



# Non-invasive hemoglobin estimation with outcome prediction via deep learning analysis of ECG-derived cardiac micro-dynamics

Chung-Chi Yang<sup>1,2,3,4</sup> · Chin-Sheng Lin<sup>2</sup> · Chin Lin<sup>5,6,7</sup> · Wei-Ting Liu<sup>2</sup> · Zih-Yin Lai<sup>3,4</sup> · Yung-Jen Chuang<sup>3,4</sup> 

Received: 13 August 2025 / Accepted: 17 November 2025  
© The Author(s) 2025

## Abstract

**Background** Anemia is a global health issue, especially in resource-limited areas, where traditional hemoglobin (Hb) testing is invasive and costly. This study aimed to develop an electrocardiogram–hemoglobin (ECG-Hb) deep learning model (DLM) for detecting anemia and assess its impact on all-cause mortality and new-onset heart failure.

**Methods** This retrospective study analyzed ECGs and corresponding Hb levels from two hospitals. The DLM was trained on 388,166 ECGs from 187,202 patients and tested on 24,279 and 29,247 patients in internal and external sets, respectively. Anemia was defined as moderate ( $Hb \leq 10$  g/dL) or severe ( $Hb \leq 8$  g/dL). Diagnostic performance was evaluated using receiver operating characteristic (ROC) curve analysis, and an 8-year follow-up assessed mortality and heart failure risk with Cox regression.

**Results** The areas under the ROC curves (AUCs) for detecting moderate-to-severe anemia were 0.8545 (internal) and 0.8243 (external), with sensitivities of 65.9% and 71.0%, and specificities of 84.8% and 77.4%, respectively. ECG-Hb performed better in detecting severe anemia ( $AUC = 0.9038/0.8766$ ) than in mild anemia. Pearson correlations between ECG-Hb and Hb were 0.56 (internal) and 0.53 (external). Key ECG features, including heart rate variability, significantly influenced ECG-Hb. Patients with severely low ECG-Hb had higher risks of mortality (hazard ratio [HR]: 1.71, 95% confidence interval [CI]: 1.42–2.06) and heart failure (HR: 2.47, 95% CI: 2.07–2.94) compared to those with standard ECG-Hb levels.

**Conclusion** The ECG-Hb DLM offers strong diagnostic and prognostic potential for anemia and cardiovascular risks, making it a valuable, non-invasive screening tool in low-resource settings.

**Keywords** Anemia detection · Deep learning model · Electrocardiogram · Heart rate variability · Prognostic value

✉ Yung-Jen Chuang  
yjchuang@life.nthu.edu.tw

<sup>1</sup> Division of Cardiovascular Medicine, Taoyuan Armed Forces General Hospital, Taoyuan 32551, Taiwan, ROC

<sup>2</sup> Division of Cardiology, Department of Internal Medicine, Tri-Service General Hospital, National Defense Medical University, Taipei, Taiwan, ROC

<sup>3</sup> School of Medicine, National Tsing Hua University, Hsinchu 300044, Taiwan, ROC

<sup>4</sup> Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu 300044, Taiwan, ROC

<sup>5</sup> Medical Technology Education Center, College of Medicine, National Defense Medical University, Taipei, Taiwan, ROC

<sup>6</sup> Departments of Artificial Intelligence of Things, Tri-Service General Hospital, National Defense Medical University, Taipei, Taiwan, ROC

<sup>7</sup> College of Public Health, National Defense Medical University, Taipei, Taiwan, ROC

## Introduction

Anemia stands as a pressing global health concern, affecting over 800 million individuals, predominantly in resource-limited regions [1]. The World Health Organization (WHO) has proactively outlined a comprehensive strategy to expedite the mitigation of anemia [2]. Anemia can be triggered by a myriad of acute and chronic conditions, encompassing nutritional deficiencies due to subpar diets or impaired nutrient absorption, infections, gastrointestinal bleeding, malignancies, and chronic renal diseases, among others [3]. Patients may not exhibit symptoms in instances of moderate or insidious anemia. Consequently, diagnosis with only patient histories and examinations becomes easier once the condition is uncompensated and complications occur [4]. Anemia can escalate the risk of cardiovascular diseases and mortality [5, 6], necessitating prompt intervention to enhance patient prognosis. While anemia can be diagnosed

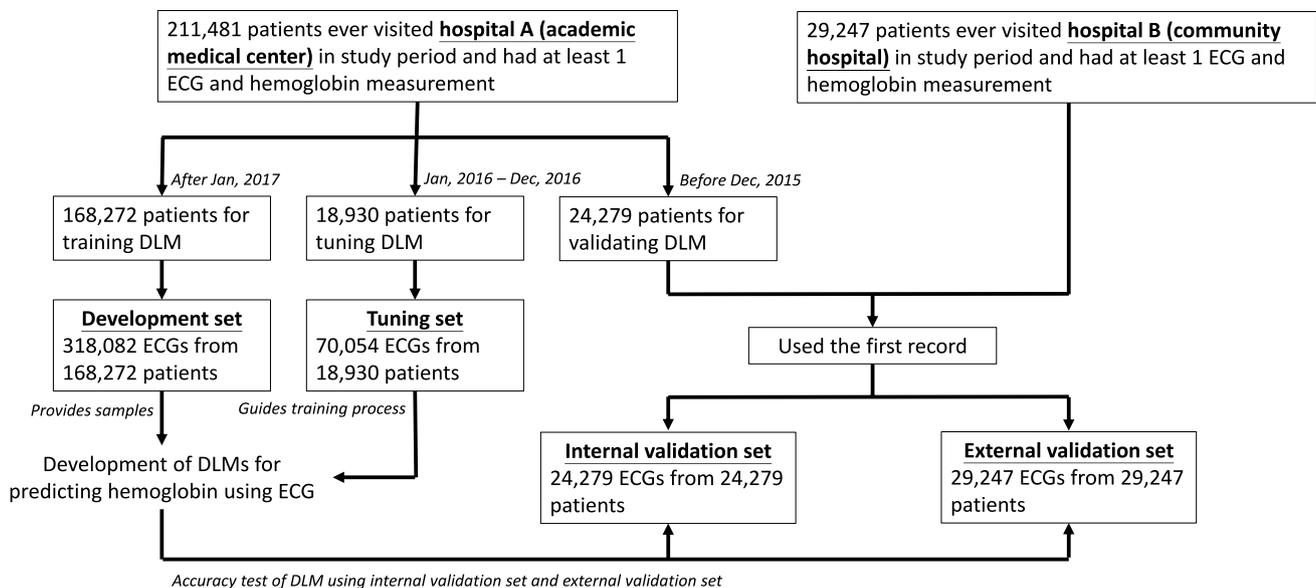
through laboratory blood tests [7], the invasive nature of the procedure and the substantial costs associated with blood testing equipment signifies that many patients with asymptomatic anemia in various regions remain undiagnosed and untreated [8].

The association between heart rate variability (HRV) and anemia has been extensively reported [9, 10]. Furthermore, numerous studies have indicated that anemia can alter electrocardiogram (ECG) morphology, suggesting that a mismatch between oxygen demand and supply in the myocardium influences ECG patterns [11, 12]. Given the cost-effectiveness and non-invasive nature of electrocardiography, it is particularly suitable for anemia screening in resource-limited regions. Moreover, the evolution of wearable ECG devices has broadened their accessibility and usage in cloud computing. Despite the potential association between HRV, ECG alterations, and anemia, diagnosing anemia based solely on ECG remains challenging.

With the advent of deep learning models (DLMs), these advanced algorithms have been validated to detect conditions, such as acute myocardial infarction, [13, 14] pericarditis, [15] digoxin toxicity, [16] pneumothorax, [17] aortic dissection, [18] thyrotoxic periodic paralysis, [19] and dyskalemias [20], using ECGs, often surpassing the diagnostic capabilities of cardiologists. Furthermore, prior studies have illustrated that artificial intelligence (AI)-enhanced ECGs detect previous history of left ventricular dysfunction even in patients with a normal ejection fraction [21–24]. This capability is not limited to identifying such precursors;

AI-empowered ECGs have also pinpointed patient groups at elevated mortality risk, even when their serum potassium concentrations are within the normal range [25]. In addition, when combined with AI, ECGs can estimate "heart age" as an indicator of cardiovascular health [26] and accurately predict short-term mortality [27, 28]. Leveraging AI, ECGs hold the potential to forecast future disease risks in ostensibly healthy individuals and facilitate early preventive measures, thereby mitigating disease risks. Such advancements underscore the significance of primary prevention, health promotion, and targeted protective measures.

