BMC Medical Informatics and Decision Making

RESEARCH

Open Access

Electrocardiogram-based deep learning score for coronary artery calcification reclassifies cardiovascular risk and identifies screening candidates

Changho Han[1], Seng Chan You[1,2], Hyung-Chul Lee[3], Jin Young Park[4,5,6], Hong-Seok Lim[7], ChulHyoung Park[8], Hui-Nam Pak[9], Oh-Seok Kwon[2,9], Songsoo Kim[1], Jung-Sun Kim[10*] and Dukyong Yoon[1,2,6,11*]

Abstract

Background Coronary artery calcium (CAC) scoring is included in major guidelines to guide statin decisions when conventional cardiovascular risk assessments are inconclusive. While computed tomography for CAC is impractical for routine use, electrocardiograms (ECGs) are widely available, offering broader opportunities for early detection. Extending our prior work demonstrating an ECG-based deep learning model for CAC prediction (ECG-CAC model), we aimed to develop a more robust model, while also comprehensively assessing its clinical utility.

Methods Using nearly 200,000 standard 12-lead ECGs from Severance Hospital (SH), we developed an ECG-CAC model producing a risk score reflecting the likelihood of CAC. We utilized data from three health checkup centers and the United Kingdom Biobank (UKB) to evaluate its performance in predicting CAC, potential for guiding targeted CAC screening and statin therapy decisions, and association with major adverse cardiovascular events (MACE). MACE was defined as a composite of myocardial infarction, ischemic stroke, and cardiovascular death.

Results The ECG-CAC model achieved areas under the receiver operating characteristic curve of 0.741 (95% confidence interval: 0.732–0.749)/0.837 (95% confidence interval: 0.821–0.851) for predicting CAC scores > 0/≥400, with robust external validation. Among Pooled Cohort Equations (PCE) low-risk individuals, for whom statins are not recommended unless CAC-positive, CAC prevalence rose from 24.9% to 49.8% in those reclassified as ECG-CAC model high-risk ($P < 0.001$), and among PCE moderate-risk individuals, for whom statins are recommended unless CAC-negative, CAC absence rose from 35.4% to 53.9% in those reclassified as ECG-CAC model low-risk ($P < 0.001$), supporting targeted CAC screening in these groups. MACE incidence rate per 1000 person-years was higher in PCE low-risk individuals reclassified as ECG-CAC model high-risk than in PCE moderate-risk individuals reclassified as ECG-CAC model low-risk (SH: 6.4 vs. 3.4, $P = 0.026$; UKB: 8.1 vs. 5.1, $P = 0.183$), suggesting that statin therapy could be

*Correspondence:
Jung-Sun Kim
kjs1218@yuhs.ac
Dukyong Yoon
dukyong.yoon@yonsei.ac.kr

Full list of author information is available at the end of the article

considered in the former and potentially deferred in the latter, pending further clinical evaluation. ECG-CAC model score was an independent MACE risk factor and improved prediction when combined with conventional tools.

Conclusions The ECG-CAC model serves as an indicator for CAC and shows potential as a biomarker of coronary-cerebrovascular atherosclerotic burden, providing incremental value to conventional cardiovascular risk stratification tools.

Keywords Coronary artery calcification, Coronary artery calcium score, Deep learning, Electrocardiogram, Artificial intelligence

Background

Coronary artery calcium (CAC) is a representative marker of coronary atherosclerosis and one of the strongest predictors of atherosclerotic cardiovascular disease (ASCVD) events [1–3]. The American College of Cardiology/American Heart Association (ACC/AHA) Guideline on the Management of Blood Cholesterol in 2018 [4], incorporated the CAC score when considering statin therapy for primary prevention in individuals with borderline or intermediate cardiovascular disease (CVD) risk, as assessed by the Pooled Cohort Equations (PCE) [5]. The CAC score serves as a sequential decision aid when uncertainty remains after standard risk assessment using traditional risk factors. In such cases, the presence of CAC tips the balance toward initiating statin therapy, while the absence of CAC may lead to recommendations to withhold statins.

However, routine CAC measurement faces significant barriers. Computed tomography (CT)-based CAC scoring is costly, involves radiation, and is often inaccessible in resource-limited settings. These limitations highlight the need for alternative methods to identify the disease early in a cost-effective and accessible way.

An electrocardiogram (ECG) is a universal and cost-effective diagnostic tool commonly utilized in various health evaluations. Numerous studies have demonstrated the successful application of artificial intelligence (AI) to ECGs [6–8]. A previous study has demonstrated that an ECG-based deep learning analysis can predict CAC, achieving areas under the receiver operating characteristics curve (AUROC) of 0.696 and 0.802 for predicting CAC scores > 0 and ≥ 400, respectively [9]. However, this study had some limitations. First, it only demonstrated predictive performance for CAC without evaluating clinical utility, including its impact on decision-making in CVD primary prevention and its association with clinical outcomes. Second, while CAC screening via an ECG-based deep learning model would have utility for CVD primary prevention, validation was conducted in a tertiary hospital setting without focusing on primary prevention candidates. Third, the relatively small dataset (8,178 ECGs from 5,765 individuals) used for training and validation may have limited the model's generalizability.

In this study, we aimed to expand upon and address the limitations of the previous study. We developed an ECG-based deep learning model for CAC prediction (ECG-CAC model) based on a larger standard 12-lead ECG dataset than the previous study. We validated the model using health checkup data from multiple institutions and conducted multinational retrospective cohort analyses for cardiovascular events (Fig. 1A). Through these analyses, we evaluated the feasibility of integrating this ECG-CAC model score into clinical workflows by assessing its potential for guiding targeted CAC screening and statin therapy decisions. Additionally, we assessed its association with cardiovascular events.

Methods

Data sources and labeling

Figure 1 provides an overview of the study and datasets. CAC scores were extracted using regular expressions from CT readings at Severance Hospital (SH, Seoul, South Korea) between November 2005 and August 2022, and from CT readings at the SH Health Checkup Center between December 2010 and August 2022.

Standard 12-lead ECGs from SH were extracted if recorded within 30 days before or after CAC measurements and were labeled with the nearest CAC score (Additional file 1: Figure S1). A 30-day window is valid given that CAC progression is slow enough to remain stable within this time frame [10]. These ECGs formed the model development dataset, which was split 8:2 into training and validation sets using a random stratified split based on CAC presence, while ensuring no individual overlap.

Standard 12-lead ECGs from SH Health Checkup Center with same-visit CAC measurements were extracted and labeled with their corresponding CAC scores, forming the test dataset (Additional file 1: Figure S2). Standard 12-lead ECGs from SH Health Checkup Center without CAC measurements (September 2012 to September 2016) formed the cohort analysis dataset. To prevent data leakage, appropriate exclusions were made to ensure no individual overlap across the SH and SH Health Checkup Center datasets.

