CLINICAL RESEARCH

Improved prediction of sudden cardiac death in patients with heart failure through digital processing of electrocardiography

Yasuyuki Shiraishi [1]*[†], Shinichi Goto [2][†], Nozomi Niimi [1], Yoshinori Katsumata [3]*, Ayumi Goda [4], Makoto Takei [5], Mike Saji [6], Motoaki Sano[1], Keiichi Fukuda [1], Takashi Kohno [4], Tsutomu Yoshikawa[6], and Shun Kohsaka [1]

[1]Department of Cardiology, Keio University School of Medicine, 35 Shinanomachi Shinjuku-ku, Tokyo 160-8582, Japan; [2]One Brave Idea and Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA; [3]Institute for Integrated Sports Medicine, Keio University School of Medicine, Tokyo, Japan; [4]Department of Cardiovascular Medicine, Kyorin University Faculty of Medicine, Tokyo, Japan; [5]Department of Cardiology, Saiseikai Central Hospital, Tokyo, Japan; and [6]Department of Cardiology, Sakakibara Heart Institute, Tokyo, Japan

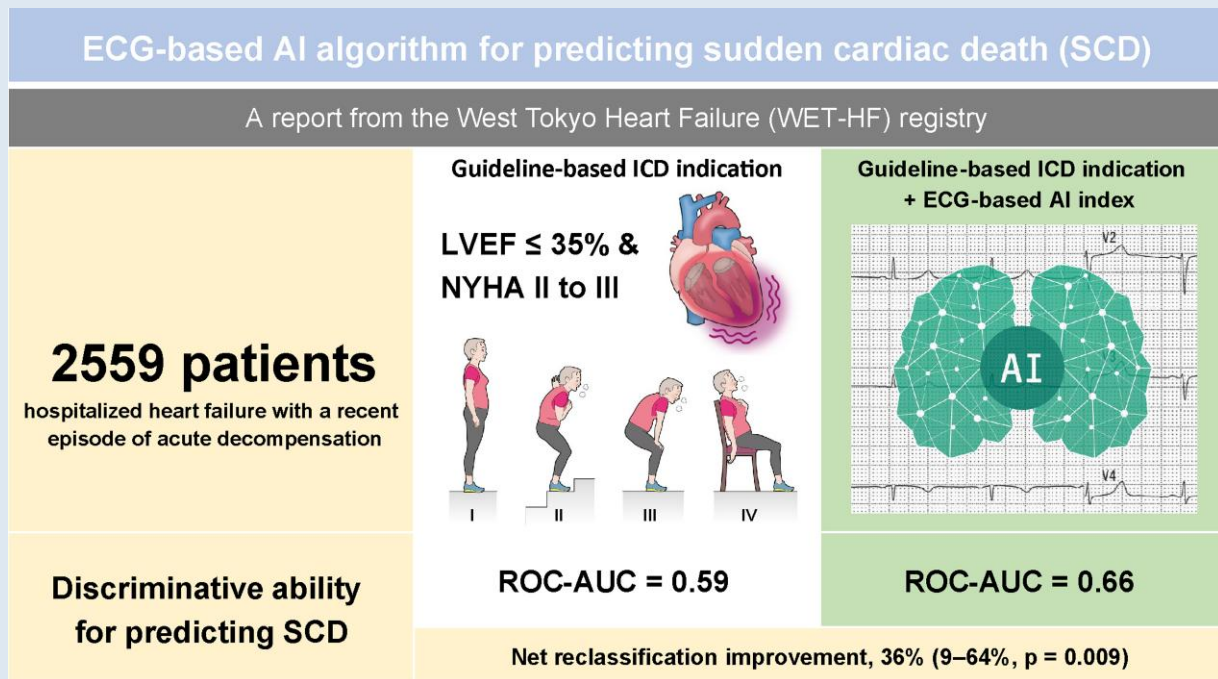|	
Aims	Available predictive models for sudden cardiac death (SCD) in heart failure (HF) patients remain suboptimal. We assessed whether the electrocardiography (ECG)-based artificial intelligence (AI) could better predict SCD, and also whether the combination of the ECG-AI index and conventional predictors of SCD would improve the SCD stratification among HF patients.
Methods and results	In a prospective observational study, 4 tertiary care hospitals in Tokyo enrolled 2559 patients hospitalized for HF who were successfully discharged after acute decompensation. The ECG data during the index hospitalization were extracted from the hospitals' electronic medical record systems. The association of the ECG-AI index and SCD was evaluated with adjustment for left ventricular ejection fraction (LVEF), New York Heart Association (NYHA) class, and competing risk of non-SCD. The ECG-AI index plus classical predictive guidelines (i.e. LVEF ≤35%, NYHA Class II and III) significantly improved the discriminative value of SCD [receiver operating characteristic area under the curve (ROC-AUC), 0.66 vs. 0.59; $P = 0.017$; Delong's test] with good calibration ($P = 0.11$; Hosmer–Lemeshow test) and improved net reclassification [36%; 95% confidence interval (CI), 9–64%; $P = 0.009$]. The Fine–Gray model considering the competing risk of non-SCD demonstrated that the ECG-AI index was independently associated with SCD (adjusted sub-distributional hazard ratio, 1.25; 95% CI, 1.04–1.49; $P = 0.015$). An increased proportional risk of SCD vs. non-SCD with an increasing ECG-AI index was also observed (low, 16.7%; intermediate, 18.5%; high, 28.7%; P for trend = 0.023). Similar findings were observed in patients aged ≤75 years with a non-ischaemic aetiology and an LVEF of >35%.
Conclusion	To improve risk stratification of SCD, ECG-based AI may provide additional values in the management of patients with HF.

* Corresponding authors. Tel: +81 3 3353 1211; fax: +81 5 5843 6167. *E-mail address:* yasshiraishi@keio.jp (YS); Tel: +81 3 3353 1211; fax: +81 5 5843 6167. *E-mail address:* goodcentury21@gmail.com (YK)
[†] These authors contributed equally to the study.

Graphical Abstract

ECG-based AI algorithm for predicting sudden cardiac death (SCD)

A report from the West Tokyo Heart Failure (WET-HF) registry

2559 patients hospitalized heart failure with a recent episode of acute decompensation

Guideline-based ICD indication
LVEF ≤ 35% & NYHA II to III

Guideline-based ICD indication + ECG-based AI index

Discriminative ability for predicting SCD

ROC-AUC = 0.59

ROC-AUC = 0.66

Net reclassification improvement, 36% (9–64%, p = 0.009)

What's new?

