practice. For example, EF was binarized to $\leq$35% and >35%. Certain variables had a U-shaped relationship with clinical prognosis, such as potassium (K^+) level [65]. Therefore, two cut-off points were used for further analyses. Moreover, some cutoff points were different for males and females, such as high-density lipoprotein cholesterol; therefore, all cutoff points were presented as two values for each sex.

2.4. CVD-related outcomes

The outcomes of interest were all-cause death, new-onset acute myocardial infarction (AMI), new-onset stroke, new-onset coronary artery disease (CAD), and new-onset HF. Mortality events were defined based on updates from EMRs of the hospitals. Moreover, data for live visits were censored at the patient's last known hospital live encounter to limit bias due to incomplete records. The other outcomes were based on a new diagnosis according to the corresponding International Classification of Diseases, Ninth Revision and Tenth Revision (ICD-9 and ICD-10, respectively) as follows: AMI, ICD-9 codes 410.x and ICD-10 codes I21.x; stroke, ICD-9 codes 430.x to 438.x and ICD-10 codes I60.x to I63.x; CAD, ICD-9 codes 410.x to 414.x and 429.2 and ICD-10 codes I20.x to I25.x; HF, ICD-9 codes 428.x, 398.91, and 402.x1, and ICD-10 code I50.x. Patients with at least one record of less than or equal to 35% EF were also considered to have HF. For each outcome, patients with a corresponding diagnosis before the first ECG examination were excluded from follow-up analysis. Therefore, at-risk patients in the internal/external validation sets for each outcome were as follows: death (26,048/30,738), AMI (24,011/27,618), stroke (22,758/24,967), CAD (21,131/23,170), and HF (23,301/26,381). The number of events during the median following years of 5.3/2.7, 5.4/3.1, 5.0/2.7, 4.3/2.4, and 5.2/2.9 on death, AMI, stroke, CAD, and HF, respectively, in internal/external validation sets were 679/987 (2.6%/3.2%), 353/388 (1.5%/1.4%), 1,704/1,724 (7.5%/6.9%), 2,904/2,869 (13.7%/12.4%), and 1,103/1,444 (4.7%/5.5%).

2.5. The implementation of DLMs

The major architecture in this study is a siamese convolutional neural network with shift invariance of morphological features [26,31], called ECGIDNet. It uses ECG sequence data to calculate high-order features that are used for identity recognition and called identification (ID) embeddings. The architecture of ECGIDNet is shown in Fig. 3, and the total number of learnable parameters in ECGIDNet is approximately 2.97 million. We assumed that DLM can only use cardiovascular-related features to constitute information, and high-order features, ID embeddings, may be useful for further tasks. We used DLM to learn ID embeddings from ECGs belonging to one or two individuals. Each ECG was recorded using 12 standard leads, consisting of 5,000 sequences. This 5,000 $\times$ 12 matrix was used as the input without cropping. We developed an architecture composed of 12 ECG lead blocks as ECG12Net [26]. The 5,000 signals from each lead were spliced and input to the ECG lead block. The weights of the ECG lead blocks are shared to avoid overfitting. We added two additional non-shared layers comprising of a lead-specific module for each ECG lead block, primarily used for feature extraction from each lead.

As the downsampling size of each ECG lead block was 256, 19 high-order features belonged to each ECG signal. The general integrating method for high-order features is average or max pooling. However, we used the routing-by-agreement algorithm to integrate feature maps of each output of the final convolution layer for each lead [66]. Our preliminary data showed that the performance of routing-by-agreement was better than that of the traditional pooling layer. The size of the feature maps of each lead are 1 $\times$ 1 $\times$ 128 after routing-by-agreement. Finally, all the feature maps were flattened as vectors of 1 $\times$ 1,536. These vectors are ID embeddings that may be useful for further tasks.

ID embeddings were used to recognize individual identities based on the Euclidean distance. The L2 distances of the two identity embeddings are large when obtained from distinct people and small if they are from the same person. Accordingly, the loss function of ECGIDNet is as follows:

$$loss(f_i, f_k, y_{ik}, \theta) = \begin{cases} \frac{1}{2}\|f_i - f_k\|_2^2 & if\ y_{ik} = 0 \\ \frac{1}{2}\max(0, \theta - \|f_i - f_k\|_2)^2 & if\ y_{ik} = 1 \end{cases}$$

where f_i and f_k represent the ID embeddings of the two ECGs. y_{ik} is equal to zero when two ECGs are from the same individual and equal to one when two ECGs are from different individuals. The Euclidean distance should be greater than the hyperparameter θ when two ECGs are from different individuals. The θ value was set to 1 in our experiments. If two ECGs were from the same individual, the distance was close to zero.

We used 80% of the ECG data in the development set to optimize the parameters of ECGIDNet and used the other 20% as the tuning subset to select the final model. The ECGIDNet was trained using adaptive moment estimation (Adam), which is an improved optimization method for stochastic gradient descent. This method minimizes the output of the loss function by iteratively updating the parameters of the neural network using a small part of the data over the training dataset every iteration. We trained our models using Adam with standard parameters ($B_1 = 0.9$, $B_2 = 0.999$), mini-batch size of 32 ECG pairs, and weight decay of 0.01. The

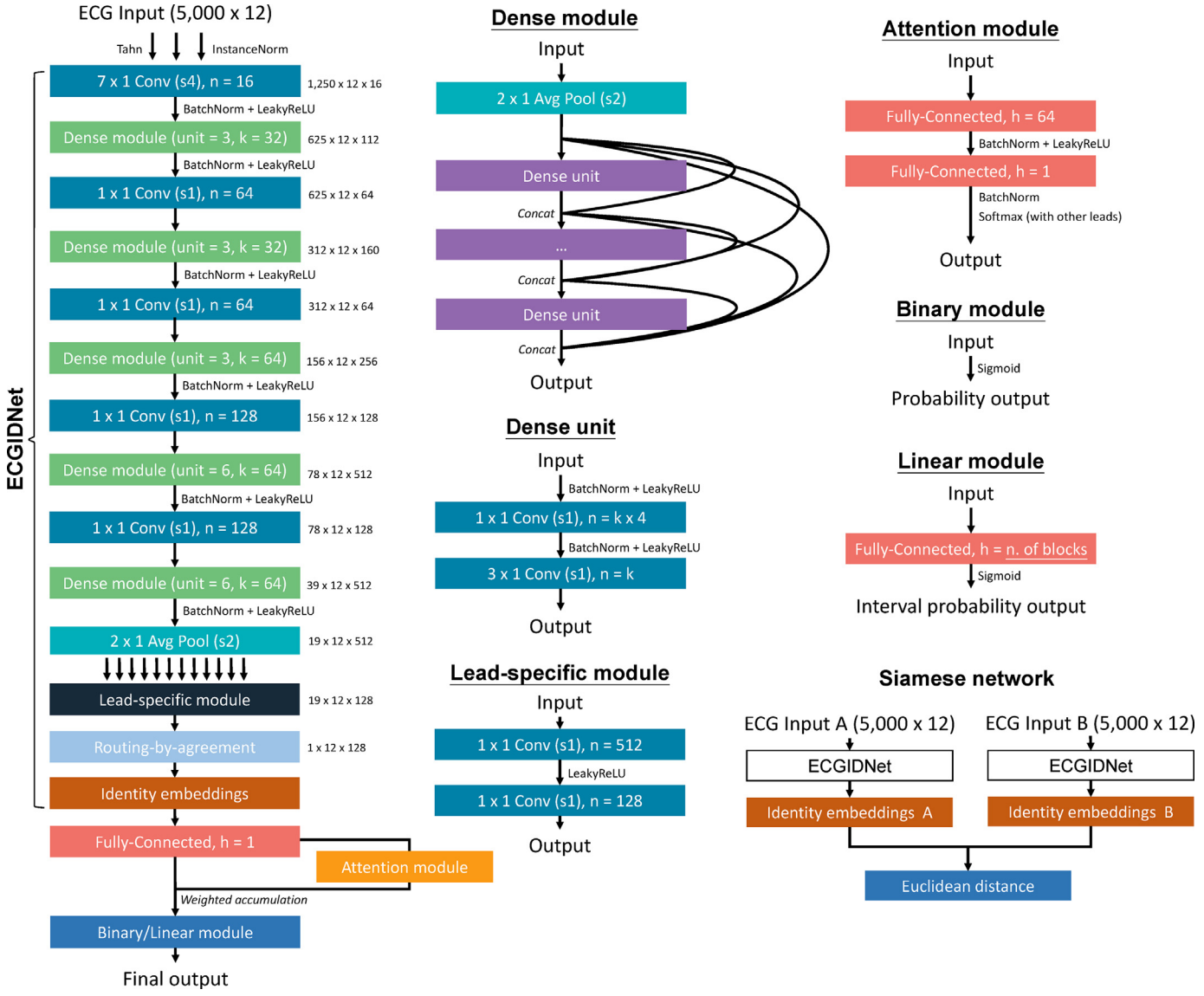

Fig. 3. Model architectures of the ECGIDNet, Siamese network, and predicted modules.

model parameters were initialized using the Xavier initialization method. Due to the extreme imbalance between two ECGs from the same patients and from different patients in our development set, an equal number of ECGs from the same person and from a different person were sampled for each minibatch during training, comprising 16 ECG pairs for the same and difference person, each. We trained ECGIDNet for 90 epochs (single pass over the full training data) with an initial learning rate of 0.0001, which decreased by a factor of 10 at the 30^{th} and 60^{th} epochs. Moreover, a simple form of data augmentation was used to reduce overfitting, which involved adding Gaussian noise with a mean of zero and variance of one into the input of the ECG sequences. ECG12Net was implemented using the mxnet package (R package version 1.3.0). We calculated the area under the receiver operating characteristic curve (AUC) after each epoch using the tuning subset, and the final model was selected based on the highest AUC.