External validation utilized data from the Health Checkup Centers at Yongin Severance Hospital (YSH,

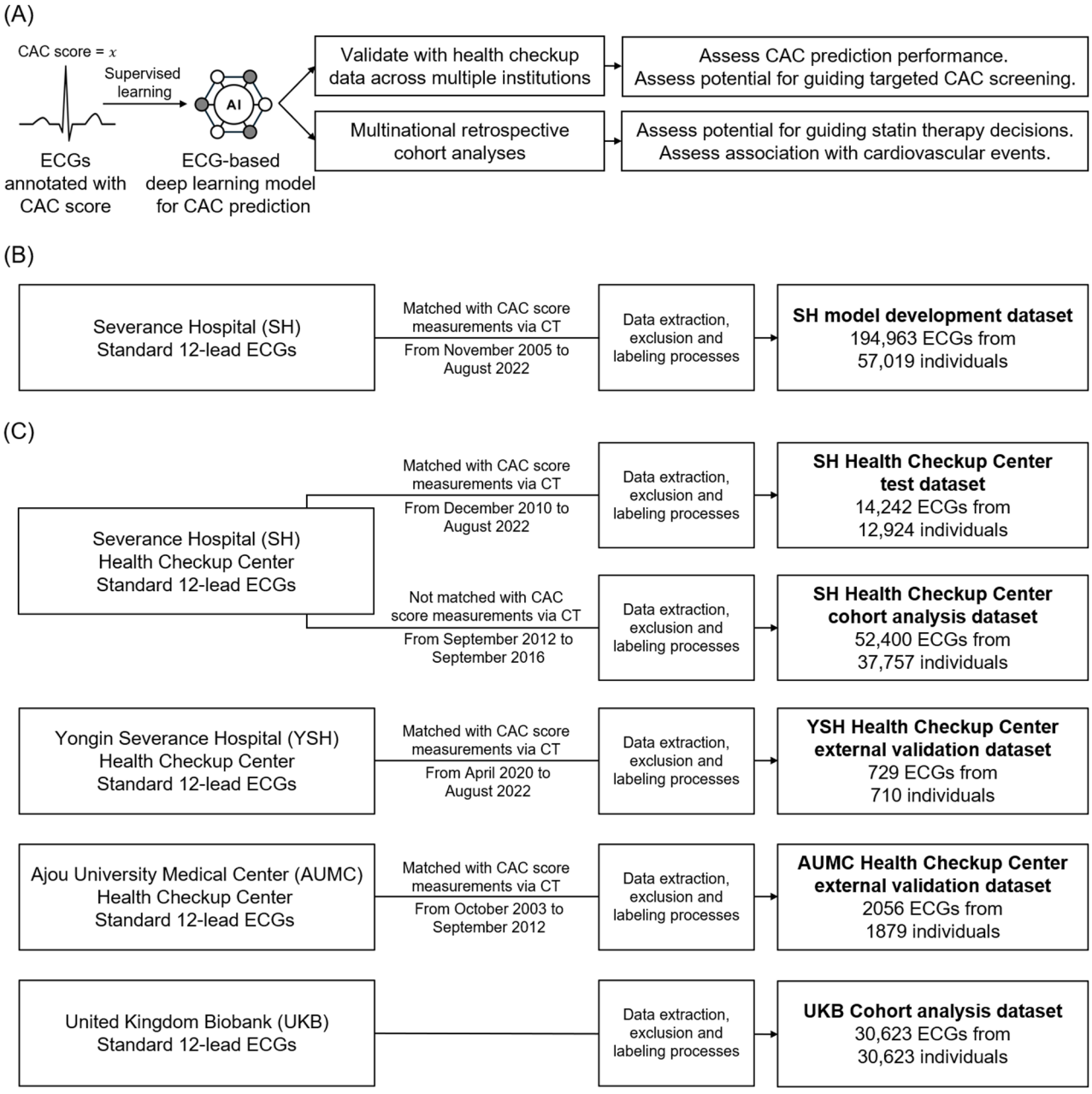

Fig. 1 Overview of the study and the datasets. (**A**) We developed an ECG-based deep learning model for CAC prediction (ECG-CAC model) using CAC score-annotated ECGs. We validated it using health checkup data from multiple institutions, and conducted multinational retrospective cohort analyses for cardiovascular events. Through these analyses, we evaluated the feasibility of integrating our ECG-CAC model into clinical workflows by assessing its potential for guiding targeted CAC screening and statin therapy decisions. Additionally, we assessed its association with cardiovascular events. (**B**) Overview of the dataset used for ECG-CAC model development, including training and internal validation. (**C**) After developing the model using the dataset shown in (**B**), we applied the ECG-CAC model to the datasets presented in (**C**) to generate ECG-CAC model scores and perform subsequent analyses. CAC: coronary artery calcium; ECG: electrocardiogram; AI: artificial intelligence; SH: Severance Hospital; YSH: Yongin Severance Hospital; AUMC: Ajou University Medical Center; UKB: United Kingdom Biobank; CT: computed tomography

Yongin, South Korea) (April 2020 to August 2022) and Ajou University Medical Center (AUMC, Suwon, South Korea) (October 2003 to September 2012) (Additional file 1: Figure S3). CAC scores were obtained from the Health Checkup Centers at YSH and AUMC using the same methodology as at SH. Standard 12-lead ECGs with same-visit CAC measurements were collected and labeled with their corresponding CAC scores.

We utilized data from the United Kingdom Biobank (UKB), a large-scale longitudinal biomedical database representing the general population of the United Kingdom, comprising approximately 500,000 participants

aged 40–69 at enrollment (started in 2006) [11]. In 2015, the UKB began an imaging study targeting 20% of the cohort, during which standard 12-lead ECG data were collected and used in our analysis (Additional file 1: Figure S4, Additional file 1: Table S1).

ECG-CAC model development

We utilized the raw waveforms of the standard 12-lead ECGs as input and adopted the 1-dimensional EfficientNet-B0 for our ECG-CAC model architecture (Additional file 1: Figure S5) [12]. We set the training objective of the ECG-CAC model to predict CAC score > 0, a binary classification task. The ECG-CAC model generates a continuous score between 0 and 1 representing the likelihood of CAC, with higher values indicating greater confidence in the presence of CAC. All datasets consisted of 10-second ECG recordings and underwent identical preprocessing. The signals, originally at 250–500 Hz, were standardized to 500 Hz by upsampling the 250 Hz signals via linear interpolation. Each lead was z-normalized (mean = 0, standard deviation = 1). Since four of the six limb leads can be linearly derived from any two leads using the Einthoven law and the Goldberger Eqs [13, 14]., we used eight leads (I, II, V1-V6) as input. Hyperparameter optimization was achieved through comprehensive grid search. The model was trained on the training set, with early stopping based on validation loss (validation set) to prevent overfitting, and then applied to the test dataset, external validation datasets, and cohort analysis datasets. The model produces an output score ranging from 0 to 1, where higher scores indicate a greater likelihood of CAC as predicted by the model. Detailed strategies for data augmentation, class imbalance handling, and ensemble modeling are described in Additional file 1: Method S1.

Performance metrics and risk categorization

We evaluated the AUROC and area under the precision-recall curve (AUPRC) of the ECG-CAC model for predicting CAC score > 0. We also evaluated the model's performance in predicting CAC score ≥ 400, a clinically recognized threshold indicating high cardiovascular event risk and classified as "severe disease" [3, 15], using the same model output score but with labels set to identify whether CAC scores exceeded 400.