A recent investigation pioneered the development of a DLM capable of detecting anemia, achieving an area under the receiver operating characteristic (ROC) curve (AUCs) exceeding 0.85, relying solely on 12-lead ECGs [29]. However, the model's relatively low positive predictive value (PPV) of <20% and its lack of clarity regarding false-positive findings to precursors of cardiovascular diseases has curtailed its practical applicability. Moreover, no comprehensive analysis has elucidated the underlying mechanisms by which this "black-box" AI model accurately identifies anemia. This study aimed to develop a DLM that employs ECG as a novel biomarker, termed electrocardiogram–hemoglobin (ECG-Hb). Beyond conducting analyses of HRV and ECG patterns to enhance the transparency of the inner workings of AI, we also aimed to explore its implications for all-cause mortality and the onset of heart failure. This study sought to establish guidelines for prospective, large-scale



**Fig. 1** Development, tuning, internal validation, and external validation sets generation and ECG labeling of hemoglobin. Schematic of the data set creation and analysis strategy, which was devised to assure a robust and reliable data set for training, validating, and testing of the network. Once a patient's data were placed in one of the

data sets, the data is used only in that set, avoiding 'cross-contamination' among the training, validation, and test data sets. The details of the flow chart and the usage of each data set are described in the Materials and Methods

**Table 1** Baseline characteristics

	Development set	Tuning set	Internal validation set	External validation set
<b>Hemoglobin profile</b>				
Hb (g/dL)	12.9 ± 2.3	12.5 ± 2.4	13.6 ± 2.0	13.3 ± 2.1
Hb ≤ 8	39,568(12.4%)	11,879(17.0%)	1303(5.4%)	2120(7.2%)
8 < Hb ≤ 10	56,488(17.8%)	15,477(22.1%)	3066(12.6%)	4498(15.4%)
10 < Hb	222,026(69.8%)	42,698(61.0%)	19,910(82.0%)	22,629(77.4%)
<b>Demography</b>				
Sex (male)	169,991(53.4%)	35,506(50.7%)	12,974(53.4%)	14,702(50.3%)
Age (years)	58.4 ± 18.5	64.0 ± 17.7	52.6 ± 17.8	57.6 ± 20.6
BMI (kg/m <sup>2</sup> )	24.3 ± 4.2	24.2 ± 4.3	24.3 ± 4.1	24.1 ± 4.2
SBP (mmHg)	133.7 ± 26.9	138.6 ± 28.4	133.9 ± 26.5	137.8 ± 27.5
DBP (mmHg)	78.7 ± 16.8	78.0 ± 17.7	79.1 ± 16.3	77.5 ± 17.1
<b>Disease history</b>				
DM	57,918(18.2%)	21,563(30.8%)	3313(13.6%)	5726(19.6%)
HTN	88,605(27.9%)	34,048(48.6%)	6122(25.2%)	10,191(34.8%)
HLP	75,594(23.8%)	27,822(39.7%)	5015(20.7%)	8376(28.6%)
CKD	59,654(18.8%)	25,383(36.2%)	2014(8.3%)	3587(12.3%)
AMI	9825(3.1%)	4235(6.0%)	258(1.1%)	289(1.0%)
STK	33,534(10.5%)	13,244(18.9%)	1733(7.1%)	3209(11.0%)
CAD	53,596(16.8%)	22,089(31.5%)	3154(13.0%)	4855(16.6%)
HF	22,158(7.0%)	10,475(15.0%)	1038(4.3%)	1661(5.7%)
AF	12,175(3.8%)	6424(9.2%)	521(2.1%)	821(2.8%)
COPD	33,223(10.4%)	13,764(19.6%)	2239(9.2%)	4362(14.9%)

Hb, hemoglobin; BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; DM, diabetes mellitus; HTN, hypertension; HLP, hyperlipidemia; CKD, chronic kidney disease; AMI, acute myocardial infarction; STK, stroke; CAD, coronary artery disease; HF, heart failure; AF, atrial fibrillation; COPD, chronic obstructive pulmonary disease

implementation in resource-limited regions for screening purposes.

## Materials and methods

### Data source

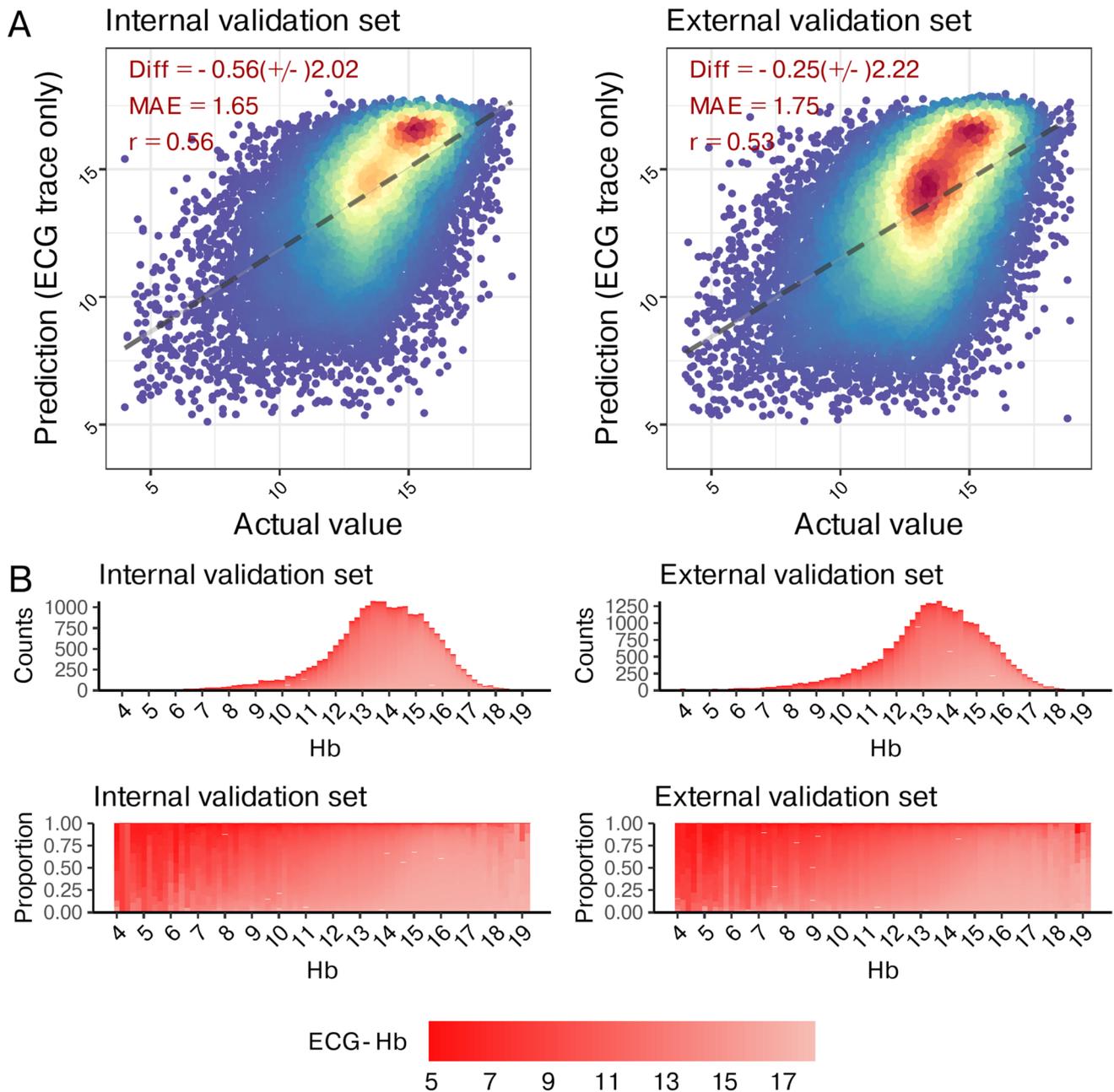
This retrospective study was approved by the Ethics Committee of the Tri-Service General Hospital (C202105049) to collect medical data. The study aimed to formulate a DLM and subsequently validate its internal and external efficacies. ECGs were obtained from two hospitals within the Tri-Service General Hospital system: an academic medical institution in Neihu District (Hospital A) and a community-based hospital in Zhongzheng District (Hospital B). The collection period spanned from January 1, 2010 to September 30, 2021. Each ECG was annotated based on the nearest Hb, which varied between 5.0 and 18.0 g/dL. The nearest numerical boundary replaced values outside this range. ECGs that lacked an Hb test within a 12-h window were omitted from the study. No other exclusion criteria were applied.