We constructed DLMs to predict each patient's characteristics, and the architecture was based on an extension of ECGIDNet as shown in Fig. 3. A fully connected layer (FC) with the attention module was used after high-order features of the ECGIDNet. The architectures used to predict continuous and binary variables is according to different variable types of patient characteristics. For binary prediction, the neuron of the FC was set to one, followed by sigmoid output, which provided a probability ranging from zero to one. Binary cross-entropy was used as the loss function. For continuous prediction, category-wise encoding was used to encode the labels. Continuous variables were converted into ordinal variables that included 20 categories. Twenty value intervals were generated based on the lower and upper bounds described in Table C.1, and the values outside the range were limited. For example, a case older than 80 years old is encoded as [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]. Cases aged 55 years old will be encoded as [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], and so on. We used FCs with 20 neurons for 20 categories of ordinal variables, and each neuron was followed by a sigmoid function. Binary cross-entropy was used as the loss function for each sigmoid output.

Similar training details were based on ECGIDNet, and more details of the DLMs for predicting patient characteristics are provided in Appendix A. The only two changes were oversampling and model selection strategies. To handle class imbalances in the patient characteristic label, oversampling was applied to ensure that the DLM could learn from the minority classes. For the binary and

continuous variables converted into ordinal variables, the sampling weight was the inverse of the number of classes for each mini-batch during training. During training, the performance of the DLM was evaluated for each epoch in the tuning subset. For continuous variables, mean absolute error was used. For the binary variable, the AUC was used. The model with the best performance on the tuning subset was selected as the final DLM.

2.6. Implementation of machine learning models

The eXtreme gradient boosting model (XGBoost) and elastic net were used to build the prediction model using high-order features. The high-order features in this study included the ECG measures and ID embedding. We used xgboost (R package version 0.71.2) and glmnet (R package version 2.0–16) for implementation. The number of rounds of XGBoost was set to 150, which achieved the best performance in the preliminary data, and the remaining parameters were set to default. We used the glmnet() function to implement elastic net and selected the parameter α with the best model in the tuning subset. Parameter α determines the weighting between Lasso and Ridge. Candidate α was set to [0.0, 0.1, 0.2, …, 0.9, 1.0], and the remaining parameters were set to their default values. We used only the development set to train these models, and the performance assessment of each patient characteristic was conducted only once using the internal and external validation sets.

2.7. Statistical analysis for identity recognition

We first analyzed the identification performance and followed by performing misclassification analysis and adjustment analysis. Misclassification analysis evaluated the differences in patient characteristics when the identity of ECGs pairs was misclassified. Adjustment analysis evaluated the identification performance after excluding the ECG pairs with large differences in patient characteristic. These experiments are indicated in Fig. 1 b.

Identity recognition performance was based on the ECG pairs. All analyses were based on duplicated internal and external validation sets. All possible pairs of ECGs from the same individual were calculated. However, we only sampled 1,000,000 ECG pairs and from different individuals, owing to the extensive computing time. To determine a suitable threshold to distinguish positive from negative, we used the value with the largest Youden's index in the tuning subset. All analyses were based on the same threshold to calculate sensitivity and specificity.

Moreover, we performed misclassification analysis by analyzing its effect on the difference of the corresponding ECG states when the identity of ECGs pairs was misclassified. The differences in patient characteristics were evaluated for the correct and incorrect pairs in the validation sets, which were compared using standardized mean differences (SMDs). The details are provided in Appendix A.

2.8. Statistical analysis for use of ID embeddings

The following analyses were based on the internal and external validation sets, which included only the earliest ECGs of each patient. We performed principal component analysis (PCA) on ID embeddings and R-squared measurements among ID embeddings, patient characteristics, and ECG measures. These experiments are indicated in Fig. 1 b. We visualized the relationship between ID embeddings and patient characteristics using PCA. We obtained a matrix of variable loadings in the tuning subset and used the same matrix to rotate the ID embeddings in the internal and external validation sets. To preserve the relative space of the ID embeddings, they were not shifted to zero, centered, or scaled before PCA.

To analyze the association between patient characteristics, ECG measures, and ID embeddings, we evaluated the explained variance (R square) among three of them. The estimation of the explained variance is described in Appendix A.

We conducted machine learning models to predict patient characteristics, as indicated in Fig. 1 b. ID embeddings and ECG measures were used as the input of machine learning model, respectively. We assessed the correlations between the prediction and patient characteristics using XGBoost and elastic nets in the internal and external validation sets. The AUCs were used to quantify performance based on the pre-assigned cutoff points. We further conducted path analysis which analyzed the direct and indirect relationships between each prediction with ID embeddings and corresponding patient characteristics, as shown in Fig. 1 b. The relationships were investigated using path analysis, an analysis of multiple regressions based on the hypothesized model. The details of the path analysis are provided in Appendix A. In this study, the hypothesized model had both direct and indirect effects. The direct effect is the direct estimation of patient characteristics on their real value. The indirect effect predicts patient characteristics based on the estimation of other patient characteristics, which is a mediator in our hypothesized model. Path diagram illustrated in Fig. D.6 shows an example of the estimation of brain natriuretic peptide (BNP) and estimated glomerular filtrate rate (eGFR). The path coefficients were calculated based on the correlation matrix using Pearson's correlation, polychoric correlation, or polyserial correlation, where appropriate, implemented using the polycor package (R package version 4.6–14).

2.9. Statistical analysis for deep learning models

We evaluated the model performance of TL and randomly initiated parameters in the internal and external validation sets, as indicated in Fig. 1 b. TL is a method to overcome the limitations of a small training sample size. Thus, we compared sample sizes of 1,000 and 10,000 ECG samples (if sufficient ECG samples with labels were available). For 1,000 samples, we randomly sampled 800 fixed ECGs from the development set for each DLM and 200 fixed ECGs from the tuning subset. Differences in the AUCs were evaluated.

We used the Bland–Altman plots to present differences in each predictive ability using different methods. As shown in Fig. 1 b, we selected the best method with the best performance in the tuning subset as an integrated strategy. These methods are using DLM with randomly initiated parameters, DLM with TL based on identity identification, and elastic net using ID embedding as input, as illustrated in Fig. D.9. The final performance of patient characteristic prediction is shown using scatter plots and receiver operating characteristic (ROC) curves, which was evaluated once in both validation sets. The decision threshold for the binary variable was based on the best cutoff point in the tuning set.

2.10. Statistical analysis for CVD-related outcomes

We employed a Cox proportional hazards model to predict the risk of CVD-related outcomes in the tuning subset, and the prediction of each model was directly applied to the internal and external validation sets. We employed univariable Cox models for actual patient characteristics, univariable Cox model for ECG-based characteristics, and multivariable Cox model to integrate them. For variables with a U-shaped relationship to clinical prognosis, we used the pspline() function with 2 degrees of freedom using the survival package (R package version 2.43–3) to fit the U-shaped re-

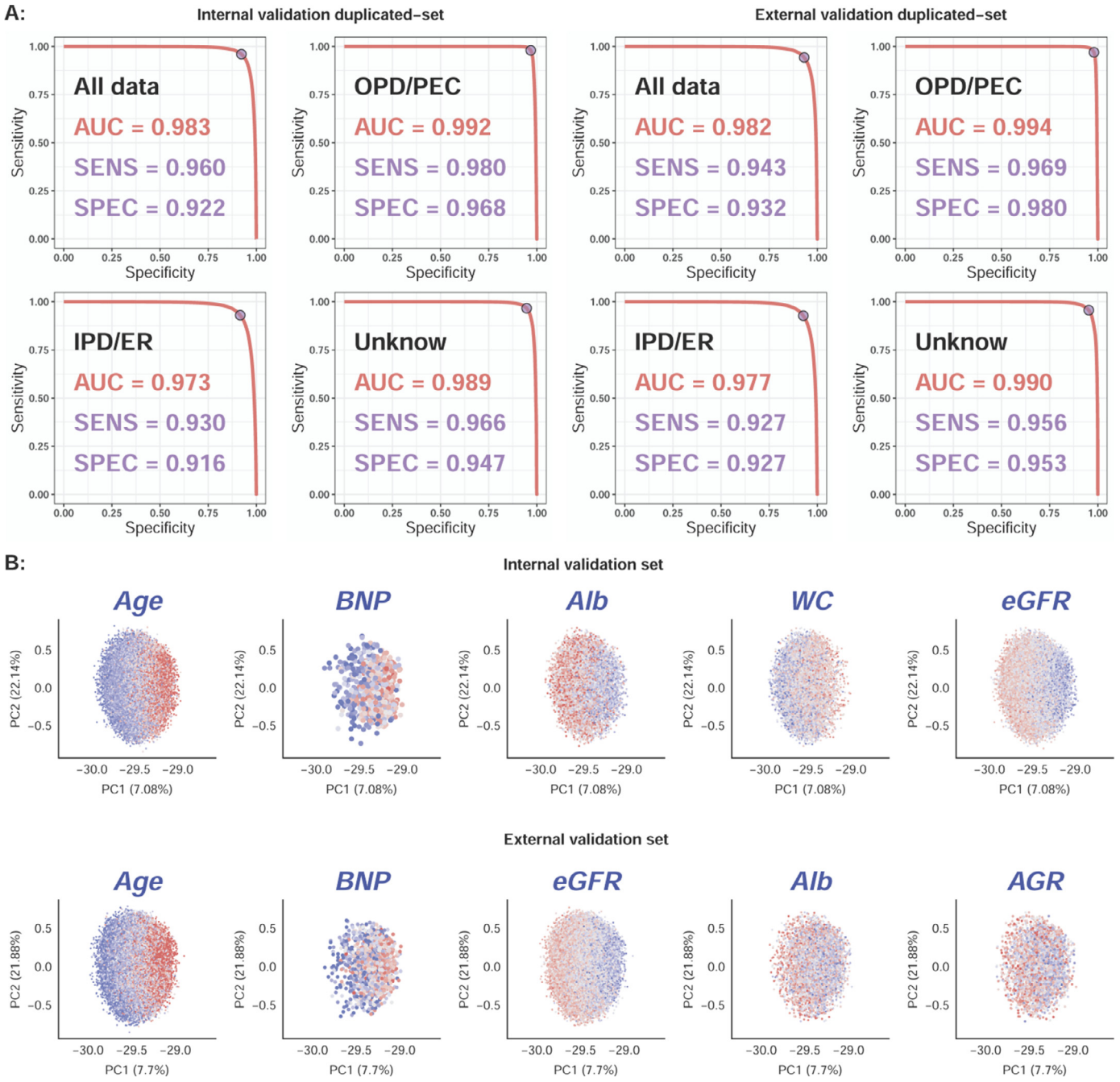

Fig. 4. Identity recognition of ECGIDNet. **a.** ROC curve analysis of individual identity recognition in validation sets; **b.** PCA of the correlation between ID embeddings and patient characteristics. The category variables are shown in red and blue, and a continuous blue to red gradient displays the increased continuous variables;