We categorized PCE risk scores into low-, moderate-, and high-risk groups using the 7.5% and 20% thresholds, following current cholesterol management guidelines [4, 5]. In the PCE low-risk category, guidelines indicate statins are initially not recommended, but the presence of CAC can support the decision to initiate statin therapy [4, 5], therefore it would be advantageous if the ECG-CAC model could enhance the detection of CAC presence. Conversely, in the PCE moderate-risk category, guidelines indicate statins are initially recommended, but the absence of CAC can guide the decision to withhold statin therapy [4, 5], thus it would be beneficial if the ECG-CAC model could improve the identification of CAC absence. Accordingly, in the validation set, we defined ECG-CAC model high-risk at the threshold where the positive predictive value (PPV) for a CAC score > 0 is 0.850. Similarly, we defined ECG-CAC model low-risk at the threshold where the negative predictive value (NPV) for a CAC score > 0 is 0.850. Individuals falling between these thresholds were categorized as ECG-CAC model moderate-risk. At the determined thresholds, we calculated accuracy, sensitivity, specificity, PPV, NPV, and the F1 score for CAC score > 0. We evaluated how ECG-CAC model-based risk classification, following initial PCE-based CVD risk classification, alters CAC prevalence to assess its potential for guiding targeted CAC screening.

Multinational retrospective cohort analyses

Multinational retrospective cohort analyses were conducted on major adverse cardiovascular event (MACE) occurrence using datasets from two countries: the SH Health Checkup Center cohort analysis dataset and the UKB cohort analysis dataset. MACE was defined using the International Classification of Diseases, 10th edition (ICD-10) codes as myocardial infarction (I21–I25), ischemic stroke (I63, I64), or cardiovascular death [16]. Cardiovascular death was defined using the same ICD-10 codes for myocardial infarction and ischemic stroke in mortality data.

In the SH Health Checkup Center cohort analysis dataset, diagnoses were extracted from electronic medical records (EMR) using ICD-10 codes. In South Korea, a government organization called Statistics Korea (KOSTAT) offers a service that links researchers' data with mortality data (including date and cause of death [in ICD-10 codes]), based on resident registration numbers, following specific ethical and application procedures. The SH Health Checkup Center cohort analysis dataset was linked with KOSTAT mortality data. Individuals with MACE before or within 90 days after the health checkup were excluded (washout period to exclude pre-existing cardiovascular disease detected through follow-up testing). The observation period started after the 90-day washout period and was censored at first MACE, 9 years post-health checkup, or December 2022, whichever came first. In the UKB cohort analysis dataset, the "first occurrences" variable, which maps the first known occurrence dates of each ICD-10 code from diagnoses or death records, was extracted. Individuals with MACE before ECG measurement were excluded, and data were censored at the first MACE, 6 years post-ECG, or November 2021, whichever occurred first.

We evaluated how ECG-CAC model-based risk classification, following initial PCE-based CVD risk

classification, alters MACE IRs to assess its potential role in statin therapy decision-making. The net reclassification improvement (NRI) of the ECG-CAC model-based risk classification within the PCE low- or moderate-risk categories was also evaluated. To evaluate the ECG-CAC model score as an independent risk factor for MACE, we performed Cox regression using the ECG-CAC model score along with variables included in the PCE—age, sex, diabetes mellitus, hypertension, smoking status, total cholesterol, high-density lipoprotein (HDL) cholesterol, and systolic blood pressure—as independent variables. The predictive ability of the combined PCE plus the ECG-CAC model score was compared to PCE alone in terms of Harrell's C-index. Subgroup analyses by sex, age (< 60, ≥ 60 years), and PCE risk category (low, moderate/high) were conducted.

Explainability analysis

Integrated Gradients (IG) were applied to the ECG-CAC model to quantify the contribution of each time point in every lead to the predicted likelihood of CAC [17]. After IG computation, R-peaks were detected using the NeuroKit2 algorithm [18], and each waveform with its corresponding IG saliency map was segmented around each R-peak (– 300 to + 400 ms, excluding truncated segments) to align R-peaks across samples. Within each group (CAC score = 0, > 0, or ≥ 400), the aligned waveforms and saliency maps were averaged across all subjects for each lead. Saliency values were then globally normalized across leads (min–max scaling to 0–1) and overlaid on the mean waveform, enabling detailed visualization of ECG segments most influential to model predictions. Finally, the summed IG attribution values across all ECGs were used to evaluate the relative contribution of each lead to CAC prediction.

Statistical analysis

The Shapiro-Wilk test assessed normality of continuous data. For comparing three or more groups, ANOVA was used for normally distributed variables, while Kruskal-Wallis tests were applied for non-normal distributions. Categorical data were compared using chi-square tests, and AUROC comparisons used the Delong test [19]. Kaplan-Meier survival curves were compared using log-rank tests, with Bonferroni corrections for pairwise comparisons. Z-tests determined p-values for incidence rate (IR) and hazard ratio differences. Model calibration was assessed using calibration curves constructed with quantile-based binning (five bins), and calibration performance was further evaluated using the Brier score and expected calibration error (ECE). All confidence intervals (CIs) were calculated using 2000 bootstrap samples (resampling with replacement), reporting the 2.5th and 97.5th percentiles. Statistical significance was defined as $P < 0.05$.

Results

Dataset characteristics (model development and test datasets)

Figure 1B and C provide an overview of the datasets, while Additional file 1: Figures S1 to S4 depict the detailed data flow for each dataset. The model development and test datasets comprised: SH model development (194,964 ECGs/57,019 individuals), SH Health Checkup Center test (14,242 ECGs/12,924 individuals), YSH Health Checkup Center external validation (729 ECGs/710 individuals), and AUMC Health Checkup Center external validation (2056 ECGs/1879 individuals). The test and external validation datasets, derived from health checkups, represented a healthier spectrum of individuals with younger ages (mean 53.1, 57.1, and 51.1 vs. 61.8 years, $P < 0.001$) and lower CAC scores (mean 59.0, 81.5, and 36.4 vs. 295.1 Agatston units, $P < 0.001$) than the model development dataset (Table 1).