As shown in Fig. 1, a systematic approach was employed to develop and validate the DLM. Within Hospital A,

211,481 patients presented with at least one paired ECG and Hb for this investigation. Subsequently, 168,272 patients who sought medical care at Hospital A after January 1, 2017 were allocated to the development cohort, contributing 318,082 ECG records for DLM training (mean ECGs per patient: 1.89). The cohort comprising 18,930 patients who visited between January 1, 2016 and December 31, 2016 was designated as the tuning set. Furnishing with 70,054 ECGs was done to facilitate the training process and establish the pivotal diagnostic threshold (mean ECGs per patient: 3.70). Before December 31, 2015, 24,279 patients seeking care were categorized into the internal validation cohort, and their initial ECGs were used for accuracy assessment and subsequent analyses. An external validation set was constructed using data from Hospital B to ascertain the generalizability of the DLM. The validation set encompassed 29,247 patients and adhered to the inclusion criteria identical to those of Hospital A.

### Data collection

Within the ECG dataset, patients were selected based on the presence of at least one standard digital ECG, characterized by a 500 Hz frequency, a duration of 10 s, and a

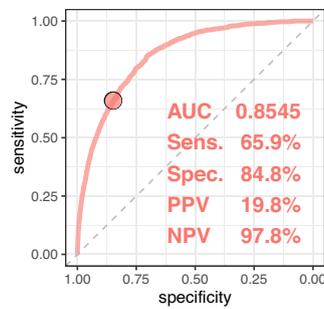
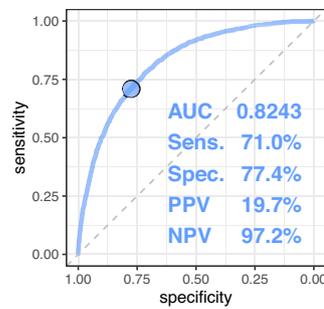
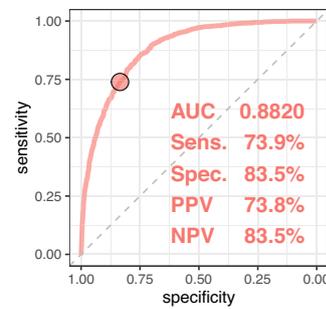
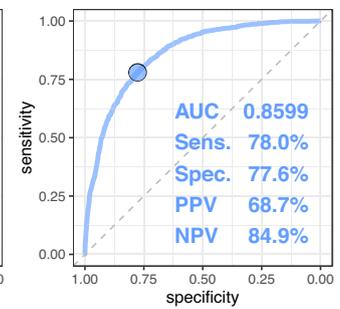
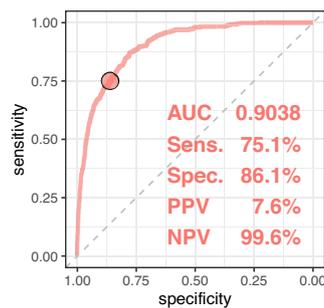
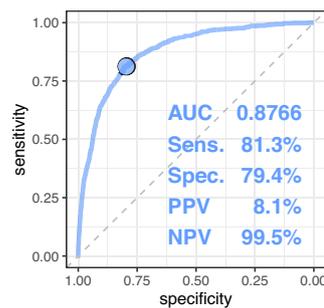
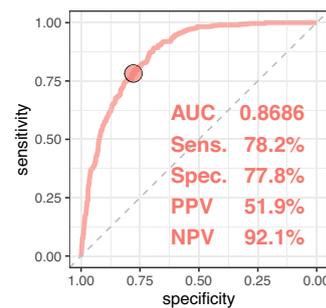
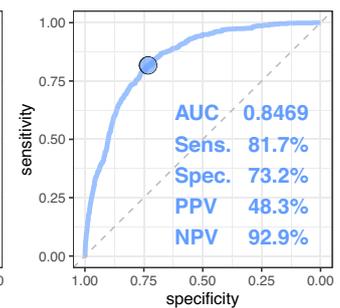


**Fig. 2** Predicted hemoglobin (ECG-Hb) and actual Hb. A) Scatter plots of ECG-Hb via ECG only compared to the actual Hb. The x-axis indicates the actual Hb, and the y-axis presents the ECG-Hb. Red points represent the highest density, followed by yellow, green, light blue, and dark blue. We presented the mean difference (Diff), Pearson correlation coefficients ( $r$ ), and mean absolute errors (MAE) to demonstrate the accuracy of DLM. The black lines with 95% con-

ference intervals are fitted via simple linear regression. B) The distributions of Hb in the internal and external validation sets. The color gradient from white to red demonstrated the ECG-Hb levels from normal to low. The panel above shows the original distribution of each dataset, and the panel below shows the distribution of ECG-Hb in each actual Hb value

12-lead configuration obtained in the supine position during the designated study timeframe. Hb levels were determined using cyanide-free sodium lauryl sulfate (SLS) methodology in the central laboratory. The DLM was trained using raw ECG traces. Both quantitative metrics and anomalous findings were derived from the ECG, resulting in 31 diagnostic

pattern classifications and 8 continuous ECG measurements. Any missing data within the ECG measurements were addressed by applying multiple imputations [30]. The proportions of missing values were 7.8% for the PR interval, 8.1% for the P-wave axis, and 1.4% for the T-wave axis. Other ECG measurements had no missing values. The 31

**Original distribution****Internal validation set****Hb  $\leq 10$  vs Hb  $> 10$** **External validation set****Hb  $\leq 10$  vs Hb  $> 10$** **Balance distribution****Internal validation set****Hb  $\leq 10$  vs Hb  $> 10$** **External validation set****Hb  $\leq 10$  vs Hb  $> 10$** **Hb  $\leq 8$  vs Hb  $> 8$** **Hb  $\leq 8$  vs Hb  $> 8$** **Hb  $\leq 8$  vs Hb  $> 8$** **Hb  $\leq 8$  vs Hb  $> 8$** 

**Fig. 3** The ROC curve of DLM predictions based on ECG to detect moderate to severe anemia. Moderate and severe anemia were defined as actual hemoglobin (Hb) of  $\leq 10$  and  $\leq 8$ , respectively. The operating point was selected based on the maximum of Youden's index in the tuning set and presented using a circle mark, and the area under ROC curve (AUC), sensitivity (Sens.), specificity (Spec.), positive