lationship. The concordance index (C-index) was used to evaluate the performance of each model.

3. Results

3.1. Statistical analysis result of identity identification

First, we evaluated the identity recognition performance of ECGIDNet. Fig. 4 a shows the ROC curves of patient identification from ECG pairs of the same or different patients. Sensitivity refers to the correct identification among different patients, and specificity refers to the correct identification of the same patient. The areas under the ROC curves (AUCs) were 0.983/0.982, with sensitivities of 0.960 and 0.943 and specificities of 0.922 and 0.932 for the internal and external validation sets, respectively. Stratified analyses of ECG sources revealed higher performance for the OPD and physical examination center (PEC) patients, with AUCs of 0.992 and 0.994 for the internal and external validation sets, respectively, than for the inpatient department (IPD) and ER patients.

We further analyzed the relationship between ID classification and patient characteristics. Fig. D. 1a shows the misclassification analyses of ECG pairs from different patients in the internal and external validation sets. We calculated the difference between the corresponding characteristics of the correct and incorrect pairs, presented as the SMD (Tables C.2 and C.3), and proposed 20 critical variables. Interestingly, the characteristic differences contributed to the correct identification of ECG pairs from different patients, even though they were not provided for unsu-

pervised learning. The most critical variables were age, followed by BNP and albumin (Alb) in the internal validation set and age, followed by BNP and eGFR, in the external validation set. The critical variables were highly consistent for both the validation sets.

To further evaluate whether patient characteristics contribute to ECGIDNet identity recognition, we excluded ECG pairs with standard deviation >1 of each variable (Tables C.4 and C.5), expecting to increase the specificity of identifying the same patients due to fewer characteristic differences in ECG pairs and decrease the sensitivity of identifying different patients due to difficulty in discrimination. We calculated the sensitivity and specificity after excluding these ECG pairs and compared them to the performance without excluding pairs. As shown in Fig. D. 1b, sensitivity decreased from 0.9597 to 0.9442, and specificity increased from 0.9224 to 0.9358 after excluding the ECG pairs with large age differences in the internal validation duplicated-set. A consistent trend of decreased sensitivity and increased specificity in the internal and external validation duplicated-sets was noted, except for unchanged patient characteristics in our EMRs, such as waist circumference (WC), sex, and drinking status (Drink), which had only one record for each patient. This evidence emphasizes that high-order features extracted by ECGIDNet, called identity (ID) embeddings, are correlated with certain patient characteristics.

We further visualized the associations between ID embeddings and patient characteristics using PCA, as shown in Fig. 4 b. It shows the top five variables with the largest SMDs in Fig. D. 1a in the internal and external validation sets. Continuous color gradients for age are shown for both validation sets. The distinguished color gradients elucidate a strong relationship between ID embeddings and the selected characteristics. Results of PCA for all patient characteristics are shown in Figs. D.2 and D.3.

By estimating the explained variance, we analyzed the information on ID embeddings to predict patient characteristics (Appendix B). The ID embeddings contain much more information regarding patient characteristics than ECG measures, as indicated in Fig. D.4. Therefore, we directly used ID embeddings to estimate individual patient characteristics. To evaluate the superiority of ID embeddings over ECG measures, we applied them as inputs to the machine learning models. Tables C. 6 and C. 7 show the detailed performance of elastic net and XGBoost using ID embeddings and ECG measures. The prediction accuracy of elastic net using ID embeddings was better than that of XGBoost, followed by XGBoost using ECG measures, and elastic net using ECG measures, further confirming the superiority of ID embeddings. As shown in Fig. D. 5a, we used the Bland–Altman plots to summarize the AUC comparisons of four strategies for each patient characteristic. Elastic net had better performance than XGBoost, with McNemer odds ratios (ORs) of 3.56/3.56 in the internal and external validation sets, respectively, and mean AUC differences of 0.033/0.030, indicating a more linear relationship between ID embeddings and patient characteristics. Notably, XGBoost performed better than elastic net using ECG measures, implying a nonlinear relationship between ECG measures and patient characteristics, thereby elucidating better latent cardiovascular feature extraction by unsupervised ECGIDNet than manual ECG measures. We conducted path analysis (examples are illustrated in Fig. D.6) to investigate the direct and indirect effects of the hypothesized model. Fig. D. 5b shows path analysis of the 15 critical variables with the highest AUCs in elastic net using ID embeddings, and each column of heatmap shows the size of path coefficient for these patient characteristics. Certain predictive abilities were directly enhanced by corresponding patient characteristics in internal/external validation sets, including gender (100.0%/100.0%), age (100.0%/100.0%), BNP (100.0%/79.9%), and Alb (85.7%/90.1%). However, the predictive ability due to certain patient characteristics was found to be indirect. For example, 37.8% and 35.1% prediction of eGFR resulted from the direct effect, while a majority was mediated by ECG age (39.7% and 40.6%, respectively). Similar results were reported for hemoglobin (Hb) indirectly predicted by ECG hematocrit (HCT), ECG gender, and ECG red blood cells (RBCs). In summary, this analysis demonstrates that although certain ECG information is lost, ID embeddings are superior to ECG measures when predicting patient characteristics, and each ECG-based patient characteristic included indirect cardiovascular information.

3.2. Performance analysis of transfer learning from identity embeddings

We used ECGIDNet as a pretrained TL model to enhance feature expression using supervised learning with annotations. To predict patient characteristics, we constructed DLMs based on ECGIDNet with additional neuron layers. We compared the performance of DLMs with pretrained weights from ECGIDNet (with TL) and DLMs with initialized parameters at random (without TL). To demonstrate the effects of TL on a limited sample size, we compared the predictive abilities of patient characteristics with different sample sizes of 1,000, 10,000, and the full sample of the original data. The AUC performance with and without TL for each patient characteristic with different sample sizes is shown in Fig s . D.7 and D.8. As shown in Fig. 5 a, the selected patient characteristic predictions significantly improved, which demonstrated the superiority of using TL based on identity identification in predicting hyperkalemia, hypokalemia, lower eGFR, lower free calcium, and higher BNP. Collectively, the DLM trained with TL was superior to that trained without TL, particularly for a smaller sample size. The Bland–Altman plot summarizes the performance improvement of TL for the internal and external validation sets (Fig. 5 b). The ORs of AUC improvement were 2.21/2.16, 1.17/1.41, and 1.42/1.11 in the internal/external validation sets using training sizes of 1,000, 10,000, and full sample, respectively, which were inversely related to the sample size. Moreover, the averages of the improved AUCs with TL were 0.023/0.019, 0.003/0.005, and 0.003/0.008 for the internal/external validation sets with sample sizes of 1,000, 10,000, and the full sample, respectively, clearly demonstrating a significant improvement in the performance for a small sample size. Most researchers determine the best training strategy, such as with or without TL, based on the performance of the model after tuning the subsets. Our ID embeddings provided another strategy for direct prediction with an elastic net and were used to determine the suitable models for the tuning subset (Fig. D.9). Fig. 5 c shows significant improvement of our integrated strategy in full sample size training compared to traditional DLM (without TL). Using our training strategy, the proportions of increased and decreased characteristic predictions were 49%/49% and 26%/27% for the internal/external validation sets, respectively, with an average AUC improvement of 0.010/0.015. This analysis highlights the superiority of our ECG training strategy over the traditional method.