Model performance

The ECG-CAC model, developed using standard 12-lead ECGs, generates a continuous score between 0 and 1 representing the likelihood of CAC, with higher values indicating greater confidence in the presence of CAC. Our ECG-CAC model showed AUROCs of 0.741 (95% CI: 0.732–0.749) and 0.837 (95% CI: 0.821–0.851)

Table 1 Dataset characteristics (model development and test datasets)

	SH model development dataset *N* = 194,963	SH Health Checkup Center test dataset *N* = 14,242	YSH Health Checkup Center external validation dataset *N* = 729	AUMC Health Checkup Center external validation dataset *N* = 2056	*P*-Value
Number of individuals, *n*	57,019	12,926	710	1879	
Sex (male), n (%)	108,828 (55.8%)	8,502 (59.7%)	427 (58.6%)	1540 (74.9%)	< 0.001
Age, years, mean ± SD	61.8 ± 13.3	53.1 ± 10.0	57.1 ± 10.7	51.1 ± 8.2	< 0.001
CAC score, Agatston units, mean ± SD	295.1 ± 920.9	59.0 ± 212.2	81.5 ± 274.8	36.4 ± 152.3	< 0.001
CAC score > 0, n (%)	120,837 (62.0%)	5394 (37.9%)	341 (46.8%)	606 (29.5%)	< 0.001
CAC score ≥ 400, n (%)	34,637 (17.8%)	542 (3.8%)	44 (6.0%)	52 (2.5%)	< 0.001

SH: Severance Hospital; YSH: Yongin Severance Hospital; AUMC: Ajou University Medical Center; SD: standard deviation; CAC: coronary artery calcium

in predicting CAC score > 0 and ≥ 400, respectively, in the SH Health Checkup Center test dataset (Fig. 2). The respective AUPRCs were 0.623 (95% CI: 0.608–0.638) and 0.249 (95% CI: 0.212–0.287). Additional file 1: Table S2 displays performance metrics at thresholds defining ECG-CAC model risk categories. The median and interquartile range (IQR) of the ECG-CAC model score progressively increased with higher CAC score groups (Additional file 1: Figure S6).

In the external validations, the ECG-CAC model maintained its efficacy, achieving AUROCs of 0.712 (95% CI: 0.674–0.748)/0.711 (95% CI: 0.688–0.735) and AUPRCs of 0.676 (95% CI: 0.621–0.725)/0.482 (95% CI: 0.441–0.523) for predicting CAC score > 0 in the YSH/AUMC Health Checkup Center datasets, and AUROCs of 0.800 (95% CI: 0.732–0.860)/0.799 (95% CI: 0.742–0.849) and AUPRCs of 0.260 (95% CI: 0.147–0.404)/0.096 (95% CI: 0.058–0.155) for predicting CAC score ≥ 400. AUROC

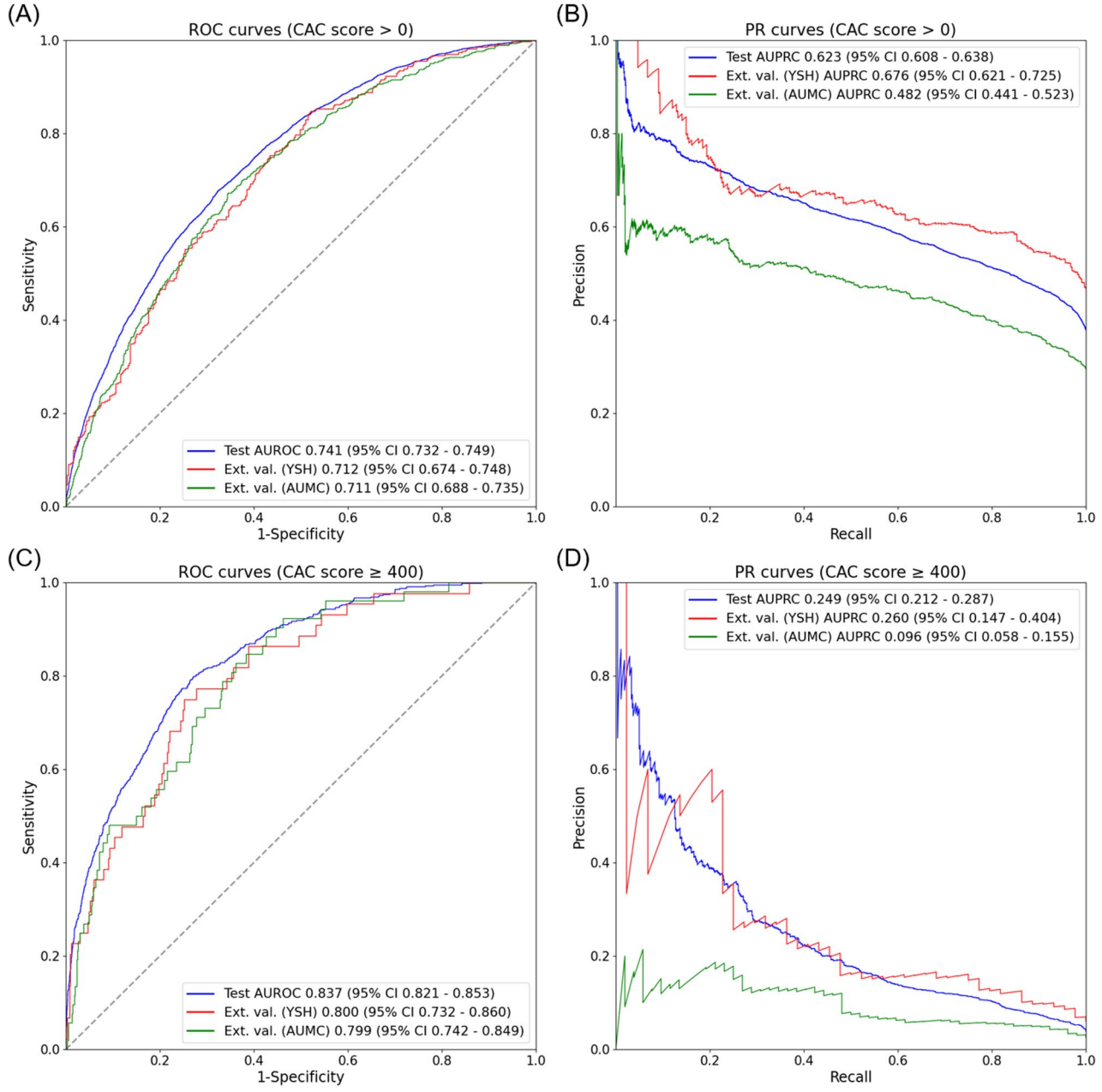

Fig. 2 ROC and PR curves of the ECG-CAC model in the health checkup datasets. (**A**, **B**) ROC and PR curves for predicting CAC score > 0. (**C**, **D**) ROC and PR curves for predicting CAC score ≥ 400. ROC: receiver operating characteristic PR: precision-recall; ECG: electrocardiogram; CAC: coronary artery calcium; ECG-CAC model: ECG-based deep learning model for CAC prediction; AUROC: area under the ROC curve; AUPRC: area under the PR curve; CI: confidence interval; Ext. val.: external validation; YSH: Yongin Severance Hospital; AUMC: Ajou University Medical Center