- Current approaches for the assessment of the risk of sudden cardiac death (SCD) are mainly based on left ventricular ejection fraction (LVEF; ≤ 35%) and functional status, and remain suboptimal.
- SCDs also occur frequently in patients without severe left ventricular dysfunction, and there is currently no risk stratification algorithm for SCD in patients with heart failure (HF) with mildly reduced or preserved ejection fraction (LVEF >35%).
- This observational study showed that electrocardiography-based artificial intelligence significantly improved the discrimination and reclassification of SCD in patients with HF when added to the conventional guideline-directed indication for implantable cardioverter-defibrillator.
- In particular, the predictive ability was observed in patients aged ≤ 75 years with a non-ischaemic aetiology and an LVEF of >35%.

Introduction

Heart failure (HF) is associated with poor quality of life and premature death.[1] At present, the incidence of HF is high in both Asian (1.2–6.7%) and Western countries (1–14%).[2–4] As the prevalence is expected to increase, HF is expected to generate a substantial global public health burden.

Sudden cardiac death (SCD), typically caused by lethal arrhythmias, is responsible for ~50% of all cardiovascular deaths in patients with HF.[5] Although implantable cardioverter-defibrillators (ICDs) are used to reduce the risk of SCD, the implantation procedure is invasive, and ~50% of patients with ICD implantation experience inappropriate shocks, reducing the quality of life and increasing mortality rates.[6] Therefore, the accurate assessment of SCD risk in patients with HF is paramount for clinical decision-making to ensure appropriate device application.

Current approaches to assessing SCD risk are mainly based on left ventricular ejection fraction (LVEF ≤ 35%) and the New York Heart Association (NYHA) functional classification and remain suboptimal,[7,8] resulting in ICD over- and underuse.[9] Furthermore, despite the fact that SCD also occurs frequently in patients without severe left ventricular dysfunction, there is currently no risk stratification algorithm for SCD in patients with an LVEF of >35%.[10] Artificial intelligence (AI) is a promising technology for deriving a statistical model from information-rich yet complex datasets. Electrocardiography-based AI (ECG-AI) models constructed by convolutional neural networks (CNNs) have shown promise for detecting diseases,[11] predicting cardiac function,[12] and estimating prognosis.[13]

We aimed to investigate whether AI models trained on 12-lead ECG will enable the detection of important features for classifying the risk of SCD and improve the risk stratification of patients with HF. The ability of ECG-based AI to predict the incidence of SCD was tested in multicentre HF registry data that captured patients hospitalized for acute decompensation and were successfully discharged.

Methods

Ethical approval/informed consent

According to the Ethical Guidelines for Medical and Health Research Involving Human Subjects and Personal Information Protection Law in Japan, verbal informed consent was obtained from all participants before the study enrolment. All patient information was recorded using an electronic data-capturing system in a de-anonymized form.

Data source and data collection

Using data from a prospective observational study [the West Tokyo Heart Failure Registry (WET-HF)], we analyzed hospitalized patients with HF who required urgent treatment for acute decompensation in four tertiary care hospitals (Keio University Hospital, Kyorin University Hospital, Sakakibara Heart Institute, and St. Luke's International Hospital) within the metropolitan area of Tokyo, Japan.[8] The WET-HF was launched in January 2006 and consisted of six tertiary care hospitals as of December 2017. To ensure a robust assessment of the care and outcomes, baseline data and outcome measures were collected from the medical records. To ensure the accuracy and completeness of the information, the treating physicians were contacted by dedicated clinical research coordinators. Data were entered into an electronic data-capturing system, which has a robust data query engine and system validations for data quality. Exclusive on-site auditing by the investigators (Y.S. and S.K.) ensured the proper registration of each patient. The objective and detailed design was provided by the University Hospital Medical Information Network (UMIN000001171). The study protocol was approved by the ethical review committee of each centre, and the study was conducted in accordance with the Declaration of Helsinki.

Patient demographics, medical history, laboratory and other tests (such as electrocardiogram and echocardiogram), medications, procedures, and clinical outcomes during hospitalization and after discharge with a minimum 2-year follow-up were recorded. The diagnosis of acute HF was made at the discretion of the treating cardiologists at each institution based on the Framingham criteria. In this cohort, NYHA functional class was evaluated at discharge by individual cardiologists at each institution and reviewed by the investigators (Y.S., S.K., T.K., Y.N., A.G., and T.Y.). Left ventricular ejection fraction on echocardiography was assessed using the modified Simpson's method during the index hospitalization after stabilization of the HF signs and symptoms.

All 12-lead ECG data of the patients were reported and extracted from the electronic medical record system of each institution. All the participating hospitals used the same vendor for ECGs (Nihon Koden, Tokyo, Japan) throughout the study period. Electrode placement is based on a standard 12-lead ECG in all hospitals, which has 15 voltage-time traces, including those 2.5 s in duration for all 12 leads and those 10 s in duration for leads V1, II, and V5. The ECG data were stored as measurements of the time-series voltage at a sampling rate of 500 Hz. Of the six hospitals, two were not included in the present study because the ECG data could not be extracted from the database. Patients who were registered in the remaining hospitals underwent ECG at the time of discharge.

Study cohort

The patient selection and exclusion criteria as well as the group allocations are shown in Supplementary material online, *Figure S1*. A total of 2559 patients who were successfully discharged between 2006 and 2017 were included in the present analysis. Patients were allocated to one of three cohorts (derivation, validation, or test) based on the hospital of recruitment. The ECG data of the patients at the time of discharge were reported and extracted from the electronic medical record systems. When multiple ECG data were available for the same patient, the latest data were included in the analysis.

Model training

The AI model to predict 3-year SCD was constructed using a combination of CNN and long short-term memory (LSTM), a variant of recurrent neural network (RNN) (see Supplementary material online, *Figure S2*). The details of the architecture have been published previously.[14] In brief, the neural network is a mathematical model that is constructed with units that simulate the function of neuron cells. The unit takes multiple inputs and multiplies the input with internal weights. After summing the results, the number is inputted into a function called the activation function. The training process adjusts the internal weights such that the output approaches the given label. The architecture used in the present study consists of a neural network that stacks up four layers of one-dimensional CNN suitable for detecting 'shape patterns' with a relatively low computational cost and two layers of LSTM suitable for learning time-series data in detail with a relatively high computational cost. The model was trained with the data from the derivation cohort to minimize the binary cross-entropy loss with the RMSProp optimizer with an initial learning rate of 0.0001. The model's performance

was calculated using data from the validation dataset at the end of each epoch. The final model was chosen as the one that performed best for 50 epochs in the validation cohort (see Supplementary material online, *Figure S3*). To ensure that the model works on data that were never seen during training and the model selection procedure, the performance of the final model (i.e. the ECG-AI index) was calculated only once using data from the test cohort. Finally, we implemented gradient-weighted class activation mapping (Grad-CAM) to identify which regions in ECG were based on the prediction of the neural network model (see Supplementary material online, *Figure S4*). The model was trained using Tensorflow framework version 2.2.0 with Python version 3.6.8.