We used an integrated strategy with the full sample size to predict 70 patient characteristics via ECG and presented the performance as the ROC curve for dichotomous variables and scatter plots for continuous variables. Fig. 6 a shows the ROC curves for predicting patient gender, low EF, higher BNP, lower eGFR, and lower free calcium. The algorithm provided AUCs of 0.982/0.968 for gender distinction, which is better than previous reports with an AUC of 0.969 [18]. Moreover, our model exhibited high discrimination between EF $\leq$35% and EF >35% [AUCs = 0.942/0.951 for the internal/external validation sets] with sensitivities of 0.859/0.824 and specificities of 0.913/0.924, which is better than the highest AUC of 0.933 from previous reports [27,28,67]. Notably, ECG was first applied to predict higher BNP and lower eGFR, with AUCs of 0.921/0.899 and 0.881/0.840 in the internal/external validation sets, respectively. Satisfactory AUCs of 0.823/0.832 for higher

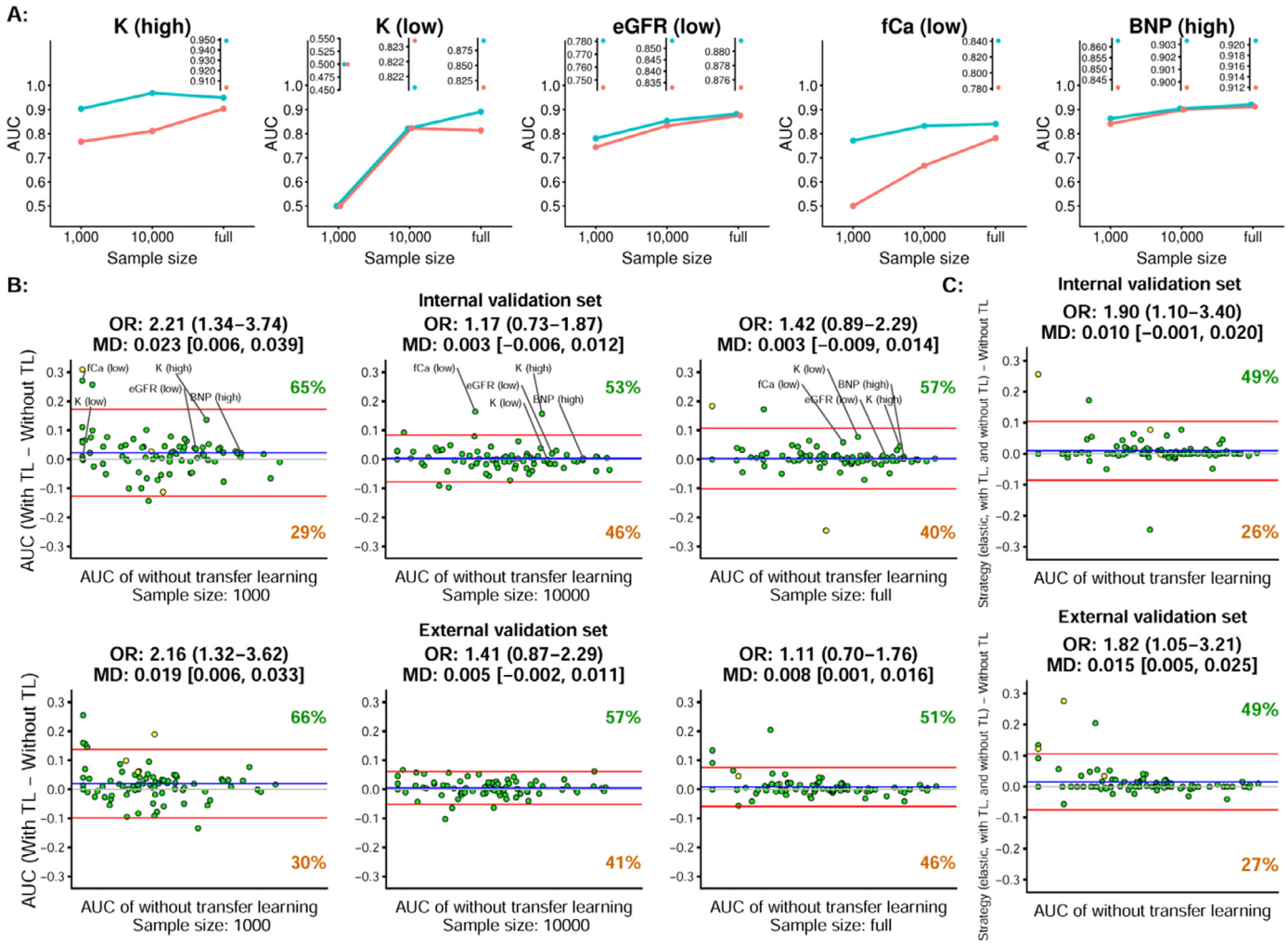

Fig. 5. Performance improvement of TL over ID embeddings. **a.** Selected comparisons of AUCs from DLMs with TL and without TL in the internal validation set; **b.** Performance comparison summary between DLMs with and without TL in the internal and external validation sets in a series of sample sizes. The McNemer odds ratios (OR) were calculated based on the proportion comparison of >0 (dark green) vs. <0 (brown) in y-axis, and the mean differences (MD) were calculated based on the mean of AUC difference. The full sample size of patient characteristics with more than 10,000 is indicated in green, and the full sample size with less than 10,000 is indicated in yellow; **c.** Comparison of correlations using integrated strategy and traditional DLM studies (without TL) for internal and external validation sets.

free triiodothyronine (fT3) detection were emphasized, although a large training sample size (n = 2,406) was lacking in our training strategy. Fig. 6 b shows the complete AUCs of all patient characteristics with corresponding clinically significant cutoff points (Table C.1) for the internal/external validation sets. Our strategy was superior in the above long-term stable patient characteristics and demonstrated better AUCs of 0.949/0.992 for hyperkalemia detection than those in a previous study [25,26]. Figs. D.10 and D.11 shows the complete ROC curves of all patient characteristics. Importantly, 10 (12.2%)/8 (9.8%) and 62 (75.6%)/52 (63.4%) clinically important features were identified using ECG with AUCs of >0.9 and >0.7, respectively, which demonstrates the extensive role of DLM-enhanced ECG interpretation. Fig s . D. 12 and D. 13 shows the complete scatter plots of the prediction compared to actual value for all patient characteristics in the internal/external validation sets. For age prediction, our model showed mean absolute errors (MAEs) of 6.109/6.471 years and correlations of 0.886/0.892 on the internal/external validation sets. These values are significantly better than the MAE of 6.9 and correlation of 0.837 in a previous study [18]. A total of 22 (31.4%) and 15 (21.4%) patient characteristics had a correlation of >0.5, demonstrating the quantitative capacity of our model for disease severity. Although certain patient characteristics were not highly correlated with the corresponding ECG parameters, we are confident that DLMs with our training strategy will fully extract the maximum correlations between ECG data and patient characteristics to achieve state-of-the-art detection levels.

Estimation errors may exist in latent cardiovascular features for the prediction of CVD outcomes, as described previously [18,28]. Therefore, we simultaneously used ECG-based and actual patient characteristics to predict the five CVD-related outcomes. We compared the C-index of the separate Cox model using actual, ECG age, and the combination of both as predictors the internal validation set, as shown in Fig. D.14. The C-indices of real age were 0.785, 0.735, 0.740, 0.666, and 0.789 for death, acute myocardial infarction (AMI), stroke, coronary artery disease (CAD), and heart failure (HF) predictions, respectively. The ECG-age provided significantly higher C-indices of 0.758 and 0.803 for AMI and HF predictions, respectively, than real age. The C-indices of the model integrating both the real and ECG ages were significantly higher than those using real age alone for all outcomes of interest. We also applied the risk matrixes to assess the risk of corresponding age and ECG age groups on CVD outcomes, as shown in Fig. D.14. In patients belonging to the younger real age group (normal group, <65 years old), 3.79-, 4.27-, 2.77-, 2.10-, and 4.43-fold risks of death, AMI, stroke, CAD, and HF, respectively, were pre-