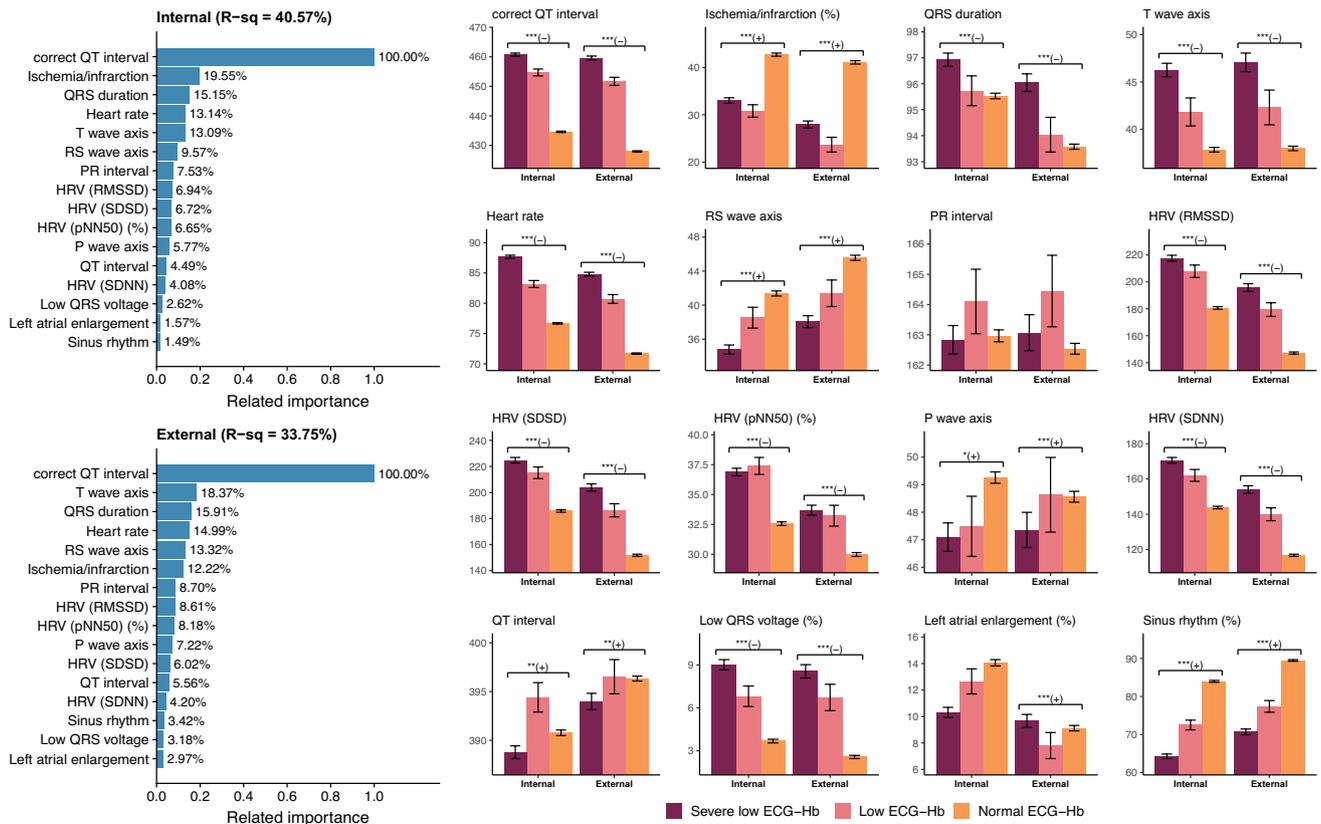
predictive value (PPV), and negative predictive value (NPV) were calculated based on it (left panel). Due to the different distributions of Hb in the internal and external validation sets, we generated the balance dataset for each set to ensure the same number of cases in different values of Hb (right panel)

clinical diagnostic patterns were discerned from the structured findings and predicted by key terminologies consistent with the Philips system. Based on a previous study, the HRV parameters were calculated based on 10-s raw ECG traces, including standard deviation of normal-to-normal intervals (SDNN), root mean square of successive differences (RMSSD), standard deviation of successive differences (SDSD), and percentage of successive NN intervals that differ by more than 50 ms (pNN50) defined based on a previous study [31]. Anemia classifications were established, with moderate anemia defined by an Hb level of  $\leq 10$  g/dL and severe anemia characterized by an Hb level of  $\leq 8$  g/dL.

Disease histories were delineated according to the International Classification of Diseases (ICD), Ninth and Tenth Revision (ICD-9 and ICD-10, respectively). The classifications are as follows: diabetes mellitus (DM, ICD-9 codes 250.x and ICD-10 codes E08.x to E13.x), hypertension (HTN, ICD-9 codes 401.x to 404.x and ICD-10 codes I10.x to I16.x), hyperlipidemia (HLP, ICD-9 codes 272.x and ICD-10 codes E78.x), chronic kidney disease (CKD, ICD-9 codes 585.x and ICD-10 codes N18.x), acute myocardial infarction (AMI, ICD-9 codes 410.x and ICD-10 codes I21.x), stroke (STK, ICD-9 codes 430.x to 438.x

and ICD-10 codes I60.x to I63.x), coronary artery disease (CAD, ICD-9 codes 410.x to 414.x, and 429.2, and ICD-10 codes I20.x to I25.x), heart failure (HF, ICD-9 codes 428.x and ICD-10 codes I50.x), atrial fibrillation (AF, ICD-9 codes 427.31 and ICD-10 codes I48.x), and chronic obstructive pulmonary disease (COPD, ICD-9 codes 490.x to 496.x and ICD-10 codes J44.9).

This study focused on complications in all-cause mortality and new-onset heart failure. Survival duration was ascertained using ECG data as a reference point for all-cause mortality. Patient status (deceased or alive) was determined via electronic medical records, which were updated with each hospital interaction. Furthermore, to mitigate potential bias arising from incomplete documentation, data pertaining to live visits were censored during the most recent documented hospital encounter. Relevant ICD codes were used to characterize the incidence of new-onset heart failure. Notably, patients diagnosed with heart failure prior to the ECG date were excluded from the new-onset heart failure complication analysis.



**Fig. 4** Relationship between selected ECG features and predicted hemoglobin (ECG-Hb). The related importance is based on the information gained by the XGB model, and the R-square (R-sq) is the coefficient of determination used in selected ECG features for predicting ECG-Hb. AI-ECG prediction was classified as normal ECG-

Hb, low ECG-Hb, and severe low ECG-Hb based on the operating points, the same as the previous ROC curve analysis. The analyses were conducted both in the internal and external validation sets. (\*: p for trend < 0.05; \*\*: p for trend < 0.01; \*\*\*: p for trend < 0.001)

## Deep learning model training

In this study, we employed the previously proposed ECG-12Net, an 82-layer convolutional neural network, as referenced in previous studies [32, 33]. For this investigation, we adopted the identical architectural framework to train a DLM to estimate ECG-Hb. The training specifics were consistent with methodologies delineated in prior studies [34–37]. Furthermore, an oversampling procedure was implemented, predicated on the inverse prevalence of each Hb interval, incrementing by 0.5 g/dL, ranging from 5.0 to 18.0 g/dL within the development set.

## Statistical analysis and model performance assessment

Patient demographics were delineated as means with standard deviations, patient counts, or percentages, where appropriate. All statistical evaluations were performed using R software (version 3.4.4) with a predetermined significance threshold of  $p < 0.05$ . To juxtapose the actual Hb with the

ECG-Hb, scatter plots showing mean differences and standard deviations (Diff), Pearson correlation coefficient ( $r$ ), and mean absolute error (MAE) were employed. Additionally, the diagnostic efficacy for varying degrees of anemia (moderate and severe) was assessed in both the internal and external validation sets. Metrics, such as area under the receiver operating characteristic curve (AUC), sensitivity (Sens.), specificity (Spec.), positive predictive value (PPV), and negative predictive value (NPV), were elucidated. The optimal operating point was determined based on the maximum Youden index value within the tuning set. Given the disparate Hb distributions in the validation sets, balanced datasets were curated for each set, ensuring uniform case numbers across Hb values.

Furthermore, the outcomes from the XGBoost (XGB) model were presented, offering appropriate variable rankings to elucidate the nexus between discernible features and ECG-Hb. Multivariate Cox proportional hazard models were used to investigate the association between ECG-Hb levels and outcomes. Comparative metrics include hazard ratios, each accompanied by a 95% confidence interval.

## Results

Table 1 shows the distribution of Hb levels, fundamental demographics, and historical disease data across the development, tuning, internal validation, and external validation sets. The average ages for these sets are  $58.4 \pm 18.5$ ,  $64.0 \pm 17.7$ ,  $52.6 \pm 17.8$ , and  $57.6 \pm 20.6$  years, respectively. Among these cohorts, 53.4%, 50.7%, 53.4%, and 50.3% were male. The proportion of individuals with severe anemia ( $Hb \leq 8$ ) in these sets is 12.4%, 17.0%, 5.4%, and 7.2%, respectively. Meanwhile, the prevalence of moderate anemia ( $8 < Hb \leq 10$ ) in each set is 17.8%, 22.1%, 12.6%, and 15.4%, respectively.

Figure 2a shows a scatter plot comparing laboratory hemoglobin (Lab-Hb) to ECG-Hb. In the internal validation set, the MAE between Lab-Hb and ECG-Hb was 1.65, exhibiting a correlation coefficient of 0.56 and a Diff of  $-0.56 \pm 2.02$ . Notably, the accuracy marginally diminished in the external validation set, with an MAE of 1.75, a correlation coefficient of 0.53, and a Diff of  $-0.25 \pm 2.22$ . The upper segment of Fig. 2b shows the actual Lab-Hb distribution in both internal and external sets. Most patients in the internal validation set had a lab-Hb range of 12–16 g/dL, and a similar distribution was observed in the external validation set. While there appears to be no overt distinction in the Lab-Hb distribution between the internal and external validation sets, it is imperative to note that the external validation set (22.6%) exhibits a higher proportion of patients with Lab-Hb of  $\leq 10$  compared to the internal validation set (18.0%). For a balanced representation, the distribution of ECG-Hb for each lab-Hb value was visualized, revealing comparable proportions in both validation sets, as shown in the lower segment of Fig. 2b. In conclusion, a discernible trend was observed: As lab-Hb decreased, a higher proportion exhibited a correspondingly lower ECG-Hb.