Ascertainment and classification of sudden cardiac death or implantable cardioverter-defibrillator events

The outcome measure was the composite of SCD and ICD activation (i.e. both shock and anti-tachycardia pacing). To ensure SCD assessment accuracy, the WET-HF registry was supported by a central study committee that adjudicated the mode of death. All deaths were reviewed by the investigators and then categorized into those in need of adjudication or those in which the mode of death could be defined clearly. Central committee members (Y.S., S.K., T.K., Y.N., A.G., and T.Y.) reviewed the abstracted records and adjudicated modes of death. SCD was defined as unexpected and otherwise unexplained death in a previously stable patient or death from documented or presumed cardiac arrhythmia without a clear non-cardiovascular cause, including patients who were comatose and then died after attempted resuscitation.[15] Patients who died and had been out of contact for more than 24 h were classified as 'unknown death'.[15] All other causes of death were classified as non-SCD. In addition, appropriate ICD activation was ascertained through device interrogation during regular check-ups or at a suspected instance of an arrhythmic episode and SCD.

Statistical analysis

With respect to descriptive statistics, continuous variables are presented as medians and interquartile ranges, while categorical variables are presented as frequencies and percentages. For baseline characteristics, the three cohorts (i.e. derivation, validation, and test) were compared using the Kruskal–Wallis rank sum test for continuous variables and the χ^2 test or Fisher's exact test for categorical variables as appropriate.

The discriminative ability of the conventional guideline-directed indication for ICD (i.e. LVEF ≤35% and NYHA Class II and III) or its combination with the ECG-AI index (ECG-AI combined model) for predicting composite SCD events over 3 years was evaluated using the receiver operating characteristic area under the curve (ROC-AUC) with logistic regression analysis and the pairwise Delong's test. The model's calibration performance was assessed by comparing the predicted and observed probabilities for the four groups using the Hosmer–Lemeshow test. The model's reclassification of the composite SCD events was also assessed with net reclassification improvement using the 3-year estimated probabilities of the composite SCD events.

Next, the model's discriminative abilities were assessed in the pre-specified subgroups [i.e. age (≥ 75 vs. <75 years), sex (male vs. female), aetiology (ischaemic vs. non-ischaemic), and presence or absence of atrial fibrillation (AF)]. In the subset of patients divided by LVEF, the discriminative ability of the ECG-AI index, as well as the frequencies of composite SCD vs. non-SCD events, was assessed separately: ≤35%, 35–50%, and ≥50%. Furthermore, we performed sensitivity analyses using several LVEF cut-off values (45, 55, and 60%).

For the survival analysis to assess time-to-event, we first evaluated the cumulative incidence of composite SCD and non-SCD events using the Aalen–Johansen estimator divided by the risk of the conventional guideline-directed indication for ICD and the ECG-AI combined model. The optimal thresholds for risk categories (low vs. high risk) for the ICD indication and the ECG-AI combined model were determined using the Youden index. Next, we examined the association between the ECG-AI index and composite SCD events using univariate and multivariate Fine and Gray models, which estimated the incidence of outcomes over time and calculated subdistributional hazard ratios, accounting for the competing risk of

Table 1 Patient backgrounds according to derivation, validation, and test cohorts

Value	Derivation cohort (n = 1028)	Validation cohort (n = 454)	Test cohort (n = 1077)	P-value
Age, years	78 (69–84)	73 (60–81)	77 (67–84)	<0.001
Men, n (%)	606 (58.9)	288 (63.4)	638 (59.2)	0.400
Body mass index, kg/m^2	21.5 (19.3–24.1)	21.8 (19.5–24.1)	21.1 (18.7–23.8)	<0.001
Systolic blood pressure, mmHg	110 (100–120)	108 (97–120)	113 (100–128)	<0.001
Heart rate, bpm	70 (60–78)	71 (61–80)	70 (61–80)	<0.001
LVEF, %	48 (32–59)	40 (30–56)	48 (34–59)	<0.001
LVEF ≤ 35%, n (%)	321 (31.5)	209 (46.2)	290 (26.9)	<0.001
NYHA Class II and III, n (%)	771 (75.2)	416 (92.4)	990 (93.4)	<0.001
Ischaemic aetiology, n (%)	279 (27.1)	115 (25.3)	305 (28.3)	0.006
Previous HF hospitalization	337 (32.7)	146 (31.7)	257 (27.5)	<0.001
Atrial fibrillation, n (%)	607 (59.0)	210 (46.1)	428 (39.7)	<0.001
Hypertension, n (%)	653 (63.5)	266 (58.5)	766 (71.1)	<0.001
Diabetes mellitus, n (%)	296 (28.8)	145 (31.8)	428 (39.7)	<0.001
Stroke, n (%)	140 (12.9)	66 (14.2)	141 (13.1)	0.552
COPD, n (%)	27 (2.5)	42 (7.4)	67 (6.0)	<0.001
Haemoglobin, g/dL	11.9 (10.5–13.3)	12.5 (10.7–14.3)	11.9 (10.6–13.4)	<0.001
Blood urea nitrogen, mg/L	20.1 (15.2–29.4)	23.9 (18.0–32.2)	24.4 (17.9–35.2)	<0.001
eGFR, mL/min/1.73 m^2	51.8 (37.7–64.5)	50.3 (36.3–62.4)	49.2 (30.9–66.0)	0.001
Sodium, mEq/L	139 (137–141)	139 (137–141)	139 (136–141)	0.522
Potassium, mEq/L	4.3 (4.0–4.6)	4.4 (4.1–4.7)	4.3 (4.0–4.7)	<0.001
Uric acid, mg/L	6.6 (5.3–7.8)	6.9 (5.6–8.2)	7.2 (5.7–8.6)	<0.001
Albumin, mg/L	3.6 (3.3–3.9)	3.7 (3.3–4.0)	3.4 (3.0–3.7)	<0.001
BNP, pg/mL*	N/A	244 (116–479)	247 (128–508)	
NT-proBNP, pg/mL*	1830 (1109–3612)	N/A	N/A	
Loop diuretics, n (%)	820 (79.8)	315 (69.2)	821 (76.2)	<0.001
ACEi or ARB, n (%)	640 (62.3)	315 (69.2)	687 (63.8)	0.018
Beta-blocker, n (%)	788 (76.7)	360 (79.2)	815 (75.7)	0.304
MRA, n (%)	307 (29.9)	193 (34.9)	378 (34.9)	<0.001
Digitalis, n (%)	73 (7.1)	54 (7.0)	77 (7.0)	0.011
Amiodarone, n (%)	75 (7.2)	42 (8.8)	114 (10.5)	0.044
ICD, n (%)	52 (5.1)	38 (8.4)	32 (3.0)	<0.001
CRT, n (%)	19 (1.9)	32 (7.1)	13 (1.2)	<0.001