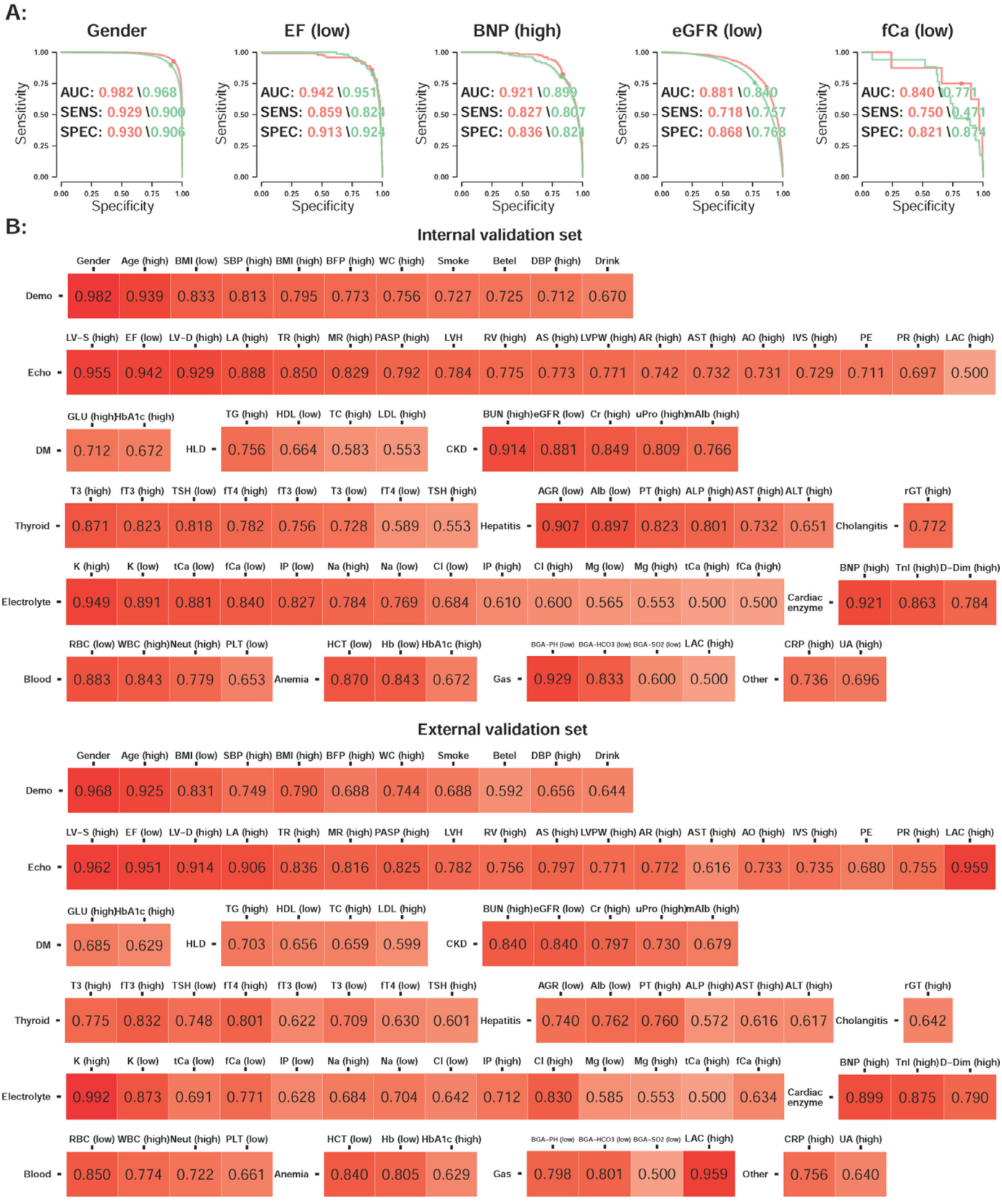

Fig. 6. Final prediction performance of each patient characteristic using the integration training strategy. **a.** ROC curves of specificity against sensitivity to identify patients by gender, an EF of ≤35%, BNP of ≥500 ng/L, eGFR of ≤60 ml/min, and fCa of ≤0.88 mmol/L for the internal and external validation sets. The operating point was determined by the tuning subset. **b.** Completed AUCs list of patient characteristics stratified by the corresponding attributes with clinically significant cutoff points for the internal and external validation sets.

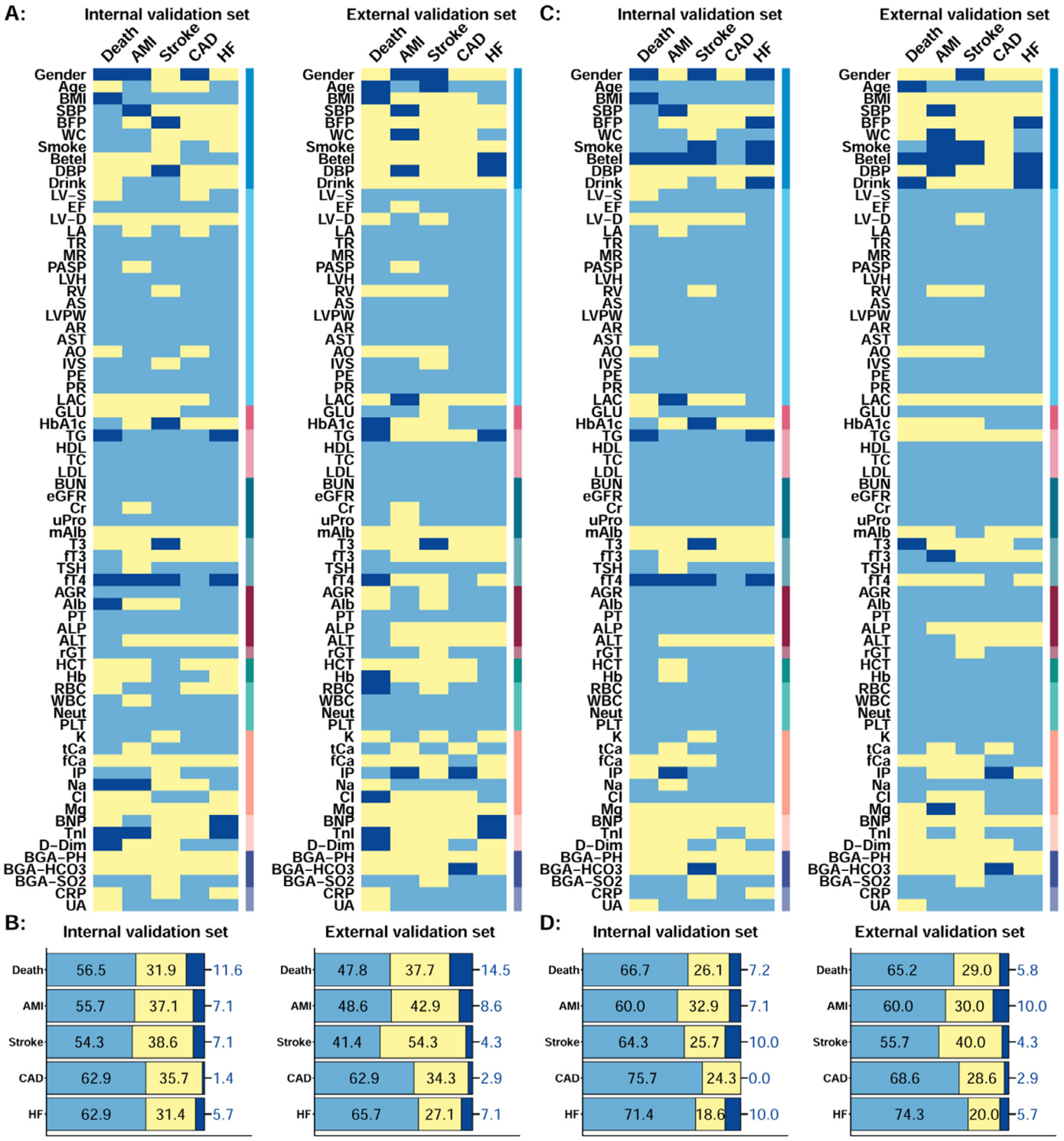

Fig. 7. Risk effect analysis of the actual patient characteristics, ECG-based characteristics, and their combinations. **a.** Heatmap summary of C-index comparison between actual patient and ECG-based characteristics. Dark blue indicates that the actual patient characteristics had a significantly higher C-index, and light blue indicates that ECG-based characteristics were superior. Yellow indicates no significant difference between their C-indices; **b.** Proportion of significant differences between the actual patient and ECG-based characteristics on each outcome. **c.** Heatmap summary of the C-index comparison between the actual patient characteristics and corresponding integrations; **d.** Proportion of significant differences between the actual patient characteristics and corresponding integrations on each outcome.

dicted when the DLM defined the ECG as the elderly group (AI-abnormal, false-positive) compared to the younger ECG group classified by the DLM (AI-normal, true negative). CVD outcome prediction for all 70 patient characteristics for the internal and external validation sets is shown in Figs. D.14 and D.153. We applied a heatmap to visualize the C-index comparison between actual patient and ECG-based characteristics (Fig. 7 a). Although most ECG-based demographic parameters did not perform significantly better than the actual parameters in predicting the five outcomes, ECG-based echocardiographic parameters, lipid profiles, renal function status, and blood cell counts clearly demonstrated the superiority of the outcome predictive capacities compared to the actual pa-

rameters. Fig. 7 b summarizes the proportion of significant differences between actual patient and ECG-based characteristics, and it shows that ECG-based characteristics provide more information for CVD outcome prediction. 56.5%/47.8%, 55.7%/48.6%, 54.3%/41.4%, 62.9%/62.9%, and 62.9%/65.7% ECG-based characteristics with significantly higher C-indices for death, AMI, stroke, CAD, and HF, respectively, were reported for the internal and external validation sets, which were superior to actual patient characteristics, with proportions of 11.6%/14.5%, 7.1%/8.6%, 7.1%/4.3%, 1.4%/2.9%, and 5.7%/7.1, respectively. Fig. 7 c shows the C-index comparison between each actual patient characteristic and the corresponding integrations of the ECG and actual parameters. ECG-based characteristics provided significantly more information on outcome predictions. Fig. 7 d summarizes the proportion of significant differences between the integrated model and actual patient characteristics. In summary, the integration model provided more information on most outcome predictions than actual patient characteristics alone.

4. Discussion

Based on TL from ID embeddings, our study developed an extensive ECG interpretation system for 70 patient characteristics that exhibited higher accuracy than previous studies and included several critical markers previously unaddressed, thereby providing extra information on the prediction of future CVD. Importantly, we highlight the application of DLM with TL to improve the accuracy of patient characteristic prediction, particularly in smaller databases.

The application of unsupervised ID recognition is an extraordinary breakthrough that successfully overcomes the hinderances of TL in ECG analyses. Compared to other TL strategies in ECG-based DLM, our unsupervised identity recognition method, which can pre-extract CVD-related features, is easily accessible, has sophisticated strengths, and can be used on large networks. Previous studies have chosen pre-trained tasks using abundant label data, gender, to pre-train DLM, and fine-tuned it to predict rare genetic heart disease [68] or Alzheimer's [69]. We did not choose pre-trained tasks, such as sex classification or age estimation, although that label is not manually annotated since the prediction of gender or age may only be based on a specific part of the biological meaning. By learning identity identification, the DLM can extract various CVD-related features, as observed in our PCA. Thus, previous TL strategies using single annotated labels, such as gender, may have limitations for extensively improving the performance of ECG-based DLM for diagnosing CVD, and our method could enhance the ability of DLM to comprehensively diagnose CVD in ECG analysis.