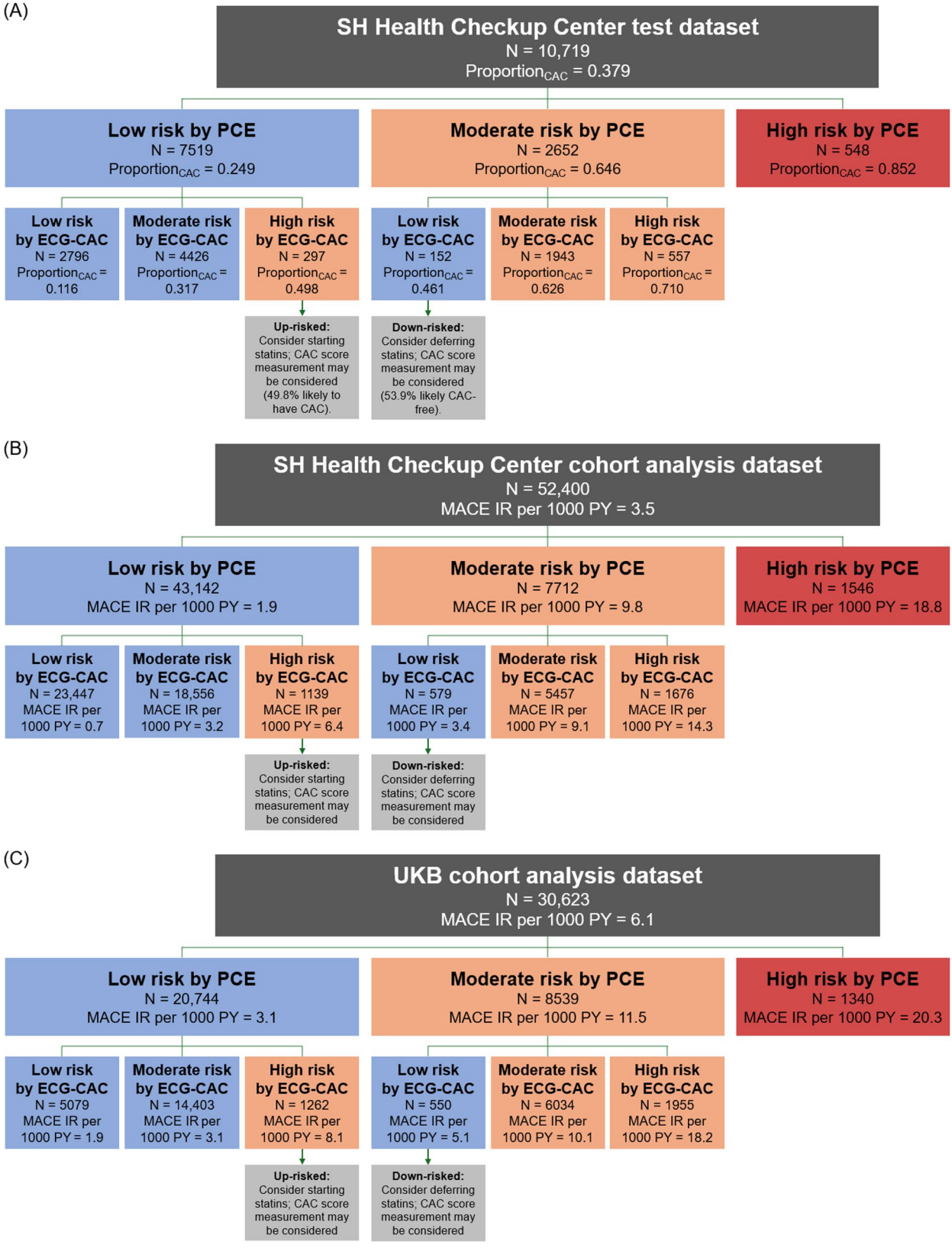

Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 ECG-CAC model-based risk classification within each PCE risk category. "Proportion$_{CAC}$" represents the proportion of individuals with CAC. (**A**) In the SH Health Checkup Center test dataset, in the PCE low-risk category, further classification to ECG-CAC model high-risk increased CAC presence (24.9% to 49.8%, $P < 0.001$), while in the PCE moderate-risk category, further classification to ECG-CAC model low-risk increased CAC absence (35.4% to 53.9%, $P < 0.001$). (**B**) In the SH Health Checkup Center cohort analysis dataset, MACE IR per 1000 PY was higher in PCE low-risk individuals further classified to ECG-CAC model high-risk compared to PCE moderate-risk individuals further classified to ECG-CAC model low-risk (6.4 vs. 3.4, $P = 0.026$). (**C**) In the UKB cohort analysis dataset, MACE IR per 1000 PY was also higher in PCE low-risk individuals further classified to ECG-CAC model high-risk compared to PCE moderate-risk individuals further classified to ECG-CAC model low-risk, though the difference was not statistically significant (8.1 vs. 5.1, $P = 0.183$). ECG: electrocardiogram; CAC: coronary artery calcium; ECG-CAC model: ECG-based deep learning model for CAC prediction; PCE: Pooled Cohort Equations; SH: Severance Hospital; MACE: major adverse cardiovascular event; IR: incidence rate; PY: person-year; UKB: United Kingdom Biobank

comparisons between test and external validation datasets predominantly showed no significant differences (Delong test [unpaired, two-sided]: SH vs. YSH $P = 0.137$, SH vs. AUMC $P = 0.020$ for CAC score > 0; SH vs. YSH $P = 0.266$, SH vs. AUMC $P = 0.186$ for CAC score ≥ 400). No statistically significant differences in AUROC were observed across demographic subgroups, suggesting consistent performance of the model among these subgroups (Additional file 1: Table S3).

In the SH test dataset, the ECG-CAC model showed excellent calibration for detecting CAC presence (ECE = 0.019 [95% CI: 0.013–0.027], Brier score = 0.197 [95% CI: 0.194–0.199]), closely following the ideal diagonal (Additional file 1: Figure S7). Calibration remained acceptable in the YSH external validation dataset (ECE = 0.064 [95% CI: 0.038–0.098], Brier score = 0.218 [95% CI: 0.205–0.231]). Calibration performance in the AUMC external validation dataset was suboptimal (ECE = 0.105 [95% CI: 0.086–0.123], Brier score = 0.197 [95% CI: 0.190–0.204]), with predicted probabilities exceeding observed CAC prevalence across risk bins, indicating systematic overestimation of risk. Overall, these results indicate that the ECG-CAC model showed good calibration in the development and some external validation settings, while calibration performance varied across institutions.

Among individuals in the SH Health Checkup Center test dataset for whom the PCE was calculated ($n = 10{,}719$), PCE categorized 7519 as low-risk, 2652 as moderate-risk, and 548 as high-risk (Fig. 3A). In the PCE low-risk group, 49.8% of those labeled high-risk by the ECG-CAC model had CAC, exceeding the overall CAC > 0 proportion in the PCE low-risk group (24.9%, $P < 0.001$). In the PCE moderate-risk group, 53.9% of those labeled low-risk by the ECG-CAC model were CAC-free, exceeding the overall CAC = 0 proportion in the PCE moderate-risk group (35.4%, $P < 0.001$).

Dataset characteristics (cohort analyses datasets)

The cohort analyses datasets comprised (Fig. 1): SH Health Checkup Center cohort analysis (52,400 ECGs/37,757 individuals), and UKB cohort analysis (30,623 ECGs/30,623 individuals). In the SH Health Checkup Center cohort analysis dataset, the lower-risk groups by the ECG-CAC model represented a healthier spectrum of individuals than the higher-risk groups, characterized, for example, by younger ages (median 40 to 52 to 61 years, $P < 0.001$), fewer comorbidities (diabetes mellitus: 1.0% to 5.7% to 14.8%, $P < 0.001$; hypertension: 3.2% to 17.6% to 41.0%, $P < 0.001$), lower systolic blood pressure (median 114 to 123 to 129 mmHg, $P < 0.001$), lower PCE score (median 0.6 to 3.9 to 10.8, $P < 0.001$), and lower MACE IR per 1000 person-years (PY) (0.8 to 4.9 to 13.0, $P < 0.001$) (Table 2). A similar trend was observed in the UKB cohort analysis dataset.