*BNP was measured in the validation and test cohorts, while NT-proBNP was measured in the derivation cohort.
LVEF, left ventricular ejection fraction; NYHA, New York Heart Association; HF, heart failure; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate; BNP, B-type natriuretic peptide; NT-proBNP, N-terminal pro-B-type natriuretic peptide; ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; MRA, mineralocorticoid receptor antagonist; ICD, implantable cardioverter-defibrillator; CRT, cardiac resynchronization therapy.

non-SCD. To assess the model's discrimination of the survival analysis, we used Harrel's c-statistics for the Fine and Gray model. We transformed the ECG-AI index (the output of the neural network model) to be standardized (mean = 0; standardized deviation = 1). Finally, a survival analysis that did not consider the competing risk was also performed with the time-dependent ROC-AUC.

To visualize the proportional risk of composite SCD vs. non-SCD events, we divided the patients into three groups by tertile of predicted risk by the ECG-AI index alone and the ECG-AI combined model and calculated the prevalence of SCD and non-SCD events by group using a trend test. All analyses were conducted using the tidyverse, tidymodels, pROC, PredictABEL, survminer, cmprsk, riskRegression, and survival packages of R version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria, 2008).

Results

Patient characteristics

The baseline characteristics of the derivation, validation, and test cohorts are shown in *Table 1*. The median patient age and LVEF were 73–78 years and 40–48%, respectively. The proportions of patients with NYHA functional Class II and III were 75.2, 92.4, and 93.4% in the derivation, validation, and test cohorts, respectively. Medical therapies for HF were similarly implemented during the index hospitalization across each cohort: angiotensin-converting enzyme inhibitors or angiotensin receptor blockers, 62.3–69.2%; beta-blockers, 75.7–79.2%; and mineralocorticoid receptor antagonists, 29.9–34.9%. The
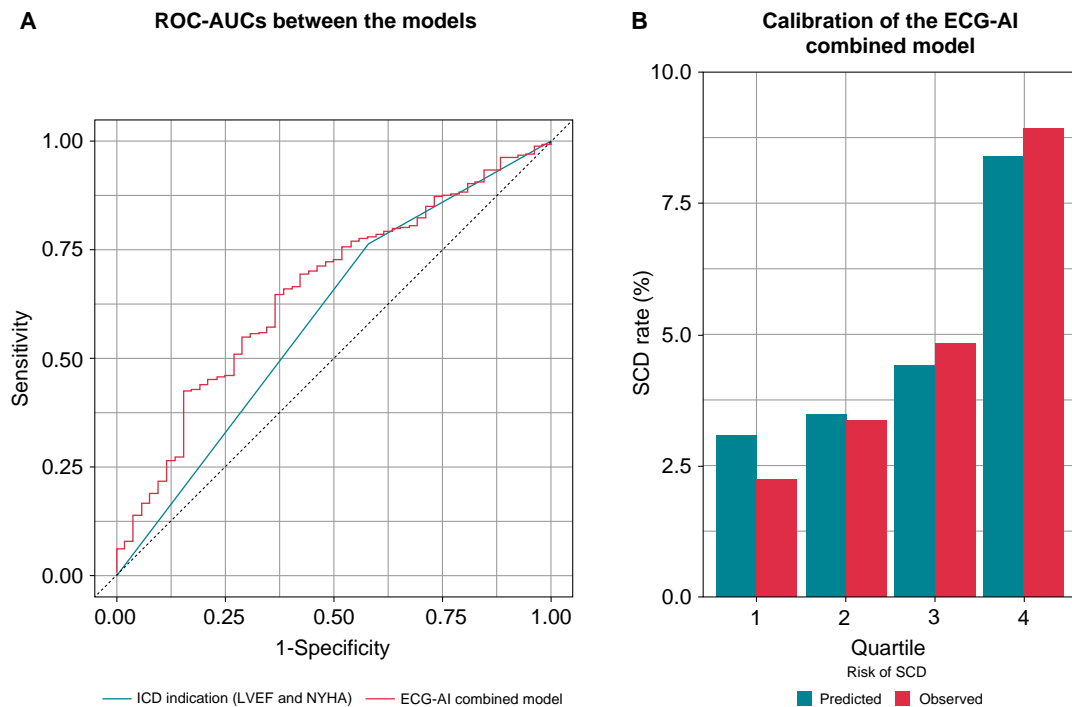
Figure 1 Discrimination and calibration of risk models for predicting SCD events. Comparison of ROC-AUC (*A*) shows the conventional guideline-directed indication for ICD (LVEF ≤ 35% and NYHA Class II and III [as a binary variable]; AUC = 0.59 [95% CI, 0.52–0.66]) vs. the ECG-AI combined model (ECG-AI index [as a continuous variable] + ICD indication; AUC = 0.66 [95% CI, 0.58–0.73]) for predicting 3-year composite SCD events (*P* = 0.017 for Delong's test). Calibration of the ECG-AI combined model is shown (*B*) by dividing four bins based on the quartiles of the model-predicted risk (*P* = 0.11 for the Hosmer–Lemeshow test). CI, confidence interval; ECG-AI, electrocardiography-based artificial intelligence; ICD, implantable cardioverter-defibrillator; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association; ROC-AUC, receiver operating characteristic area under the curve; SCD, sudden cardiac death.

prevalence of ICD implantation was low at 3.0–8.4%. Overall, there were 236 (21.9%) deaths [48 (20.3%) SCDs and 188 (79.7%) non-SCDs] and 4 ICD activations occurred in 1077 patients who were included in the test cohort in 3 years following the hospitalization events.

ECG-based artificial intelligence performance

The ROC curve analysis, without considering time-to-event, showed a good discriminative ability for predicting SCD over 3 years of the ECG-AI index with 0.62 [95% confidence interval (CI), 0.54–0.70] of ROC-AUC compared to the conventional guideline-directed indication for ICD. *Figure 1A* shows that the addition of the ECG-AI index to the conventional indication for ICD, inclusive of LVEF and NYHA class, significantly improved the discrimination [ROC-AUC = 0.66 (95% CI, 0.58–0.73) vs. 0.59 (95% CI, 0.52–0.66) with and without the ECG-AI index, *P* = 0.017, Delong's test] compared to the conventional indication for ICD alone [sensitivity = 0.635 and 0.423, and specificity = 0.648 and 0.760 (χ^2 test, 8.28, *P* < 0.001)]. The ECG-AI combined model showed good calibration (*Figure 1B*; *P* = 0.11, Hosmer–Lemeshow test). Furthermore, the addition of the ECG-AI index to the conventional indication for ICD improved the indices of reclassification [net reclassification improvement, 36% (9–64%; *P* = 0.009)].