Figure 3 (left panel) shows the proficiency of the DLM in identifying moderate and severe anemia. For the detection of moderate-to-severe anemia, the AUC was 0.8545, accompanied by a Sens of 65.9%, Spec of 84.8%, PPV of 19.8%, and NPV of 97.8%. Notably, the AUC was augmented to detect severe cases, registering 0.9038, with a Sens of 75.1% and a Spec of 86.1%. The raw accuracies in the external validation set declined slightly compared to those in the internal validation set. Specifically, for moderate anemia, the metrics were AUC=0.8243, Sens=71.0%, Spec=77.4%, PPV=19.7%, and NPV=97.2%. The AUC for detecting severe cases (0.8766) was marginally reduced in the external set compared to that of the internal set. Considering the higher prevalence of moderate cases in the external validation set (22.6%) than in the internal validation set (18.0%), it may be more challenging for the DLM to differentiate severe cases

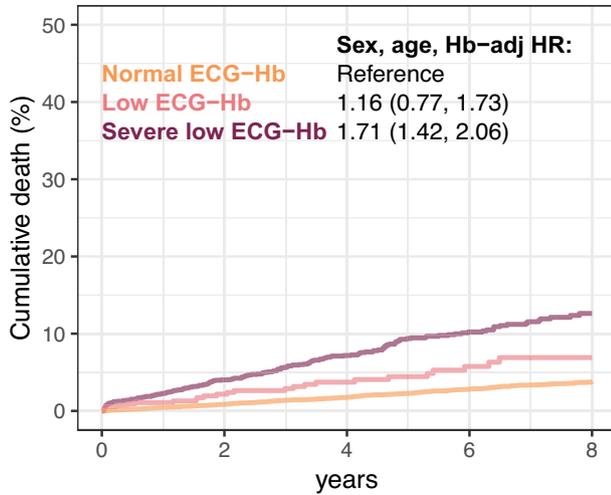
from moderate ones compared to normal cases. Therefore, a balanced analysis was performed for each set, considering the respective distributions depicted in Fig. 3 (right panel). The AUCs for the external validation set (0.8599 for moderate anemia and 0.8469 for severe anemia) remained slightly lower than those of the internal validation set (0.8820 for moderate anemia and 0.8686 for severe anemia). This discrepancy might stem from variations in patient characteristics between the external and internal validation sets; it was not caused by the different distributions of Lab-Hb.

Extended Fig. 1 delineates the significance of the relationship between all ECG features and ECG-Hb as determined by the information gained from the XGB model. The coefficients of determination (R-square) values based on all ECG features were 40.90% and 34.33% for the internal and external validation sets, respectively. Figure 4 shows the 12 most noticeable ECG features pertinent to ECG-Hb. The R-square values derived from these features were congruent with those obtained using all the ECG features, registering 40.57% for the internal validation set and 33.75% for the external validation set. Within the internal validation set, the ECG features of paramount importance to ECG-Hb included corrected QT interval, ischemia/infarction markers, QRS duration, T-wave axis, heart rate, and RS-wave axis in descending order. Notably, ECG measurements and HRV played key roles in establishing ECG-Hb levels, which had more significance than ECG morphology. The top six features in the external validation set mirrored those of the internal set, albeit with a slight reordering. The corrected QT interval may be the most important ECG feature for establishing ECG-Hb. Specifically, the corrected QT interval was the longest in the severely low group, followed by the low group and the normal group.

Figure 5 shows the disparity in new-onset cardiovascular events between false-positive and true-negative cases, underscoring the prognostic value of ECG-Hb levels. Within the internal validation set, the incidence of all-cause mortality in the severely low ECG-Hb group was 4.0% at 2 years and increased to 12.6% at 8 years, markedly elevated compared to the normal ECG-Hb group, which registered at 0.8% and 3.8%, respectively, yielding an adjusted HR of 1.71 (95% CI: 1.42–2.06). A conspicuous dose–response relationship was evident, transitioning from the HR of the low ECG-Hb group (1.16, 95% CI: 0.77–1.73) to that of the severe low ECG-Hb group. This correlation was corroborated using an external validation set. Further analysis on the onset of heart failure revealed HRs of 2.20 (95% CI: 1.61–3.01) and 2.47 (95% CI: 2.07–2.94) for the low ECG-Hb group and severe low ECG-Hb group, respectively, within the internal validation set. Notably, a pronounced dose–response relationship persisted. Parallel analyses conducted on the external

### Internal validation set

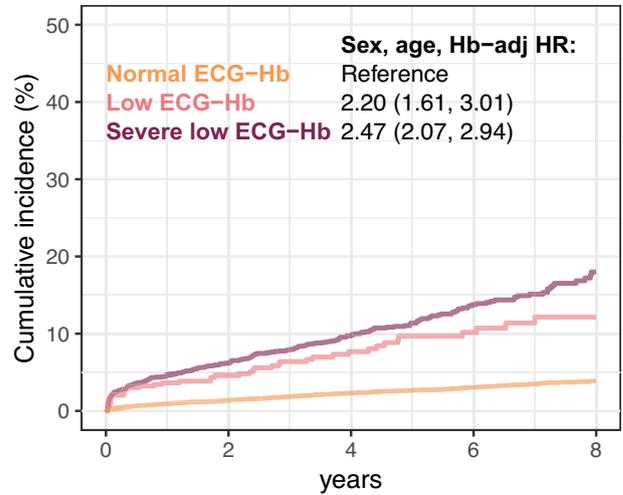
All-cause mortality



Number at risk/event rate (%)

19492	13057	9992	7237	2810
(0.0%)	(0.8%)	(1.8%)	(2.8%)	(3.8%)
670	427	294	192	68
(0.0%)	(2.2%)	(3.7%)	(5.8%)	(8.3%)
2814	1834	1255	745	277
(0.0%)	(4.0%)	(7.2%)	(10.2%)	(12.6%)

New-onset heart failure

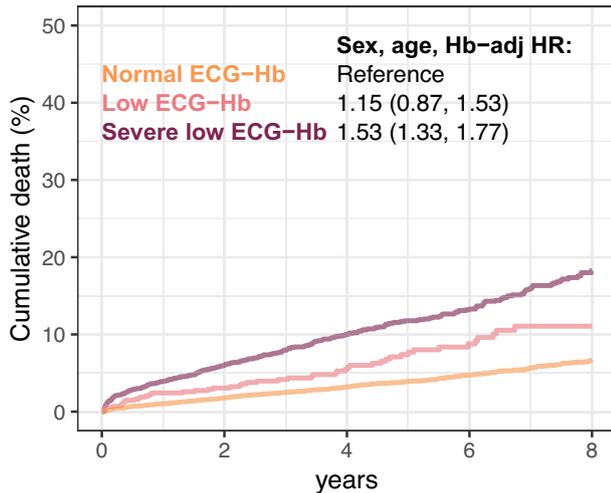


Number at risk/event rate (%)

15993	12466	9178	6605	2381
(0.0%)	(1.4%)	(2.3%)	(3.0%)	(3.8%)
558	389	255	164	64
(0.0%)	(4.6%)	(7.7%)	(10.2%)	(12.2%)
2244	1617	1015	607	202
(0.0%)	(6.2%)	(9.8%)	(13.8%)	(18.0%)

### External validation set

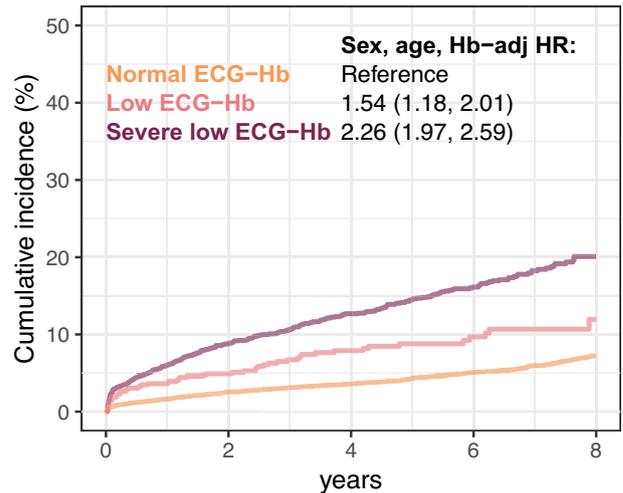
All-cause mortality



Number at risk/event rate (%)