Multinational retrospective cohort analyses

In the SH Health Checkup Center cohort analysis dataset, during a median follow-up of 7.9 years (IQR: 6.9–8.9 years), 987 individuals (1428 ECGs) experienced a MACE. The PCE categorized 43,142 as low-risk, 7712 as moderate-risk, and 1546 as high-risk (Fig. 3B). Individuals classified as low-risk by PCE but high-risk by the ECG-CAC model had a MACE IR of 6.4 per 1000 PY, exceeding 3.4 in those with PCE moderate-risk but ECG-CAC model low-risk ($P = 0.026$). Kaplan-Meier curves showed significant differences based on ECG-CAC model risk categories within both PCE low-risk and moderate-risk groups (Fig. 4A, within PCE low-risk group, log-rank test $P < 0.001$; Fig. 4B, within PCE moderate-risk group, log-rank test $P < 0.001$). The net NRI for risk reclassification based on ECG-CAC model within PCE low- or moderate-risk categories was 2.3% (95% CI: 1.0% – 3.7%) (Additional file 1: Table S4). A 10% absolute increase in the ECG-CAC model score was associated with a hazard ratio for MACE of 1.128 (95% CI: 1.093–1.164) after adjusting for clinical variables included in the PCE—age, sex, diabetes mellitus, hypertension, smoking status, total cholesterol, HDL cholesterol, and systolic blood pressure (Fig. 5, Additional file 1: Table S5). The combined PCE plus ECG-CAC model score, which yielded a C-index of 0.798 (95% CI: 0.789–0.808), surpassed the PCE by a difference of 0.012 (95% CI: 0.008–0.016) in the C-index (Additional file 1: Table S6).

In the UKB cohort analysis dataset, during a median follow-up of 3.5 years (IQR: 2.5–5.0 years), 699 individuals experienced a MACE. The PCE categorized 20,744 as low-risk, 8539 as moderate-risk, and 1340 as high-risk (Fig. 3C). Individuals classified as low-risk by PCE but high-risk by the ECG-CAC model had a MACE IR of 8.1

Table 2 Dataset characteristics (cohort analysis datasets)

	SH Health Checkup Center cohort analysis dataset				UKB cohort analysis dataset			
	Low-risk by ECG-CAC *N* = 24,061	Moderate-risk by ECG-CAC *N* = 24,814	High-risk by ECG-CAC *N* = 3525	*P*-value	Low-risk by ECG-CAC *N* = 5671	Moderate-risk by ECG-CAC *N* = 21,224	High-risk by ECG-CAC *N* = 3728	*P*-value
Age, years, median [IQR]	40 [34–47]	52 [45–58]	61 [54–68]	< 0.001	50 [45–56]	56 [50–61]	60 [55–64]	< 0.001
Sex (male), n (%)	9563 (39.7%)	17,212 (69.4%)	2841 (80.6%)	< 0.001	1456 (25.7%)	10,190 (48.0%)	2827 (75.8%)	< 0.001
DM, n (%)	242 (1.0%)	1411 (5.7%)	520 (14.8%)	< 0.001	75 (1.3%)	476 (2.2%)	220 (5.9%)	< 0.001
HTN, n (%)	771 (3.2%)	4366 (17.6%)	1450 (41.1%)	< 0.001	244 (4.3%)	2575 (12.1%)	970 (26.0%)	< 0.001
Current smoker, n (%)	3428 (14.2%)	5415 (21.8%)	749 (21.2%)	< 0.001	324 (5.7%)	1330 (6.3%)	269 (7.2%)	0.013
SBP, mmHg, median [IQR]	114 [105–123]	123 [114–132]	129 [120–138]	< 0.001	126 [116–137]	134 [123–146]	141 [131–154]	< 0.001
Total cholesterol, mg/dL, median [IQR]	188 [167–210]	196 [174–220]	188 [166–213]	< 0.001	215 [190–243]	222 [196–250]	218 [190–248]	< 0.001
HDL cholesterol, mg/dL, median [IQR]	56 [47–66]	49 [42–58]	46 [40–54]	< 0.001	61 [51–71]	55 [47–66]	50 [43–60]	< 0.001
PCE score, median [IQR]	0.6 [0.2–1.6]	3.9 [1.8–7.5]	10.8 [6.3–18.0]	< 0.001	1.7 [0.8–3.8]	4.6 [2.2–9.1]	10.7 [5.9–15.9]	< 0.001
MACE IR per 1000 PY	0.8	4.7	12.1	< 0.001	2.2	5.6	15.5	< 0.001

SH: Severance Hospital; UKB: United Kingdom Biobank; ECG: electrocardiogram; CAC: coronary artery calcium; ECG-CAC model: ECG-based deep learning model for CAC prediction; IQR: interquartile range; DM: diabetes mellitus; HTN: hypertension; SBP: systolic blood pressure; HDL: high-density lipoprotein; PCE: Pooled Cohort Equations; MACE: major adverse cardiovascular event; IR: incidence rate; PY: person-years

per 1000 PY, exceeding the 5.1 in those with PCE moderate-risk but ECG-CAC model low-risk, although the difference did not reach statistical significance ($P = 0.183$). Kaplan-Meier curves showed significant differences based on ECG-CAC model risk categories within both PCE low-risk and moderate-risk groups (Fig. 4C, within PCE low-risk group, log-rank test $P < 0.001$; Fig. 4D, within PCE moderate-risk group, log-rank test $P < 0.001$). The NRI for risk reclassification based on ECG-CAC model within PCE low- or moderate-risk categories was 2.1% (95% CI: 0.0% – 4.3%) (Additional file 1: Table S4). A 10% absolute increase in the ECG-CAC model score was associated with a hazard ratio for MACE of 1.101 (95% CI: 1.050–1.154) after adjusting for clinical variables included in the PCE (Fig. 5, Additional file 1: Table S5). The combined PCE plus ECG-CAC model score, which yielded a C-index of 0.734 (95% CI: 0.715–0.753), surpassed the PCE by a difference of 0.010 (95% CI: 0.003–0.017) in the C-index (Additional file 1: Table S6).

Figure 5 shows the hazard ratios of the ECG-CAC model by subgroup. The association of the ECG-CAC model with MACE remained consistent across all demographic and PCE-based subgroups, with all P-values for hazard ratio comparisons between subgroups being statistically insignificant.