The success of unsupervised learning from identity extraction led to the proposal of existing unmeasured CVD factors in ECGs, confirming our hypothesis of no direct relationship between identity and ECG. Patients with the same chronological age and older ECG age exhibit a higher prevalence of hypertension, CAD, or low EF than those with equal or younger ECG age [18]. Moreover, patients with an abnormal ECG-based ejection fraction have a 4-fold higher risk of developing future left ventricular dysfunction than patients without echocardiographic left ventricular dysfunction [28]. In our study, we demonstrated that ECG-based patient characteristics had better performance in predicting all-cause mortality and CVD development. Importantly, significant trend in certain categories of patient characteristics, such as echocardiographic results, diabetes mellitus markers, and hyperlipidemia-related laboratory markers, raised a fascinating hypothesis that ECG-based DLM could predict these outcomes via certain important known and unknown patient characteristics. For example, the known relationship between echocardiographic results and these outcomes [70] and the association between diabetes mellitus and abnormal ECG [71] have been reported. Our ECG-based DLM with identity identification learning strategy could comprehensively extract important features, including risk factors that have not been well defined before, to provide better prognosis to predict these outcomes. Similar studies for extensive examination by DLM in medical images have been conducted [72], demonstrating that systemic conditions, including diabetic retinal disease diagnoses, glycated hemoglobin, total cholesterol, and triglycerides, can be identified from low-cost external eye images rather than using fundus photographs as input data, using a pre-trained DLM on ImageNet. They commented that this initial finding implied that DLM can be used to detect other underlying systemic conditions, including known and novel conditions, such as thyroid disease [73] and adverse cardiac outcomes [74], through imaging of the external or anterior segment of the eye. Our study uncovered many ECG-based characteristics and determined the significance of estimation errors in future CVD outcomes, which could be extensively applied to enhance comprehensive CVD outcome prediction.

Comprehensive screening of potential CVDs using a single ECG examination is critical in our study. In addition to acute cardiac diseases with typical ECG findings, such as myocardial infarction and arrhythmias, large amounts of ECGs in patients with long-term CVDs exhibit subtle changes, which are easily overlooked and undetected by even experienced physicians [15]. Apart from the largely missing laboratory data in traditional risk-score screening systems with potentially limited application [9], our unattended-execution ECG interpretation system exhibits simple and extensive properties to provide a holistic evaluation of CVD. Moreover, its application as an automatic and prompt alarming tool for CVD can help clinical physicians identify potential unrecognized diseases early [75]. Importantly, with DLM and our TL strategies, our approach could be applied to detect certain rare and potentially lethal CVDs that had not been established before.

This study had some limitations. First, the ECGs were obtained from a single ethnic group, although DLM might be diagnosed with ECG undisturbed by race [76]. Therefore, an international study involving different racial and ethnic groups should be conducted. Second, some variables, such as smoking status, were self-reported and may have been biased. Third, DLM parameters initialized at random may result in variable performance. The limited computing resources hindered repeated experiments, especially for tuning the hyperparameter of the DLM, evaluating the performance under different sample sizes of training data in the analysis of TL, and fine-tuning all parameters of machine learning models, which primarily used default settings.

5. Conclusion

In conclusion, we developed a novel TL method for an ECG DLM based on unsupervised identity extraction. The ECG-based DLM can extract CVD-related features via identity identification that can be provided as a pre-trained task for fine-tuning DLM for extensively predicting CVD markers, including many that were previously undisclosed. More than 50% of ECG-based characteristics significantly improved, and ECG-based characteristics can provide more prognostic information regarding the development of CVDs than real characteristics. As an inexpensive and practical tool for the screening and risk stratification of CVDs, our strategy could be further applied to detect rare diseases by ECG using DLM in the near future.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

Statements of ethical approval

This study was approved by the Institutional Review Board of Tri-Service General Hospital, Taipei, Taiwan (IRB NO. C202105049). Since we retrospectively used de-identified data collected and encrypted from the hospital to the data controller, informed consent waiver was granted for this study.

Funding

This work was supported by the Ministry of Science and Technology, Taiwan [MOST 110-2314-B-016-010-MY3 to C. Lin]; Cheng Hsin General Hospital, Taiwan [CHNDMC-111-07 to C. Lin]; and Medical Affairs Bureau [MND-MAB-110-113, MND-MAB-D-111045, and MND-MAB-C13-112050 to C. Lin].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2023.107359.

References

[1] E.J. Benjamin, P. Muntner, A. Alonso, M.S. Bittencourt, C.W. Callaway, A.P. Carson, A.M. Chamberlain, A.R. Chang, S. Cheng, S.R. Das, F.N. Delling, L. Djousse, M.S.V. Elkind, J.F. Ferguson, M. Fornage, L.C. Jordan, S.S. Khan, B.M. Kissela, K.L. Knutson, T.W. Kwan, D.T. Lackland, T.T. Lewis, J.H. Lichtman, C.T. Longenecker, M.S. Loop, P.L. Lutsey, S.S. Martin, K. Matsushita, A.E. Moran, M.E. Mussolino, M. O'Flaherty, A. Pandey, A.M. Perak, W.D. Rosamond, G.A. Roth, U.K.A. Sampson, G.M. Satou, E.B. Schroeder, S.H. Shah, N.L. Spartano, A. Stokes, D.L. Tirschwell, C.W. Tsao, M.P. Turakhia, L.B. VanWagner, J.T. Wilkins, S.S. Wong, S.S. Virani, Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association, Circulation 139 (2019) e56–e528.

[2] C.W. Tsao, A.W. Aday, Z.I. Almarzooq, A. Alonso, A.Z. Beaton, M.S. Bittencourt, A.K. Boehme, A.E. Buxton, A.P. Carson, Y. Commodore-Mensah, M.S.V. Elkind, K.R. Evenson, C. Eze-Nliam, J.F. Ferguson, G. Generoso, J.E. Ho, R. Kalani, S.S. Khan, B.M. Kissela, K.L. Knutson, D.A. Levine, T.T. Lewis, J. Liu, M.S. Loop, J. Ma, M.E. Mussolino, S.D. Navaneethan, A.M. Perak, R. Poudel, M. Rezk-Hanna, G.A. Roth, E.B. Schroeder, S.H. Shah, E.L. Thacker, L.B. VanWagner, S.S. Virani, J.H. Voecks, N.-Y. Wang, K. Yaffe, S.S. Martin, Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association, Circulation 145 (2022) e153–e639.

[3] P.A. Heidenreich, B. Bozkurt, D. Aguilar, L.A. Allen, J.J. Byun, M.M. Colvin, A. Deswal, M.H. Drazner, S.M. Dunlay, L.R. Evers, J.C. Fang, S.E. Fedson, G.C. Fonarow, S.S. Hayek, A.F. Hernandez, P. Khazanie, M.M. Kittleson, C.S. Lee, M.S. Link, C.A. Milano, L.C. Nnacheta, A.T. Sandhu, L.W. Stevenson, O. Vardeny, A.R. Vest, C.W. Yancy, 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines, Circulation 145 (2022) e895–e1032.

[4] M. Humbert, G. Kovacs, M.M. Hoeper, R. Badagliacca, R.M.F. Berger, M. Brida, J. Carlsen, A.J.S. Coats, P. Escribano-Subias, P. Ferrari, D.S. Ferreira, H.A. Ghofrani, G. Giannakoulas, D.G. Kiely, E. Mayer, G. Meszaros, B. Nagavci, K.M. Olsson, J. Pepke-Zaba, J.K. Quint, G. Rådegran, G. Simonneau, O. Sitbon, T. Tonia, M. Toshner, J.L. Vachiery, A. Vonk Noordegraaf, M. Delcroix, S. Rosenkranz, 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension, Eur Respir J (2022).

[5] P.W. Wilson, R.B. D'Agostino, D. Levy, A.M. Belanger, H. Silbershatz, W.B. Kannel, Prediction of coronary heart disease using risk factor categories, Circulation 97 (1998) 1837–1847.

[6] R.M. Conroy, K. Pyorala, A.P. Fitzgerald, S. Sans, A. Menotti, G. De Backer, D. De Bacquer, P. Ducimetiere, P. Jousilahti, U. Keil, I. Njolstad, R.G. Oganov, T. Thomsen, H. Tunstall-Pedoe, A. Tverdal, H. Wedel, P. Whincup, L. Wilhelmsen, I.M. Graham, Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project, European heart journal 24 (2003) 987–1003.

[7] R.B. D'Agostino Sr., R.S. Vasan, M.J. Pencina, P.A. Wolf, M. Cobain, J.M. Massaro, W.B. Kannel, General cardiovascular risk profile for use in primary care: the Framingham Heart Study, Circulation 117 (2008) 743–753.

[8] D.C. Goff Jr., D.M. Lloyd-Jones, G. Bennett, S. Coady, R.B. D'Agostino, R. Gibbons, P. Greenland, D.T. Lackland, D. Levy, C.J. O'Donnell, J.G. Robinson, J.S. Schwartz, S.T. Shero, S.C. Smith Jr., P. Sorlie, N.J. Stone, P.W. Wilson, H.S. Jordan, L. Nevo, J. Wnek, J.L. Anderson, J.L. Halperin, N.M. Albert, B. Bozkurt, R.G. Brindis, L.H. Curtis, D. DeMets, J.S. Hochman, R.J. Kovacs, E.M. Ohman, S.J. Pressler, F.W. Sellke, W.K. Shen, S.C. Smith Jr., G.F. Tomaselli, 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Circulation 129 (2014) S49–S73.

[9] R.S. Hira, K. Kennedy, V. Nambi, H. Jneid, M. Alam, S.S. Basra, P.M. Ho, A. Deswal, C.M. Ballantyne, L.A. Petersen, S.S. Virani, Frequency and practice-level variation in inappropriate aspirin use for the primary prevention of cardiovascular disease: insights from the National Cardiovascular Disease Registry's Practice Innovation and Clinical Excellence registry, Journal of the American College of Cardiology 65 (2015) 111–121.

[10] T.B. Murdoch, A.S. Detsky, The inevitable application of big data to health care, Jama 309 (2013) 1351–1352.

[11] J. Yeboah, R.L. McClelland, T.S. Polonsky, G.L. Burke, C.T. Sibley, D. O'Leary, J.J. Carr, D.C. Goff, P. Greenland, D.M. Herrington, Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals, Jama 308 (2012) 788–795.