21000	9912	6556	3880	1561
(0.0%)	(1.8%)	(3.2%)	(4.7%)	(6.6%)
1125	568	390	223	77
(0.0%)	(3.1%)	(5.8%)	(8.8%)	(11.1%)
5002	2406	1589	906	323
(0.0%)	(6.0%)	(10.0%)	(13.3%)	(18.2%)

New-onset heart failure



Number at risk/event rate (%)

16791	9775	6335	3607	1292
(0.0%)	(2.6%)	(3.6%)	(5.1%)	(7.2%)
884	522	348	194	63
(0.0%)	(4.9%)	(7.9%)	(9.7%)	(11.9%)
3866	2091	1347	758	249
(0.0%)	(8.8%)	(12.7%)	(16.1%)	(20.1%)

**Fig. 5** Long-term incidence of developing new-onset cardiovascular events in patients with an initially normal hemoglobin (Hb) of > 10 mg/dL stratified by AI-ECG prediction. Based on the operating points, AI-ECG prediction was classified as normal ECG-Hb, low ECG-Hb, and severe low ECG-Hb, the same as with the previous ROC curve analysis. The analyses were conducted both in internal and external validation sets. The table shows the at-risk population and cumulative risk for the given time intervals in each risk stratification

validation set yielded outcomes consistent with those of the internal set, thereby emphasizing the ability of ECG-Hb to identify cardiovascular precursors.

## Discussion

This study presents and evaluates ECG-Hb as a novel approach for detecting anemia. The model demonstrated promising diagnostic accuracy with robust correlations between ECG-Hb and laboratory-based Hb levels, particularly in distinguishing between moderate and severe anemia. In addition, this study highlights the significance of specific ECG features in predicting Hb levels and offers insights into possible mechanistic pathways. Notably, ECG-Hb exhibits prognostic significance as it is associated with all-cause mortality and the incidence of heart failure.

Compared to a previous study with an AUC exceeding 0.85 [29], our approach achieved a comparable performance in anemia detection, with AUCs surpassing 0.82. This discrepancy stems from the distribution of Hb levels. A previous study was conducted in a population with an average Hb of 13.5 g/dL and an anemia prevalence of merely 4–5% [29], contrasting sharply with our study's 18–22% prevalence. It is more challenging to differentiate between Hb levels of 9 g/dL and 11 g/dL than between those of 9 and 15 g/dL. Thus, our accuracy is similar to that of previous reports. Notably, our accuracy is not inferior to that of widely implemented screening tests, such as breast cancer screening (AUC = 0.78) [38] and fecal occult blood tests (AUC = 0.71) [39]. Given the cost-effectiveness, ubiquity, and frequent use of ECG, AI-ECG can potentially detect asymptomatic anemia with satisfactory precision. This AI model is particularly suited for anemia screening in resource-limited regions that may lack standardized testing procedures and equipment. Reports indicate anemia prevalence rates as high as 30% in women from these areas [1], making our study potentially more representative of real-world applications.

Our study determined that AI-ECG primarily relies on ECG and HRV parameters when estimating ECG-Hb levels. Anemia appears to prolong the QT interval, potentially due to subendocardial ischemia resulting from a demand/supply mismatch in the coronary physiology of anemic individuals [40]. Moreover, past studies have found that barring

high-frequency power, each 1 g/dL decrease in hemoglobin correlated with increased odds of low HRV [41]. Additionally, anemia elevates heart rate, attributed to hypoxia-stimulated chemoreceptors and heightened sympathetic activity [42]. These evidence and our findings suggest a potential correlation between ECG and anemia, although physicians may struggle to discern anemia from these subtle signs. With the advancement of deep learning, its capability to uncover unknown correlations has been validated [43]. Our study confirms the relevance of deep learning in predicting Hb from ECG, suggesting potential for future exploration of ECG's relationship with other diseases using this technology.

The primary advancement of this study over prior research lies in introducing a continuous ECG-Hb measure to assess the severity of anemia rather than relying on probability outputs from traditional machine learning models. This continuous measure provides improved interpretability of disease severity. However, note that ECG-Hb is not a direct substitute for Lab-Hb. Despite their moderate to high correlation ( $r = 0.53$ – $0.56$ ), understanding discrepancies, especially false positives, is vital. Our analysis revealed that while the NPV of our AI model consistently exceeded 97%, the PPV was considerably lower (19.7–19.8%). We identified a significant association between false-positive results and anemia-related complications. This aligns with previous findings, suggesting that AI-ECG is capable of pinpointing cardiovascular precursors [44]. Furthermore, our analysis indicated that positive AI-ECG predictions stemmed from a series of ECG changes. Patients with abnormal ECGs are at an elevated risk of cardiovascular disease. Beyond the heart changes indicated by abnormal ECGs, certain rhythms may require intervention. For instance, atrial fibrillation is linked to an increased stroke risk and requires treatment [45]. Given the potential benefits of early detection and management, physicians should prioritize positive AI-ECG predictions, extending beyond just anemia.

Globally, areas with the most significant prevalence of anemia are mostly those with substantial resource limitations [1]. ECGs are considerably more affordable and user-friendly than blood analysis machines, making them ideal for anemia screening in these areas. With the aid of AI, ECGs can simultaneously detect multiple diseases [44], offering cost-effective large-scale screening in resource-limited regions. Despite the ubiquity of blood tests in developed countries, over three million ECGs are conducted daily worldwide [46]. Consequently, some patients with anemia may only undergo ECGs without blood test verification. Opportunistic screening, an emerging concept originating from the incidental discovery of "incidentalomas" in radiology [47], has gained traction [48]. Radiologists have embraced AI models that re-analyze radiological images to identify potential diseases. Likewise, ECGs offer a similar potential [49]. A randomized controlled trial has shown that

opportunistic AI-ECG analysis can identify an additional 32% of patients with left ventricular dysfunction [50]. Future research should explore AI-ECG's potential to detect more asymptomatic anemic patients, facilitating early diagnosis and intervention to improve patient outcomes.

Although this study presents promising findings regarding the development and evaluation of ECG-Hb for the detection of anemia, it is essential to acknowledge certain limitations. First, we used retrospective data from two specific hospitals, which may limit the generalizability of the model to broader patient populations and healthcare settings. The differences in inclusion periods and baseline characteristics among the cohorts used for model training, tuning, and validation could have influenced the observed prevalence of HRV measures and ECG patterns. Additionally, this study primarily focused on the relationship between ECG-Hb and anemia, emphasizing diagnostic accuracy and prognostic value. Potential confounding variables and underlying mechanisms, such as systemic conditions affecting autonomic nervous system, QT interval or HRV, that may contribute to the observed associations were not explored. Furthermore, while ECG-Hb demonstrates promising results in anemia detection and its implications for cardiovascular outcomes, the practical implementation of this model in real-world clinical practice and resource-constrained regions requires further validation. Future studies should assess its cost-effectiveness, scalability, and sustainability, as the financial and technical resources required to establish and maintain such an AI system must be carefully evaluated. Moreover, the current study does not allow a direct comparison between the AI-ECG model and conventional clinical evaluation, including history taking, physical examination, or longitudinal ECG assessment. Integrating serial ECG changes or combining AI-ECG analysis with clinical evaluation and historical data could further enhance diagnostic performance and prognostic capabilities. Finally, this study relied on historical data. The rapidly evolving field of deep learning and big data healthcare may necessitate ongoing validation and adaptation of the ECG-Hb model to ensure its relevance and efficacy in the ever-changing healthcare landscape.

## Conclusion

ECG-Hb is useful for detecting anemia and cardiovascular risk factors. The model showed the efficacy of ECG-based biomarkers in anemia identification, underlined by a robust correlation with lab-measured Hb levels. Its diagnostic precision, especially in discerning moderate-to-severe anemia, and its prognostic relevance linked to mortality and

heart failure onset highlight the transformative potential of AI in disease screening, especially in resource-limited settings. Future studies should explore its practical deployment and physiological underpinnings to maximize its clinical applications.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11739-025-04209-6>.