Model explainability

Additional file 1: Figures S8 and S9 present IG–based saliency visualizations comparing groups (CAC score = 0 vs. > 0 and ≥ 400) The mean ECG waveforms appeared nearly identical between groups. However, the IG-based saliency maps consistently highlighted the QRS complex—particularly around the R-peak and extending into the S-wave region in V1 and V2—as the most influential segment across leads. Lead-wise attribution analysis demonstrated that limb leads (II and I) contributed most to the model's predictions, followed by precordial leads V1 and V2. The overall attribution trends were largely consistent between the CAC > 0 and CAC ≥ 400 analyses. This study followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Artificial Intelligence (TRIPOD-AI) reporting guidelines (Additional file 1: Table S7) [20].

Discussion

In this study, we developed an ECG-CAC model to predict CAC and validated its performance in health checkup settings across multiple institutions. Among PCE low-risk individuals, CAC prevalence increased in those further classified as ECG-CAC model high-risk, and among PCE moderate-risk individuals, CAC absence increased in those further classified as ECG-CAC model low-risk. MACE IR was higher in PCE low-risk individuals further classified as ECG-CAC model high-risk than in PCE moderate-risk individuals further classified as ECG-CAC model low-risk. The ECG-CAC model score was an independent risk factor for MACE and provided additional predictive value beyond the PCE.

We demonstrated that our ECG-CAC model predicting CAC could potentially be integrated into clinical workflows in two ways. First, it can guide targeted CAC screening by identifying individuals who would benefit most from it. In the PCE low-risk category, where CAC detection could initiate primary prevention according to guidelines [4, 5], the ECG-CAC model identified a small subset (2.6%–6.1% of individuals in the PCE low-risk

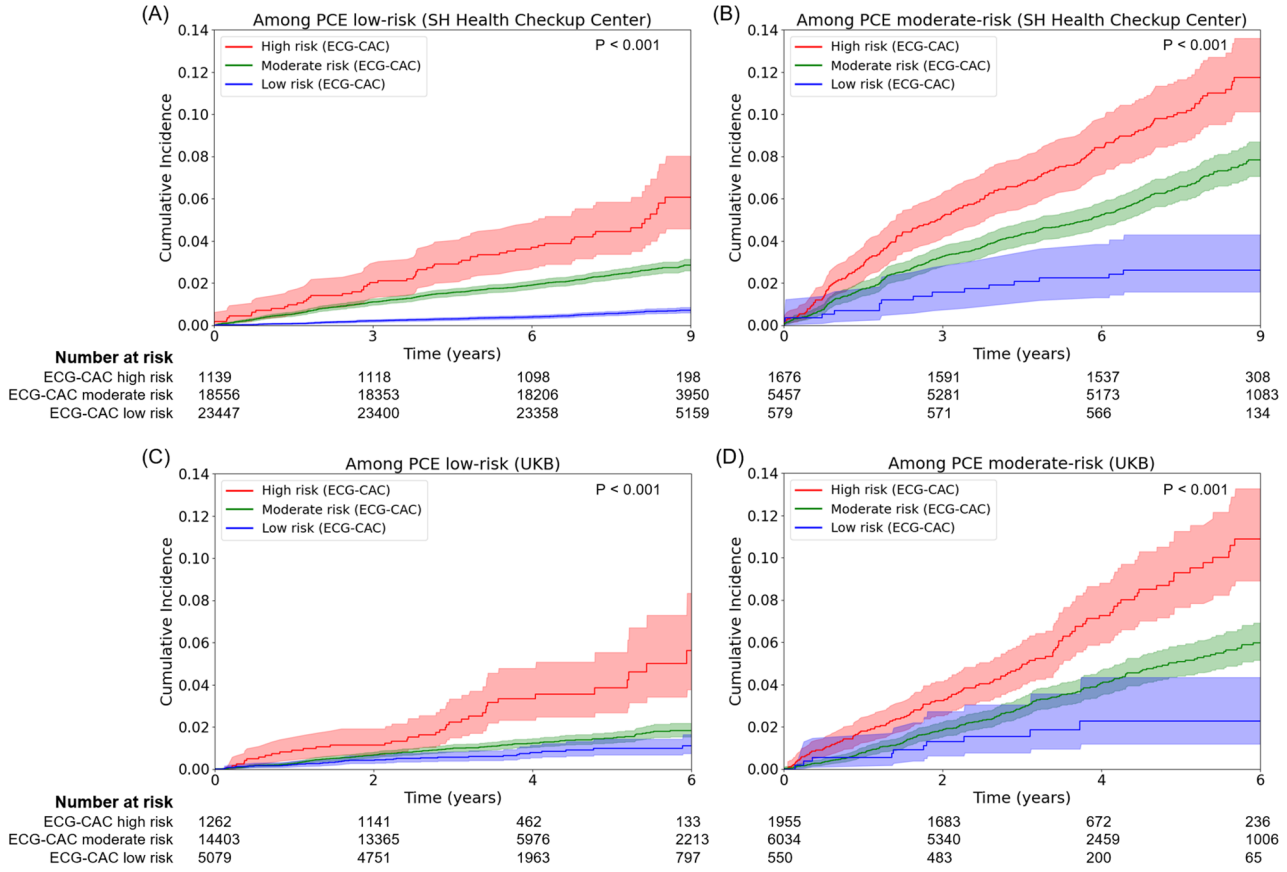

Fig. 4 Kaplan-Meier curves for MACE. Kaplan-Meier curves stratified by ECG-CAC model risk categories within the PCE low- and moderate-risk groups for the SH Health Checkup Center and UKB cohort analysis datasets: (**A**) PCE low-risk (SH Health Checkup Center), log-rank test $P < 0.001$; post-hoc pairwise comparisons: Low vs. Moderate $P < 0.001$, Moderate vs. High $P < 0.001$. (**B**) PCE moderate-risk (SH Health Checkup Center), log-rank test $P < 0.001$; post-hoc pairwise comparisons: Low vs. Moderate $P < 0.001$, Moderate vs. High $P < 0.001$. (**C**) PCE low-risk (UKB), log-rank test $P < 0.001$; post-hoc pairwise comparisons: Low vs. Moderate $P = 0.006$, Moderate vs. High $P < 0.001$. (**D**) PCE moderate-risk (UKB), log-rank test $P < 0.001$; post-hoc pairwise comparisons: Low vs. Moderate $P = 0.030$, Moderate vs. High $P < 0.001$. Bonferroni-adjusted significance level: 0.025. MACE: major adverse cardiovascular event; ECG: electrocardiogram; CAC: coronary artery calcium; ECG-CAC model: ECG-based deep learning model for CAC prediction; PCE: Pooled Cohort Equations; SH: Severance Hospital; UKB: United Kingdom Biobank

category across our study datasets) of individuals as high risk, who showed significantly higher prevalence of CAC. Similarly, in the PCE moderate-risk category, where ruling out CAC could prevent unnecessary interventions according to guidelines [4, 5], the model identified another small subset (5.7%–7.5% of individuals in the PCE moderate-risk category) as low risk, who showed significantly higher likelihood of being CAC-free. This suggests that while routine CAC measurement is impractical due to cost, radiation exposure, and limited accessibility, ECG's high availability and cost-effectiveness make our ECG-CAC model a valuable tool for identifying priority candidates for CAC scoring.