Similar findings were observed in the pre-specified subgroup analyses (*Figure 2*). In particular, compared to the conventional indication for ICD, the ECG-AI combined model showed significantly better

discrimination of the incidence of SCD among younger patients (≤75 years) with a non-ischaemic aetiology and the absence of AF.

The proportion of patients with SCD vs. non-SCD steadily decreased as LVEF increased, and the ECG-AI index showed the best discriminative ability in patients with an LVEF of 35–50% (*Figure 3*). In sensitivity analyses using several LVEF cut-offs, the ECG-AI index showed the best performance for patients with LVEF 35–50%, although a good discriminative ability was also observed among those with an LVEF of 35–60% (see Supplementary material online, *Table S1*). In addition, another sensitivity analysis stratifying patients who met or did not meet the conventional indication for ICD showed a better discriminative ability of the ECG-AI index with 0.64 (95% CI, 0.53–0.75) of ROC-AUC in those who did not meet the ICD indication (*n* = 809) compared with those who met the ICD indication (ROC-AUC = 0.51 [95% CI, 0.38–0.65]).

Competing risk analysis for adjusting non-sudden cardiac death

The results of the survival analysis accounting for the competing risk of non-SCD (see Supplementary material online, *Table S2*) show that the ECG-AI index [adjusted sub-distributional hazard ratio (sHR), 1.23; 95% CI, 1.04–1.49; *P* = 0.015], as well as the conventional guideline-directed indication for ICD (adjusted sHR, 1.98; 95% CI, 1.11–3.54; *P* = 0.02), was independently and significantly associated with the risk of SCD using the Fine–Gray models. In this competing analysis, our new model combined with the conventional indication for ICD and
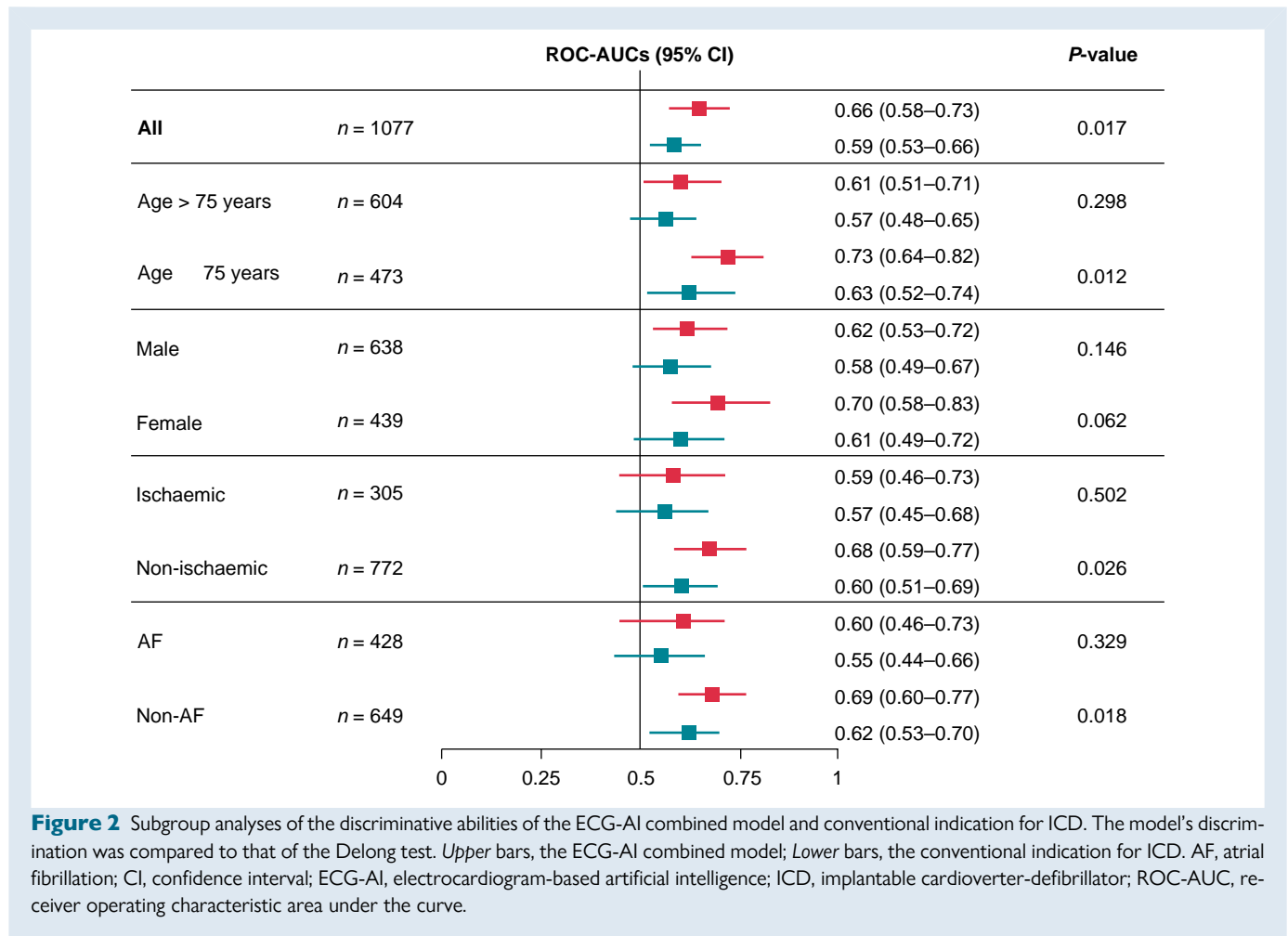
		ROC-AUCs (95% CI)		P-value
All	n = 1077		0.66 (0.58–0.73)	0.017
			0.59 (0.53–0.66)	
Age > 75 years	n = 604		0.61 (0.51–0.71)	0.298
			0.57 (0.48–0.65)	
Age ≤ 75 years	n = 473		0.73 (0.64–0.82)	0.012
			0.63 (0.52–0.74)	
Male	n = 638		0.62 (0.53–0.72)	0.146
			0.58 (0.49–0.67)	
Female	n = 439		0.70 (0.58–0.83)	0.062
			0.61 (0.49–0.72)	
Ischaemic	n = 305		0.59 (0.46–0.73)	0.502
			0.57 (0.45–0.68)	
Non-ischaemic	n = 772		0.68 (0.59–0.77)	0.026
			0.60 (0.51–0.69)	
AF	n = 428		0.60 (0.46–0.73)	0.329
			0.55 (0.44–0.66)	
Non-AF	n = 649		0.69 (0.60–0.77)	0.018
			0.62 (0.53–0.70)	