**Funding** Open access funding provided by National Tsing Hua University. None.

## Declarations

**Competing interests** The authors have nothing to disclose.

**Consent to participate** This retrospective study was approved by the Ethics Committee of the Tri-Service General Hospital (C202105049) to collect medical data. As we utilized encrypted and de-identified data from the hospital, a waiver for informed consent was granted by the data controller for this study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Mason J, Martorell R, Saldanha L, Shrimpton R (2013) Reduction of anaemia. *Lancet Glob Health* 1:e4-6. [https://doi.org/10.1016/s2214-109x\(13\)70009-3](https://doi.org/10.1016/s2214-109x(13)70009-3)
- World Health Organization. (2023) Accelerating anaemia reduction: a comprehensive framework for action. <https://www.who.int/publications/i/item/9789240074033>
- Wiciński M, Liczner G, Cadelski K, Koźnierak T, Nowaczewska M, Malinowski B (2020) Anemia of chronic diseases: wider diagnostics-better treatment? *Nutrients* 12(6):1784. <https://doi.org/10.3390/nu12061784>
- Merchant AA, Roy CN (2012) Not so benign haematology: anaemia of the elderly. *Br J Haematol* 156:173–185. <https://doi.org/10.1111/j.1365-2141.2011.08920.x>
- Han SV, Park M, Kwon YM, Yoon HJ, Chang Y, Kim H, Lim YH, Kim SG, Ko A (2019) Mild anemia and risk for all-cause, cardiovascular and cancer deaths in apparently healthy elderly Koreans. *Korean J Fam Med* 40:151–158. <https://doi.org/10.4082/kjfm.17.0089>
- Kaiafa G, Kanellos I, Savopoulos C, Kakaletsis N, Giannakoulas G, Hatzitolios AI (2015) Is anemia a new cardiovascular risk factor? *Int J Cardiol* 186:117–124. <https://doi.org/10.1016/j.ijcard.2015.03.159>
- Tyburski EA, Gillespie SE, Stoy WA, Mannino RG, Weiss AJ, Siu AF, Bulloch RH, Thota K, Cardenas A, Session W, Khoury

- HJ, O'Connor S, Bunting ST, Boudreaux J, Forest CR, Gaddh M, Leong T, Lyon LA, Lam WA (2014) Disposable platform provides visual and color-based point-of-care anemia self-testing. *J Clin Invest* 124:4387–4394. <https://doi.org/10.1172/jci76666>
8. Garcia-Casal MN, Dary O, Jefferds ME, Pasricha SR (2023) Diagnosing anemia: challenges selecting methods, addressing underlying causes, and implementing actions at the public health level. *Ann N Y Acad Sci* 1524:37–50. <https://doi.org/10.1111/nyas.14996>
  9. de Oliveira Deucher RA, de Sá Ferreirs A, da Silva Nascimento LPA, da Cal MS, Papanthasiou JV, Lopes AJ (2021) Heart rate variability in adults with sickle cell anemia during a multitasking field test. *Asian J Sports Med* 12(3):108537
  10. Lutfi MF (2011) Effects of hemoglobin concentration on heart rate variability. *Age (years)* 25:24.29
  11. Gv S, Pk S, Herur A, Chinagudi S, Patil SS, Ankad RB, Badami SV (2014) Correlation between haemoglobin level and electrocardiographic (ECG) findings in anaemia: a cross-sectional study. *J Clin Diagn Res* 8(4):BC04-06. <https://doi.org/10.7860/jcdr/2014/8966.4202>
  12. Coppler PJ, Elmer J, Calderon L, Sabedra A, Doshi AA, Callaway CW, Rittenberger JC, Dezfulian C (2015) Validation of the Pittsburgh Cardiac Arrest Category illness severity score. *Resuscitation* 89:86–92. <https://doi.org/10.1016/j.resuscitation.2015.01.020>
  13. Liu WC, Lin CS, Tsai CS, Tsao TP, Cheng CC, Liou JT, Lin WS, Cheng SM, Lou YS, Lee CC, Lin C (2021) A deep-learning algorithm for detecting acute myocardial infarction. *EuroIntervention* 17:765–773. <https://doi.org/10.4244/eij-d-20-01155>
  14. Liu WC, Lin C, Lin CS, Tsai MC, Chen SJ, Tsai SH, Lin WS, Lee CC, Tsao TP, Cheng CC (2021) An artificial intelligence-based alarm strategy facilitates management of acute myocardial infarction. *J Pers Med* 11(11):1149. <https://doi.org/10.3390/jpm11111149>
  15. Liu YL, Lin CS, Cheng CC, Lin C (2022) A deep learning algorithm for detecting acute pericarditis by electrocardiogram. *J Pers Med* 12:1150. <https://doi.org/10.3390/jpm12071150>
  16. Chang DW, Lin CS, Tsao TP, Lee CC, Chen JT, Tsai CS, Lin WS, Lin C (2021) Detecting digoxin toxicity by artificial intelligence-assisted electrocardiography. *Int J Environ Res Public Health* 18:3839. <https://doi.org/10.3390/ijerph18073839>
  17. Lee CC, Lin CS, Tsai CS, Tsao TP, Cheng CC, Liou JT, Lin WS, Lee CC, Chen JT, Lin C (2022) A deep learning-based system capable of detecting pneumothorax via electrocardiogram. *Eur J Trauma Emerg Surg* 48:3317–3326. <https://doi.org/10.1007/s00068-022-01904-3>
  18. Liu WT, Lin CS, Tsao TP, Lee CC, Cheng CC, Chen JT, Tsai CS, Lin WS, Lin C (2022) A deep-learning algorithm-enhanced system integrating electrocardiograms and chest X-rays for diagnosing aortic dissection. *Can J Cardiol*. <https://doi.org/10.1016/j.cjca.2021.09.028>
  19. Lin C, Lin CS, Lee DJ, Lee CC, Chen SJ, Tsai SH, Kuo FC, Chau T, Lin SH (2021) Artificial intelligence assisted electrocardiography for early diagnosis of thyrotoxic periodic paralysis. *J Endocr Soc* 5:bvab120. <https://doi.org/10.1210/jendso/bvab120>
  20. Lin CS, Lin C, Fang WH, Hsu CJ, Chen SJ, Huang KH, Lin WS, Tsai CS, Kuo CC, Chau T, Yang SJ, Lin SH (2020) A deep-learning algorithm (ECG12Net) for detecting hypokalemia and hyperkalemia by electrocardiography: algorithm development. *JMIR Med Inform* 2020(3):e15931. <https://doi.org/10.2196/15931>
  21. Lou YS, Lin CS, Fang WH, Lee CC, Ho CL, Wang CH, Lin C (2022) Artificial intelligence-enabled electrocardiogram estimates left atrium enlargement as a predictor of future cardiovascular disease. *J Pers Med* 12:315. <https://doi.org/10.3390/jpm12020315>
  22. Chen HY, Lin CS, Fang WH, Lou YS, Cheng CC, Lee CC, Lin C (2022) Artificial intelligence-enabled electrocardiography predicts left ventricular dysfunction and future cardiovascular outcomes: a retrospective analysis. *J Pers Med* 12:455. <https://doi.org/10.3390/jpm12030455>
  23. Chen HY, Lin CS, Fang WH, Lee CC, Ho CL, Wang CH, Lin C (2022) Artificial intelligence-enabled electrocardiogram predicted left ventricle diameter as an independent risk factor of long-term cardiovascular outcome in patients with normal ejection fraction. *Front Med* 9:870523. <https://doi.org/10.3389/fmed.2022.870523>
  24. Lee CH, Liu WT, Lou YS, Lin CS, Fang WH, Lee CC, Ho CL, Wang CH, Lin C (2022) Artificial intelligence-enabled electrocardiogram screens low left ventricular ejection fraction with a degree of confidence. *Digital Health* 8:20552076221143249. <https://doi.org/10.1177/20552076221143249>
  25. Lin C, Chau T, Lin CS, Shang HS, Fang WH, Lee DJ, Lee CC, Tsai SH, Wang CH, Lin SH (2022) Point-of-care artificial intelligence-enabled ECG for dyskalemia: a retrospective cohort analysis for accuracy and outcome prediction. *NPJ Digit Med* 5:8. <https://doi.org/10.1038/s41746-021-00550-0>
  26. Chang CH, Lin CS, Luo YS, Lee YT, Lin C (2022) Electrocardiogram-based heart age estimation by a deep learning model provides more information on the incidence of cardiovascular disorders. *Front Cardiovasc Med* 9:754909. <https://doi.org/10.3389/fcvm.2022.754909>
  27. Tsai DJ, Lou YS, Lin CS, Fang WH, Lee CC, Ho CL, Wang CH, Lin C (2023) Mortality risk prediction of the electrocardiogram as an informative indicator of cardiovascular diseases. *Digital Health* 9:20552076231187247. <https://doi.org/10.1177/20552076231187247>
  28. Chen YJ, Lin CS, Lin C, Tsai DJ, Fang WH, Lee CC, Wang CH, Chen SJ (2023) An AI-enabled dynamic risk stratification for emergency department patients with ECG and CXR integration. *J Med Syst* 47(1):81. <https://doi.org/10.1007/s10916-023-01980-x>
  29. Kwon JM, Cho Y, Jeon KH, Cho S, Kim KH, Baek SD, Jeung S, Park J, Oh BH (2020) A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. *Lancet Digit Health* 2(7):e358–e367. [https://doi.org/10.1016/s2589-7500\(20\)30108-4](https://doi.org/10.1016/s2589-7500(20)30108-4)
  30. Van Buuren S, Groothuis-Oudshoorn K (2011) Mice: multivariate imputation by chained equations in R. *J Stat Softw* 45(3):1–67
  31. Sassi R, Cerutti S, Lombardi F, Malik M, Huikuri HV, Peng CK, Schmidt G, Yamamoto Y (2015) Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society. *Europace* 17:1341–1353. <https://doi.org/10.1093/europace/euv015>
  32. Lou YS, Lin CS, Fang WH, Lee CC, Wang CH, Lin C (2023) Development and validation of a dynamic deep learning algorithm using electrocardiogram to predict dyskalaemias in patients with multiple visits. *Eur Heart J Digit Health* 4(1):22–32. <https://doi.org/10.1093/ehjdh/ztac072>
  33. Lou YS, Lin CS, Fang WH, Lee CC, Lin C (2023) Extensive deep learning model to enhance electrocardiogram application via latent cardiovascular feature extraction from identity identification. *Comput Methods Programs Biomed* 231:107359. <https://doi.org/10.1016/j.cmpb.2023.107359>
  34. Lin C, Chen CC, Chau T, Lin CS, Tsai SH, Lee DJ, Lee CC, Shang HS, Lin SH (2022) Artificial intelligence-enabled electrocardiography identifies severe dyscalcemias and has prognostic value. *Clin Chim Acta* 2022:126–134. <https://doi.org/10.1016/j.cca.2022.09.021>
  35. Lee YT, Lin CS, Fang WH, Lee CC, Ho CL, Wang CH, Tsai DJ, Lin C (2022) Artificial intelligence-enabled electrocardiography detects hypoalbuminemia and identifies the mechanism of hepatorenal and cardiovascular events. *Front Cardiovasc Med* 2022(9):895201. <https://doi.org/10.3389/fcvm.2022.895201>