[12] G.P. Diller, A. Kempny, S.V. Babu-Narayan, M. Henrichs, M. Brida, A. Uebing, A.E. Lammers, H. Baumgartner, W. Li, S.J. Wort, K. Dimopoulos, M.A. Gatzoulis, Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients, European heart journal 40 (2019) 1069–1077.

[13] U.P.S.T. Force, Screening for Cardiovascular Disease Risk With Electrocardiography: US Preventive Services Task Force Recommendation Statement, JAMA 319 (2018) 2308–2314.

[14] P. Jülicher, C. Varounis, Estimating the cost-effectiveness of screening a general population for cardiovascular risk with high-sensitivity troponin-I, European Heart Journal - Quality of Care and Clinical Outcomes 8 (2022) 342–351.

[15] X.L. Yang, G.Z. Liu, Y.H. Tong, H. Yan, Z. Xu, Q. Chen, X. Liu, H.H. Zhang, H.B. Wang, S.H. Tan, The history, hotspots, and trends of electrocardiogram, J Geriatr Cardiol 12 (2015) 448–456.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.

[17] R. Poplin, A.V. Varadarajan, K. Blumer, Y. Liu, M.V. McConnell, G.S. Corrado, L. Peng, D.R. Webster, Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning, Nature biomedical engineering 2 (2018) 158–164.

[18] Z.I. Attia, P.A. Friedman, P.A. Noseworthy, F. Lopez-Jimenez, D.J. Ladewig, G. Satam, P.A. Pellikka, T.M. Munger, S.J. Asirvatham, C.G. Scott, R.E. Carter, S. Kapa, Age and Sex Estimation Using Artificial Intelligence From Standard 12-Lead ECGs, Circ Arrhythm Electrophysiol 12 (2019) e007284.

[19] V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P.C. Nelson, J.L. Mega, D.R. Webster, Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs, JAMA 316 (2016) 2402–2410.

[20] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (2017) 115–118.

[21] D.S.W. Ting, C.Y. Cheung, G. Lim, G.S.W. Tan, N.D. Quang, A. Gan, H. Hamzah, R. Garcia-Franco, I.Y. San Yeo, S.Y. Lee, E.Y.M. Wong, C. Sabanayagam, M. Baskaran, F. Ibrahim, N.C. Tan, E.A. Finkelstein, E.L. Lamoureux, I.Y. Wong, N.M. Bressler, S. Sivaprasad, R. Varma, J.B. Jonas, M.G. He, C.Y. Cheng, G.C.M. Cheung, T. Aung, W. Hsu, M.L. Lee, T.Y. Wong, Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes, JAMA 318 (2017) 2211–2223.

[22] P. Rajpurkar, J. Irvin, R.L. Ball, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C.P. Langlotz, B.N. Patel, K.W. Yeom, K. Shpanskaya, F.G. Blankenberg, J. Seekins, T.J. Amrhein, D.A. Mong, S.S. Halabi, E.J. Zucker, A.Y. Ng, M.P. Lungren, Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists, PLoS Med 15 (2018) e1002686.

[23] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi, G.H. Tison, C. Bourn, M.P. Turakhia, A.Y. Ng, Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nat Med 25 (2019) 65–69.

[24] W.C. Liu, C.S. Lin, C.S. Tsai, T.P. Tsao, C.C. Cheng, J.T. Liou, W.S. Lin, S.M. Cheng, Y.S. Lou, C.C. Lee, C. Lin, A Deep-Learning Algorithm for Detecting Acute Myocardial Infarction, EuroIntervention (2021).

[25] C.D. Galloway, A.V. Valys, J.B. Shreibati, D.L. Treiman, F.L. Petterson, V.P. Gundotra, D.E. Albert, Z.I. Attia, R.E. Carter, S.J. Asirvatham, M.J. Ackerman, P.A. Noseworthy, J.J. Dillon, P.A. Friedman, Development and Validation of a Deep-Learning Model to Screen for Hyperkalemia From the Electrocardiogram, JAMA Cardiol 4 (2019) 428–436.

[26] C.S. Lin, C. Lin, W.H. Fang, C.J. Hsu, S.J. Chen, K.H. Huang, W.S. Lin, C.S. Tsai, C.C. Kuo, T. Chau, S.J. Yang, S.H. Lin, A Deep-Learning Algorithm (ECG12Net) for Detecting Hypokalemia and Hyperkalemia by Electrocardiography: Algorithm Development, JMIR Med Inform 8 (2020) e15931.

[27] Z.I. Attia, S. Kapa, X. Yao, F. Lopez-Jimenez, T.L. Mohan, P.A. Pellikka, R.E. Carter, N.D. Shah, P.A. Friedman, P.A. Noseworthy, Prospective validation of a deep learning electrocardiogram algorithm for the detection of left ventricular systolic dysfunction, Journal of cardiovascular electrophysiology 30 (2019) 668–674.

[28] Z.I. Attia, S. Kapa, F. Lopez-Jimenez, P.M. McKie, D.J. Ladewig, G. Satam, P.A. Pellikka, M. Enriquez-Sarano, P.A. Noseworthy, T.M. Munger, S.J. Asirvatham, C.G. Scott, R.E. Carter, P.A. Friedman, Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram, Nat Med 25 (2019) 70–74.

[29] S. Raghunath, A.E. Ulloa Cerna, L. Jing, D.P. vanMaanen, J. Stough, D.N. Hartzel, J.B. Leader, H.L. Kirchner, M.C. Stumpe, A. Hafez, A. Nemani, T. Carbonati, K.W. Johnson, K. Young, C.W. Good, J.M. Pfeifer, A.A. Patel, B.P. Delisle, A. Al-