Figure 2 Subgroup analyses of the discriminative abilities of the ECG-AI combined model and conventional indication for ICD. The model's discrimination was compared to that of the Delong test. *Upper* bars, the ECG-AI combined model; *Lower* bars, the conventional indication for ICD. AF, atrial fibrillation; CI, confidence interval; ECG-AI, electrocardiogram-based artificial intelligence; ICD, implantable cardioverter-defibrillator; ROC-AUC, receiver operating characteristic area under the curve.

the ECG-AI index showed good discriminative ability, with a 0.65 concordance index (95% CI, 0.62–0.69) using a bootstrapping technique (500 sets).

In subgroup analyses, the ECG-AI index was also independently associated with composite SCD events among each subset of patients in the Fine–Gray competing risk model (see Supplementary material online, Figure S5). The cumulative incidence calculated via Aalen–Johansen estimates further demonstrates this relationship [see Supplementary material online, Figure S6, using both the ECG-AI index (A) and the ECG-AI combined model with LVEF and NYHA class (B)]. These results were similar to the time-dependent ROC-AUC without adjustment for the competing risk of non-SCD (see Supplementary material online, Figure S7).

Proportional risk of sudden cardiac death vs. non-sudden cardiac death using the ECG-based artificial intelligence models

Figure 4 shows that the ECG-AI index alone and the ECG-AI combined model could discriminate between SCD and non-SCD across the low-, intermediate-, and high-risk patient groups. We observed an increase in the proportional risk of SCD vs. non-SCD as the ECG-AI index increased as follows: low risk, 16.7%; intermediate risk, 18.5%; high risk, 28.7% (P for trend = 0.023). A similar but sharper separation was seen in the ECG-AI combined model: low risk, 11.8%; intermediate risk, 15.7%; high risk, 36.1% (P for trend

< 0.001). Notably, the ECG-AI index found not to be associated with the risk of all-cause mortality over 3 years [odds ratio, 1.02 (95% CI, 0.90–1.16); P = 0.720].

Discussion

In the present study, the association between the ECG-AI index and SCD was evaluated. Overall, the ECG-AI index, when added to the conventional guideline-directed indication for ICD based on LVEF and NYHA functional class, significantly improved the indices of discrimination and reclassification of SCD. We also observed an increased proportional risk estimation of SCD vs. non-SCD. Importantly, similar findings were observed in subsets of patients with HF with a non-ischaemic aetiology and those with an LVEF of >35%.

The early AI model applied to the ECG data used neural network structures (i.e. multilayer perceptron) other than CNN or RNN. Improvements in computing and neural network technology have allowed the development of a deeper network pattern and, as a result, have enabled the handling of more complex data. For example, ECG-AI models using two-dimensional CNN reportedly predict age and sex and detect LV function and further latent atrial fibrillation from normal sinus rhythm ECG.[12] We reported that an ECG-based AI model combining a one-dimensional CNN with RNN (i.e. LSTM) successfully identified patients with chest pain requiring urgent revascularization in an emergency setting.[14] Recurrent neural network can theoretically learn the time-series voltage data more precisely than CNN,
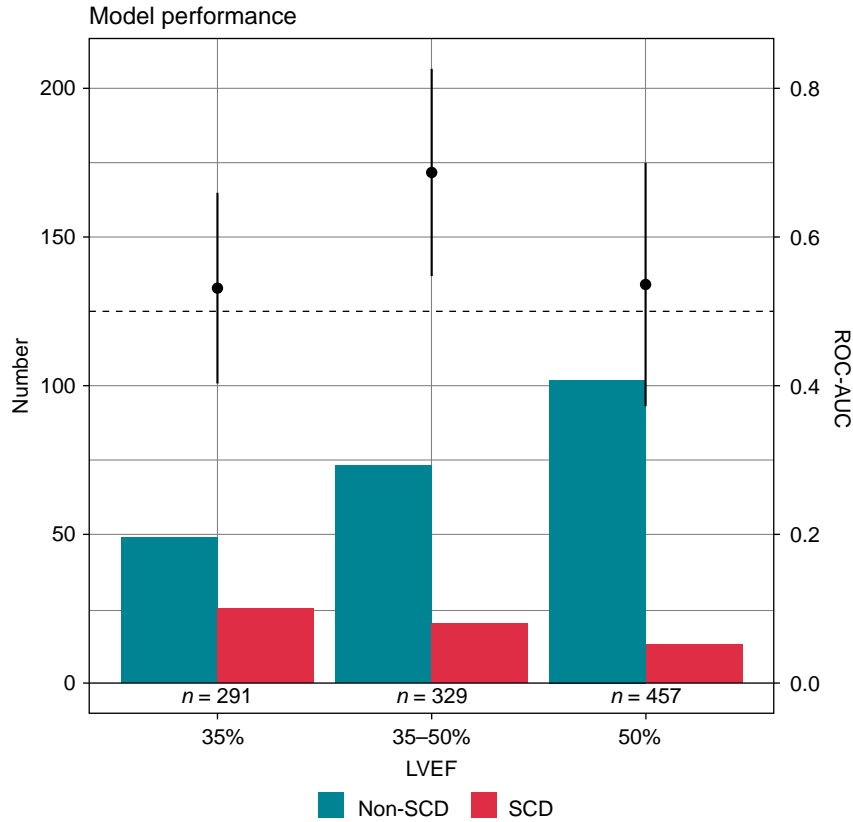
Figure 3 Discriminative ability of the ECG-AI index and frequency of SCD vs. non-SCD events by the LVEF category. The above forest plots show ROC-AUC with 95% confidence intervals of the ECG-AI index for predicting SCD events by LVEF. In patients with an LVEF of 35–50%, the ECG-AI index showed the best discriminative ability. The bar graphs represent the number of SCD and non-SCD patients according to LVEF. The proportion of patients with SCD vs. non-SCD steadily decreased as LVEF increased. ECG-AI, electrocardiography-based artificial intelligence; LVEF, left ventricular ejection fraction; ROC-AUC, receiver operating characteristic area under the curve; SCD, sudden cardiac death.
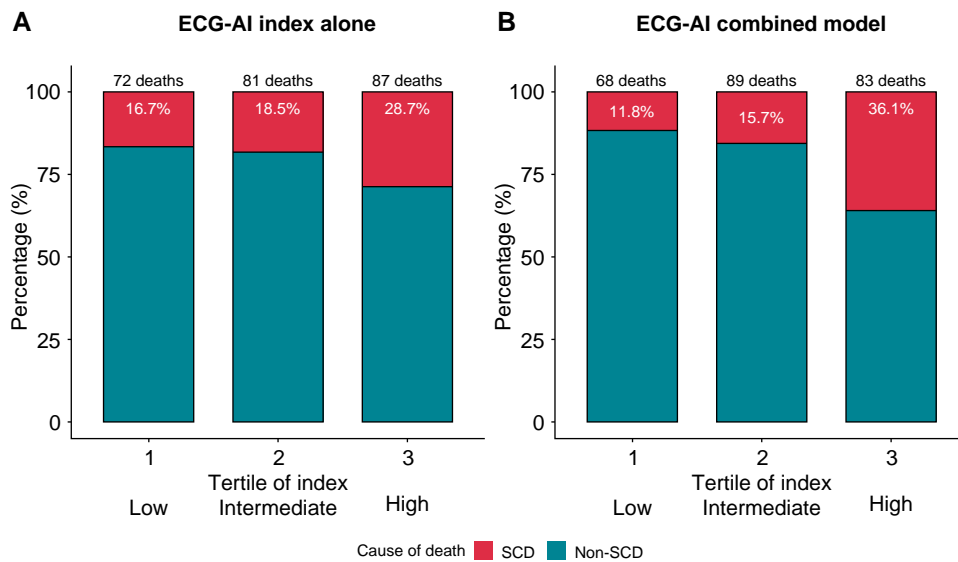
Figure 4 Proportional risk of SCD vs. non-SCD events based on the model-based risk. ECG-AI index (*A*) and ECG-AI combined model (with LVEF and NYHA class, *B*). The patients were divided into three groups by tertile of risk score (each *P* for trend < 0.05). ECG-AI, electrocardiography-based artificial intelligence; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association; SCD, sudden cardiac death.