36. Liu PY, Lin C, Lin CS, Fang WH, Lee CC, Wang CH, Tsai DJ (2023) Artificial intelligence-enabled electrocardiography detects B-type natriuretic peptide and N-terminal pro-brain natriuretic peptide. *Diagnostics* 13:2723. <https://doi.org/10.3390/diagnostic13172723>
37. Lin CS, Lee YT, Fang WH, Lou YS, Kuo FC, Lee CC, Lin C (2021) Deep learning algorithm for management of diabetes mellitus via electrocardiogram-based glycosylated hemoglobin (ECG-HbA1c): a retrospective cohort study. *J Pers Med* 11:725. <https://doi.org/10.3390/jpm11080725>
38. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LL, Bassett L, D'Orsi C, Jong R (2005) Rebner M (2005) diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 353(17):1773–1783. <https://doi.org/10.1056/NEJMoa052911>
39. Haug U, Kuntz KM, Knudsen AB, Hundt S (2011) Brenner H (2011) sensitivity of immunochemical faecal occult blood testing for detecting left- vs right-sided colorectal neoplasia. *Br J Cancer* 104(11):1779–1785. <https://doi.org/10.1038/bjc.2011.160>
40. Hünük B, Çağaç Ö, Erdoğan O, Kepez A, Mutlu B, Değertekin M (2013) Erol Ç (2013) The impact of anemia on QT interval: a population based study. *JACC* 62(18 suppl C):C146
41. Gehi A, Ix J, Shlipak M, Pipkin SS, Whooley MA (2005) Relation of anemia to low heart rate variability in patients with coronary heart disease (from the Heart and Soul study). *Am J Cardiol* 2005(12):1474–1477. <https://doi.org/10.1016/j.amjcard.2005.02.017>
42. Metivier F, Marchais SJ, Guerin AP, Pannier B, London GM (2000) Pathophysiology of anaemia: focus on the heart and blood vessels. *Nephrol Dial Transplant* 2000(3):14–18. <https://doi.org/10.1093/oxfordjournals.ndt.a027970>
43. Wang H, Fu T, Du Y, Gao W, Huang K, Liu Z, Chandak P, Liu S, Van Katwyk P, Deac A, Anandkumar A, Bergen K, Gomes CP, Ho S, Kohli P, Lasenby J, Leskovec J, Liu TY, Manrai A, Marks D, Ramsundar B, Song L, Sun J, Tang J, Veličković P, Welling M, Zhang L, Coley CW, Bengio Y, Zitnik M (2023) Scientific discovery in the age of artificial intelligence. *Nature* 620:47–60. <https://doi.org/10.1038/s41586-023-06221-2>
44. Attia ZI, Harmon DM, Behr ER, Friedman PA (2021) Application of artificial intelligence to the electrocardiogram. *Eur Heart J* 2021(46):4717–4730. <https://doi.org/10.1093/eurheartj/ehab649>
45. Gladstone DJ, Spring M, Dorian P, Panzov V, Thorpe KE, Hall J, Vaid H, O'Donnell M, Laupacis A, Côté R, Sharma M, Blakely JA, Shuaib A, Hachinski V, Coutts SB, Sahlas DJ, Teal P, Yip S, Spence JD, Buck B, Verreault S, Casaubon LK, Penn A, Selchen D, Jin A, Howse D, Mehdiratta M, Boyle K, Aviv R, Kapral MK, Mamdani M (2014) Atrial fibrillation in patients with cryptogenic stroke. *N Engl J Med* 2014(26):2467–2477. <https://doi.org/10.1056/NEJMoa1311376>
46. Shenasa M (2018) Learning and teaching electrocardiography in the 21st century: a neglected art. *J Electrocardiol*. <https://doi.org/10.1016/j.jelectrocard.2018.02.007>
47. Berland LL, Silverman SG, Gore RM, Mayo-Smith WW, Megibow AJ, Yee J, Brink JA, Baker ME, Federle MP, Foley WD, Francis IR, Herts BR, Israel GM, Krinsky G, Platt JF, Shuman WP, Taylor AJ (2010) Managing incidental findings on abdominal CT: white paper of the ACR incidental findings committee. *J Am Coll Radiol* 7(10):754–773. <https://doi.org/10.1016/j.jacr.2010.06.013>
48. Pickhardt PJ, Summers RM, Garrett JW, Krishnaraj A, Agarwal S, Dreyer KJ, Nicola GN (2023) Opportunistic screening: radiology scientific expert panel. *Radiology* 307:e222044. <https://doi.org/10.1016/j.radiol.2022044>
49. Bluemke DA, Moy L, Bredella MA, Ertl-Wagner BB, Fowler KJ, Goh VJ, Halpern EF, Hess CP, Schiebler ML, Weiss CR (2020) Assessing radiology research on artificial intelligence: a brief guide for authors, reviewers, and readers-from the Radiology Editorial Board. *Radiology* 294:487–489. <https://doi.org/10.1148/radiol.2019192515>
50. Yao X, Rushlow DR, Inselman JW, McCoy RG, Thacher TD, Behnken EM, Bernard ME, Rosas SL, Akfaly A, Misra A, Molling PE, Krien JS, Foss RM, Barry BA, Siontis KC, Kapa S, Pellikka PA, Lopez-Jimenez F, Attia ZI, Shah ND, Friedman PA, Noseworthy PA (2021) Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. *Nat Med* 27:815–819. <https://doi.org/10.1038/s41591-021-01335-4>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.