said, D. Beer, C.M. Haggerty, B.K. Fornwalt, Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network, Nat Med 26 (2020) 886–891.
[30] J.M. Kwon, Y. Cho, K.H. Jeon, S. Cho, K.H. Kim, S.D. Baek, S. Jeung, J. Park, B.H. Oh, A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study, Lancet Digit Health 2 (2020) e358–e367.
[31] Z.I. Attia, P.A. Noseworthy, F. Lopez-Jimenez, S.J. Asirvatham, A.J. Deshmukh, B.J. Gersh, R.E. Carter, X. Yao, A.A. Rabinstein, B.J. Erickson, S. Kapa, P.A. Friedman, An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction, Lancet 394 (2019) 861–867.
[32] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J. van der Laak, B. van Ginneken, C.I. Sanchez, A survey on deep learning in medical image analysis, Med Image Anal 42 (2017) 60–88.
[33] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805, (2018).
[34] A.N. Uwaechia, D.A. Ramli, A Comprehensive Survey on ECG Signals as New Biometric Modality for Human Authentication: Recent Advances and Future Challenges, IEEE Access 9 (2021) 97760–97802.
[35] R. Hoekema, G.J. Uijen, A. van Oosterom, Geometrical aspects of the interindividual variability of multilead ECG recordings, IEEE transactions on bio-medical engineering 48 (2001) 551–559.
[36] G.A. Tadesse, T. Zhu, Y. Liu, Y. Zhou, J. Chen, M. Tian, D. Clifton, Cardiovascular disease diagnosis using cross-domain transfer learning, Annu Int Conf IEEE Eng Med Biol Soc 2019 (2019) 4262–4265.
[37] P. Gopika, V. Sowmya, E.A. Gopalakrishnan, K.P. Soman, 12 - Transferable approach for cardiac disease classification using deep learning, in: B. Agarwal, V.E. Balas, L.C. Jain, R.C. Poonia, Manisha (Eds.), Deep Learning Techniques for Biomedical and Health Informatics, Academic Press, 2020, pp. 285–303.
[38] J.H. Jang, T.Y. Kim, D. Yoon, Effectiveness of Transfer Learning for Deep Learning-Based Electrocardiogram Analysis, Healthcare informatics research 27 (2021) 19–28.
[39] K. Weimann, T.O.F. Conrad, Transfer learning for ECG classification, Sci Rep 11 (2021) 5251.
[40] T.R. Wei, S. Lu, Y. Yan, Automated Atrial Fibrillation Detection with ECG, Bioengineering (Basel) 9 (2022).
[41] S. Hong, Y. Zhou, J. Shang, C. Xiao, J. Sun, Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review, Comput Biol Med 122 (2020) 103801.
[42] M. Romero, Y. Interian, T. Solberg, G. Valdes, Targeted transfer learning to improve performance in small medical physics datasets, Medical physics 47 (2020) 6246–6256.
[43] Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE transactions on pattern analysis and machine intelligence 35 (2013) 1798–1828.
[44] S. Akcay, A. Atapour-Abarghouei, T.P. Breckon, in: Ganomaly: Semi-supervised anomaly detection via adversarial training, Asian conference on computer vision, Springer, 2018, pp. 622–637.
[45] L. Shan, Y. Li, H. Jiang, P. Zhou, J. Niu, R. Liu, Y. Wei, J. Peng, H. Yu, X. Sha, S. Chang, Abnormal ECG detection based on an adversarial autoencoder, Front Physiol 13 (2022) 961724.
[46] M. Sabuhi, M. Zhou, C.P. Bezemer, P. Musilek, Applications of Generative Adversarial Networks in Anomaly Detection: A Systematic Literature Review, IEEE Access 9 (2021) 161003–161029.
[47] Y. Yazici, C.S. Foo, S. Winkler, K.H. Yap, V. Chandrasekhar, Empirical Analysis Of Overfitting And Mode Drop In Gan Training, in: 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 1651–1655.
[48] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi, G.H. Tison, C. Bourn, M.P. Turakhia, A.Y. Ng, Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nature Medicine 25 (2019) 65–69.
[49] S.W. Smith, B. Walsh, K. Grauer, K. Wang, J. Rapin, J. Li, W. Fennell, P. Taboulet, A deep neural network learning algorithm outperforms a conventional algorithm for emergency department electrocardiogram interpretation, Journal of Electrocardiology 52 (2019) 88–95.
[50] J.M. Kwon, K.H. Kim, J. Medina-Inojosa, K.H. Jeon, J. Park, B.H. Oh, Artificial intelligence for early prediction of pulmonary hypertension using electrocardiography, J Heart Lung Transplant 39 (2020) 805–814.
[51] J.M. Kwon, S.Y. Lee, K.H. Jeon, Y. Lee, K.H. Kim, J. Park, B.H. Oh, M.M. Lee, Deep Learning–Based Algorithm for Detecting Aortic Stenosis Using Electrocardiography, Journal of the American Heart Association 9 (2020) e014717.
[52] P. Elias, T.J. Poterucha, V. Rajaram, L.M. Moller, V. Rodriguez, S. Bhave, R.T. Hahn, G. Tison, S.A. Abreau, J. Barrios, J.N. Torres, J.W. Hughes, M.V. Perez, J. Finer, S. Kodali, O. Khalique, N. Hamid, A. Schwartz, S. Homma, D. Kumaraiah, D.J. Cohen, M.S. Maurer, A.J. Einstein, T. Nazif, M.B. Leon, A.J. Perotte, Deep Learning Electrocardiographic Analysis for Detection of Left-Sided Valvular Heart Disease, Journal of the American College of Cardiology 80 (2022) 613–626.
[53] W.Y. Ko, K.C. Siontis, Z.I. Attia, R.E. Carter, S. Kapa, S.R. Ommen, S.J. Demuth, M.J. Ackerman, B.J. Gersh, A.M. Arruda-Olson, J.B. Geske, S.J. Asirvatham, F. Lopez-Jimenez, R.A. Nishimura, P.A. Friedman, P.A. Noseworthy, Detection of Hypertrophic Cardiomyopathy Using a Convolutional Neural Network-Enabled Electrocardiogram, J Am Coll Cardiol 75 (2020) 722–733.
[54] Y.L. Liu, C.S. Lin, C.C. Cheng, C. Lin, A Deep Learning Algorithm for Detecting Acute Pericarditis by Electrocardiogram, J Pers Med 12 (2022).
[55] C. Lin, T. Chau, C.-S. Lin, H.-S. Shang, W.-H. Fang, D.-J. Lee, C.-C. Lee, S.-H. Tsai, C.-H. Wang, S.-H. Lin, Point-of-care artificial intelligence-enabled ECG for dyskalemia: a retrospective cohort analysis for accuracy and outcome prediction, npj Digital Medicine 5 (2022) 8.
[56] K.C. Siontis, P.A. Noseworthy, Z.I. Attia, P.A. Friedman, Artificial intelligence-enhanced electrocardiography in cardiovascular disease management, Nat Rev Cardiol (2021).
[57] T. Stracina, M. Ronzhina, R. Redina, M. Novakova, Golden Standard or Obsolete Method? Review of ECG Applications in Clinical and Experimental Context, Front Physiol 13 (2022) 867033.
[58] S. Somani, A.J. Russak, F. Richter, S. Zhao, A. Vaid, F. Chaudhry, J.K. De Freitas, N. Naik, R. Miotto, G.N. Nadkarni, J. Narula, E. Argulian, B.S. Glicksberg, Deep learning and the electrocardiogram: review of the current state-of-the-art, Europace 23 (2021) 1179–1191.
[59] G. Varoquaux, V. Cheplygina, Machine learning for medical imaging: methodological failures and recommendations for the future, npj Digital Medicine 5 (2022) 48.
[60] E.M. Lima, A.H. Ribeiro, G.M.M. Paixão, M.H. Ribeiro, M.M. Pinto-Filho, P.R. Gomes, D.M. Oliveira, E.C. Sabino, B.B. Duncan, L. Giatti, S.M. Barreto, W. Meira Jr, T.B. Schön, A.L.P. Ribeiro, Deep neural network-estimated electrocardiographic age as a mortality predictor, Nature Communications 12 (2021) 5117.
[61] K. Kusunose, Y. Hirata, T. Tsuji, J.i. Kotoku, M. Sata, Deep learning to predict elevated pulmonary artery pressure in patients with suspected pulmonary hypertension using standard chest X ray, Scientific Reports 10 (2020) 19311.
[62] V.K. Raghu, J. Weiss, U. Hoffmann, H. Aerts, M.T. Lu, Deep Learning to Estimate Biological Age From Chest Radiographs, JACC Cardiovasc Imaging 14 (2021) 2226–2236.
[63] C. Li, Y. Qin, W.H. Zhang, H. Jiang, B. Song, M.R. Bashir, H. Xu, T. Duan, M. Fang, L. Zhong, L. Meng, D. Dong, Z. Hu, J. Tian, J.K. Hu, Deep learning-based AI model for signet-ring cell carcinoma diagnosis and chemotherapy response prediction in gastric cancer, Med Phys 49 (2022) 1535–1546.
[64] E. Nilsson, A. Gasparini, J. Arnlov, H. Xu, K.M. Henriksson, J. Coresh, M.E. Grams, J.J. Carrero, Incidence and determinants of hyperkalemia and hypokalemia in a large healthcare system, International journal of cardiology 245 (2017) 277–284.
[65] C.P. Kovesdy, K. Matsushita, Y. Sang, N.J. Brunskill, J.J. Carrero, G. Chodick, T. Hasegawa, H.L. Heerspink, A. Hirayama, G.W.D. Landman, A. Levin, D. Nitsch, D.C. Wheeler, J. Coresh, S.I. Hallan, V. Shalev, M.E. Grams, Serum potassium and adverse outcomes across the range of kidney function: a CKD Prognosis Consortium meta-analysis, European heart journal 39 (2018) 1535–1542.
[66] S. Sabour, N. Frosst, G.E. Hinton, Dynamic routing between capsules, Advances in neural information processing systems (2017) 3856–3866.
[67] J.M. Kwon, K.H. Kim, K.H. Jeon, H.M. Kim, M.J. Kim, S.M. Lim, P.S. Song, J. Park, R.K. Choi, B.H. Oh, Development and Validation of Deep-Learning Algorithm for Electrocardiography-Based Heart Failure Identification, Korean circulation journal 49 (2019) 629–639.
[68] R.R. Lopes, H. Bleijendaal, L.A. Ramos, T.E. Verstraelen, A.S. Amin, A.A.M. Wilde, Y.M. Pinto, B.A.J.M. de Mol, H.A. Marquering, Improving electrocardiogram-based detection of rare genetic heart disease using transfer learning: An application to phospholamban p.Arg14del mutation carriers, Computers in Biology and Medicine 131 (2021) 104262.
[69] B. Lu, H.-X. Li, Z.-K. Chang, L. Li, N.-X. Chen, Z.-C. Zhu, H.-X. Zhou, X.-Y. Li, Y.-W. Wang, S.-X. Cui, Z.-Y. Deng, Z. Fan, H. Yang, X. Chen, P.M. Thompson, F.X. Castellanos, C.-G. Yan, A practical Alzheimer's disease classifier via brain imaging-based deep learning on 85,721 samples, Journal of Big Data 9 (2022) 101.
[70] G.J. Wehner, L. Jing, C.M. Haggerty, J.D. Suever, J.B. Leader, D.N. Hartzel, H.L. Kirchner, J.N.A. Manus, N. James, Z. Ayar, P. Gladding, C.W. Good, J.G.F. Cleland, B.K. Fornwalt, Routinely reported ejection fraction and mortality in clinical practice: where does the nadir of risk lie? Eur Heart J 41 (2020) 1249–1257.
[71] S. Zanelli, M. Ammi, M. Hallab, M.A. El Yacoubi, Diabetes Detection and Management through Photoplethysmographic and Electrocardiographic Signals Analysis: A Systematic Review, Sensors (Basel) 22 (2022).
[72] B. Babenko, A. Mitani, I. Traynis, N. Kitade, P. Singh, A.Y. Maa, J. Cuadros, G.S. Corrado, L. Peng, D.R. Webster, A. Varadarajan, N. Hammel, Y. Liu, Detection of signs of disease in external photographs of the eyes via deep learning, Nature Biomedical Engineering (2022).
[73] J.J. Dutton, Anatomic Considerations in Thyroid Eye Disease, Ophthalmic Plast Reconstr Surg 34 (2018) S7–s12.
[74] M. Christoffersen, R. Frikke-Schmidt, P. Schnohr, G.B. Jensen, B.G. Nordestgaard, A. Tybjærg-Hansen, Xanthelasmata, arcus corneae and ischaemic vascular disease and death in general population: prospective cohort study, BMJ 343 (2011) d5497.
[75] X. Yao, D.R. Rushlow, J.W. Inselman, R.G. McCoy, T.D. Thacher, E.M. Behnken, M.E. Bernard, S.L. Rosas, A. Akfaly, A. Misra, P.E. Molling, J.S. Krien, R.M. Foss, B.A. Barry, K.C. Siontis, S. Kapa, P.A. Pellikka, F. Lopez-Jimenez, Z.I. Attia, N.D. Shah, P.A. Friedman, P.A. Noseworthy, Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial, Nat Med 27 (2021) 815–819.
[76] P.A. Noseworthy, Z.I. Attia, L.C. Brewer, S.N. Hayes, X. Yao, S. Kapa, P.A. Friedman, F. Lopez-Jimenez, Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a Deep Learning Model for ECG Analysis, Circ Arrhythm Electrophysiol 13 (2020) e007988.