as it explicitly deals with data ordering. Although some complex tasks may still require an RNN, the superiority of model performance using RNN over CNN is unclear; thus, consensus is lacking about the tasks that are suitable for RNN or CNN. In the present study, we attempted to establish an ECG-AI model using RNN combined with CNN to predict the incidence of SCD in patients with HF, and our results showed its good performance beyond the conventional indication for ICD. As accurate risk prediction of SCD is essential in clinical practice, and these new approaches using AI algorithms may help clinicians provide a basis for decision-making to ensure the appropriate application of ICDs.

Previous studies reported that several ECG features (e.g. heart rate variability, T-wave alternance, early repolarization, late potential, and atrial fibrillation) are associated with SCD and can provide independent values beyond LV dysfunction.[16] In fact, the Grad-CAM showed that our neural network model looked at the QRS wave and T-wave, which seemed to focus on early repolarization, late potential, and T-wave alternance. Our ECG-based AI algorithm likely integrates specific features associated with SCD and can be accurately applied to a broader spectrum of patients with HF than classical predictive models. This is important as SCD represents a substantial burden for patients with HF and an LVEF of >35% evidenced by the fact that ~50% of SCD cases occur in the absence of severe LV dysfunction following myocardial infarction.[17] Furthermore, a retrospective analysis of 714 patients with SCD found that only one-third exhibited sufficient LV dysfunction to meet the ICD criteria.[18] Our ECG-based AI algorithm demonstrated a more refined risk stratification, in particular, an enrichment in the proportional risk estimation of SCD vs. non-SCD regardless of LVEF. As the number of patients with HF and an LVEF of >35% is increasing worldwide and the prediction of SCD in these patient populations is considered highly difficult, we believe that our findings are highly encouraging.

A known limitation of the conventional ICD indication also exists in non-ischaemic patients. The DANISH trial (The Danish Study to Assess the Efficacy of ICDs in Patients with Non-ischaemic Systolic Heart Failure on Mortality) reported that patients with HF of a non-ischaemic aetiology could not benefit from ICD as a primary prevention for SCD.[19] However, a subgroup analysis of the DANISH trial indicated a mortality benefit of ICD implantation in younger patients corresponding to a lower proportion of SCD relative to non-SCD with increasing age.[20] In the present study, the ECG-AI index showed good discriminative ability, especially among younger patients with non-ischaemic aetiology. Another factor may be that women are more likely to have a non-ischaemic aetiology (see Supplementary material online, Table S3). In addition, the ECG-AI combined model showed significantly better discrimination of the incidence of SCD events among patients without AF than among those with AF. The reason for the better predictive values of ECG-AI in these patients is uncertain, as there were no compelling differences in patient characteristics between these subgroups. Therefore, the full interpretability of ECG-AIs will be the focus of future research.

We also observed that the ECG-AI index alone and the ECG-AI combined model could discriminate between SCD and non-SCD across patients with different risk scores. As patients at increased risk of SCD are also likely to have a significantly higher risk of non-SCD mortality, those with a higher absolute risk of SCD are not always at a higher proportional risk. Importantly, previous reports demonstrated that the projected ICD benefit is relatively insensitive to absolute SCD risk but is highly sensitive to proportional risk.[7,21] Furthermore, a recent analysis identified seven novel indicators associated with SCD; however, all were associated with non-SCD to at least the same extent and, hence, do not specifically predict SCD.[22] These observations further highlight the potential impact of the ECG-AI index applied in our study. Overall, the ECD-AI index appears well poised to meet the requirements of SCD predictions, which are distinctly lacking from the conventional standards.

Strengths and limitations

This study has some strengths including the standardized assessment and adjudication of SCD by the central study committee of the WET-HF registry and the derivation and validation of the ECG-AI index using different hospitals with different patient backgrounds. This study has also several inherent limitations. First, the limited number of patients and clinical events resulted in a relatively low power to detect the incidence of fatal arrhythmic events, although our result is comparable to previously reported values, demonstrating that ~20% of patients with HF with a preserved LVEF succumbed to SCD over a 3-year follow-up period.[18] Clearly, further studies are needed to validate our ECG-AI in a larger-scale database with a longer follow-up. In addition, we did not perform substantial statistical adjustments in the multivariable models, in which only two parameters, LVEF and NYHA class, were covariables. It was not possible to adjust the model with respect to several aetiologies, such as dilated cardiomyopathy, hypertrophic cardiomyopathy, or cardiac sarcoidosis, due to limited power. Other indicators, such as sex and body mass index, are reportedly useful for discriminating SCD from non-SCD,[7] but they have not been universally confirmed as relevant tools for predicting SCD. Furthermore, it limited our ability to train on multiple institutions and validate/test on an external one, because we only had access to data from the four institutions. Third, most patients in this cohort were of East Asian ancestry; thus, the results for other regions and races need to be confirmed. Fourth, angiotensin receptor-neprilysin inhibitors and sodium glucose cotransporter 2 inhibitors, which reduce the risk of SCD and are the standard of care based on the current clinical practice guidelines, were not available in Japan for patients with HF during the study period. Fifth, the lack of information regarding changes in medical treatment during follow-up may have affected the results of this study. However, from the CHAMP-HF (Change the Management of Patients with Heart Failure) registry including outpatients in the USA, it is also true that there has been no significant improvement in the guideline-directed medication use over 12 months, despite the publication of the new HF practice guidelines and increased recognition of the validity of HF performance measures related to guideline-directed medications.[23] Finally, a pitfall of these AI models is that unidentified biases or flaws can exist in the dataset, which can lead to misclassification.

Conclusions

In this study, the AI-based assessment of ECG was tested as a new model for risk stratification of SCD in patients with HF and found to be more discriminatory than conventional standards. Specifically, we observed improved prediction of SCD in patients with an LVEF of 35–50% and a non-ischaemic aetiology as well as discrimination between SCD and non-SCD. The multifactorial nature of the ECG-AI index has allowed the creation of a more sensitive predictive model that may address the current shortcomings of capturing dynamic and proportional SCD risk in patients with HF. Further investigations are needed to validate our results in external cohorts with high-quality data inputs and ultimately compare AI-guided treatment with the standard treatment in a randomized controlled trial.

Supplementary material

Supplementary material is available at *Europace* online.

Acknowledgements

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

1. 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.
2. Sato N. Epidemiology of heart failure in Asia. Heart Fail Clin 2015;**11**:573–9.
3. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M et al. 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021;**42**:3599–726.
4. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM et al. 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 2022;**145**:e895–e1032.
5. Solomon SD, Wang D, Finn P, Skali H, Zornoff L, McMurray JJV et al. Effect of candesartan on cause-specific mortality in heart failure patients: the Candesartan in Heart Failure Assessment of Reduction in Mortality and Morbidity (CHARM) program. Circulation 2004;**110**:2180–3.
6. Poole JE, Johnson GW, Hellkamp AS, Anderson J, Callans DJ, Raitt MH et al. Prognostic importance of defibrillator shocks in patients with heart failure. N Engl J Med 2008;**359**:1009–1017.
7. Shadman R, Poole JE, Dardas TF, Mozaffarian D, Cleland JGF, Swedberg K et al. A novel method to predict the proportional risk of sudden cardiac death in heart failure: derivation of the Seattle Proportional Risk Model. Heart Rhythm 2015;**12**:2069–77.
8. Fukuoka R, Kohno T, Kohsaka S, Shiraishi Y, Sawano M, Abe T et al. Prediction of sudden cardiac death in Japanese heart failure patients: international validation of the Seattle Proportional Risk Model. Europace 2020;**22**:588–97.
9. Rohde LE, Chatterjee NA, Vaduganathan M, Claggett B, Packer M, Desai AS et al. Sacubitril/valsartan and sudden cardiac death according to implantable cardioverter-defibrillator use and heart failure cause: a PARADIGM-HF analysis. JACC Heart Fail 2020;**8**:844–55.
10. Buxton AE, Waks JW, Shen C, Chen PS. Risk stratification for sudden cardiac death in North America–current perspectives. J Electrocardiol 2016;**49**:817–23.
11. Goto S, Mahara K, Nelson LB, Ikura H, Katsumata Y, Endo J et al. Artificial intelligence-enabled, fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. Nat Commun 2021;**12**:2726.
12. Attia ZI, Kapa S, Jimenez FL, McKie PM, Ladewig DJ, Satam G et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat Med 2019;**25**:70–4.
13. Raghunath S, Cerna AEU, Jing L, vanMaanen DP, Stough J, Hartzel DN et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. Nat Med 2020;**26**:886–91.
14. Goto S, Kimura M, Katsumata Y, Goto S, kamatani T, Ichihara G et al. Artificial intelligence to predict needs for urgent revascularization from 12-leads electrocardiography in emergency patients. PLoS One 2019;**14**:e0210103.
15. Hicks KA, Tcheng JE, Bozkurt B, Chaitman BR, Cutlip DE, Farb A et al. 2014 ACC/AHA key data elements and definitions for cardiovascular endpoint events in clinical trials: a report of the American College of Cardiology/American Heart Association Task Force on clinical data standards (writing committee to develop cardiovascular endpoints data standards). Circulation 2015;**132**:302–61.
16. Liew R. Prediction of sudden arrhythmic death following acute myocardial infarction. Heart 2010;**96**:1086–94.
17. Faxén J, Jernberg T, Hollenberg J, Gadler F, Herlitz J, Szummer K. Incidence and predictors of out-of-hospital cardiac arrest within 90 days after myocardial infarction. J Am Coll Cardiol 2020;**76**:2926–36.
18. Stecker EC, Vickers C, Waltz J, Socoteanu C, John BT, Mariani R et al. Population-based analysis of sudden cardiac death with and without left ventricular systolic dysfunction: two-year findings from the Oregon Sudden Unexpected Death Study. J Am Coll Cardiol 2006;**47**:1161–1166.
19. Køber L, Thune JJ, Nielsen JC, Haarbo J, Videbæk L, Korup E et al. Defibrillator implantation in patients with nonischemic systolic heart failure. N Engl J Med 2016;**375**:1221–1230.
20. Elming MB, Nielsen JC, Haarbo J, Videbæk L, Korup E, Signorovitch J et al. Age and outcomes of primary prevention implantable cardioverter-defibrillators in patients with nonischemic systolic heart failure. Circulation 2017;**136**:1772–80.
21. Bilchick KC, Wang Y, Cheng A, Curtis JP, Dharmarajan K, Stukenborg GJ et al. Seattle Heart Failure and proportional risk models predict benefit from implantable cardioverter-defibrillators. J Am Coll Cardiol 2017;**69**:2606–18.
22. Rohde LE, Vaduganathan M, Claggett BL, Planczyk CA, Dorbala P, Packer M et al. Dynamic changes in cardiovascular and systemic parameters prior to sudden cardiac death in heart failure with reduced ejection fraction: a PARADIGM-HF analysis. Eur J Heart Fail 2021;**23**:1346–1356.
23. Bozkurt B. Reasons for lack of improvement in treatment with evidence-based therapies in heart failure. J Am Coll Cardiol 2019;**73**:2384–7.