Second, we demonstrated that the ECG-CAC model has the potential to help inform decisions regarding statin therapy. In both cohorts, individuals classified as moderate-risk by PCE but further classified as low-risk by the ECG-CAC model had lower MACE IR than those classified as low-risk by PCE but further classified as high-risk by the ECG-CAC model. These findings suggest that for individuals with discordant PCE and ECG-CAC risk classifications, further evaluation to refine the decision on statin therapy may be warranted. For instance, the model identifies a subset of PCE low-risk individuals with a MACE risk comparable to a higher PCE risk group, and vice-versa. The positive NRI values in both cohorts further validate this risk reclassification strategy.

We demonstrated that the ECG-CAC model score serves as an independent risk factor for MACE and provides incremental predictive value when combined with conventional risk stratification tools, such as the PCE. These findings suggest the potential to incorporate the ECG-CAC model score as a novel risk factor in CVD risk prediction tools. The association between the ECG-CAC model score and MACE remained consistent across all demographic and PCE-based subgroups, showing the ECG-CAC model's effectiveness across diverse populations and cardiovascular risk levels.

SH Health Checkup Center cohort analysis dataset			
	Adjusted HR (95% CI)	**P-value**	**P-value for HR comparison**
Entire cohort	1.128 (1.093 – 1.164)	<0.001	
Subgroup analyses			
Age			0.898
< 60 years	1.117 (1.071 – 1.166)	<0.001	
≥ 60 years	1.122 (1.072 – 1.175)	<0.001	
Sex			0.810
Female	1.135 (1.066 – 1.209)	0.010	
Male	1.125 (1.085 – 1.167)	<0.001	
PCE			0.750
Low risk	1.113 (1.063 – 1.165)	<0.001	
≥ Moderate risk	1.124 (1.078 – 1.173)	<0.001	
UKB cohort analysis dataset			
	Adjusted HR (95% CI)	**P-value**	**P-value for HR comparison**
Entire cohort	1.101 (1.050 – 1.154)	<0.001	
Subgroup analyses			
Age			0.624
< 60 years	1.083 (1.010 – 1.162)	0.025	
≥ 60 years	1.109 (1.041 – 1.182)	0.001	
Sex			0.659
Female	1.117 (1.025 – 1.219)	0.012	
Male	1.092 (1.032 – 1.154)	0.002	
PCE			0.344
Low risk	1.096 (1.014 – 1.185)	0.021	
≥ Moderate risk	1.149 (1.084 – 1.218)	<0.001	

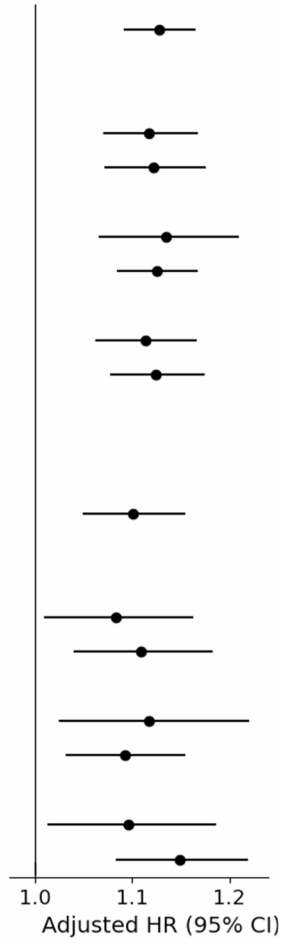

Fig. 5 Adjusted hazard ratios per 10% absolute increase in ECG-CAC model score. This figure presents the adjusted HRs, derived from Cox regression analysis for MACE, for the entire cohort and within each subgroup. The HRs were adjusted for variables included in the PCE: age, sex, diabetes mellitus, hypertension, smoking status, total cholesterol, HDL cholesterol, and systolic blood pressure. We adjusted the scale of the ECG-CAC model score from 0–1 to 0–10 by multiplying by ten for more intuitive data interpretation. The adjusted HRs now show hazard changes associated with a 10% absolute increase in ECG-CAC model score. ECG: electrocardiogram; CAC: coronary artery calcium; ECG-CAC model: ECG-based deep learning model for CAC prediction; HR: hazard ratio; MACE: major adverse cardiovascular event; PCE: Pooled Cohort Equations; HDL: high-density lipoprotein; SH: Severance Hospital; UKB: United Kingdom Biobank

ECG-based deep learning analysis for CAC prediction is most clinically relevant and generalizable when validated using health checkup data or population-representative datasets like UKB, as demonstrated in our study. Health checkups primarily target asymptomatic individuals, and since ECG is routinely measured during these checkups, they provide an ideal setting to enable opportunistic ECG-based deep learning-enabled CAC screening in the general population and optimize primary prevention strategies. In contrast, validating with outpatient data risks introducing selection bias, as diagnostic tests in outpatient settings are typically performed based on specific clinical indications. Furthermore, according to guidelines [4, 5], CAC scoring is not recommended for individuals with clinical ASCVD, who are already candidates for advanced treatments and evaluations, making ECG-based deep learning analysis for CAC prediction irrelevant for this group. To ensure clinical relevance, we excluded individuals with prior cardiovascular events from our retrospective cohort analyses.

Recent large-cohort investigations have shown that the presence of CAC, even at a subclinical stage, is independently associated with adverse cardiac remodeling, including increased left ventricular (LV) mass, larger LV volumes, and impaired diastolic function [21, 22]. Such structural changes may induce subtle conduction delays or heterogeneity in ventricular depolarization, which could manifest as minute ECG alterations

that would previously escape human detection but can now be identified through AI-enabled approaches [23]. Our ECG-CAC model primarily focused on the QRS complex—particularly the R-peak and extending into the S-wave region in V1 and V2—suggesting that CAC-related remodeling may influence both the amplitude and terminal components of ventricular activation and produce minute but AI-detectable electrophysiologic signatures. The positive correlation between CAC severity and ECG-CAC model scores in our study likely reflects progressive, AI-detectable electrophysiologic signatures of structural remodeling accompanying coronary calcification.

Our study has several strengths. First, our ECG-CAC model operates solely on ECG data without requiring additional clinical information, making it easily applicable. It includes all ECGs except those with lead misplacements, artifacts, or artificial pacemakers, enhancing broad clinical utility. Second, the model has shown consistent discriminative performance across multiple external validation datasets, demonstrating its robustness. Third, despite concerns about AI models' cross-ethnic applicability [24], our model, developed primarily with Korean data, showed similar trends in results when applied to the United Kingdom dataset, indicating cross-ethnic generalizability. Fourth, we utilized a dataset of nearly 200,000 ECGs for ECG-CAC model development. Such extensive datasets reduce the risk of overfitting, account for a broader range of variability, and improve the models' generalizability in clinical settings [25].

This study's findings should be interpreted in light of the following limitations. First, the retrospective design introduces limitations, including potential selection bias based on who decided or was recommended to undergo CAC scoring during health checkups, which cannot be precisely known. Future prospective studies are necessary to confirm the findings of the study. Second, while the ACC/AHA PCE is a widely used CVD risk prediction model, its application to Korean and United Kingdom cohorts may introduce inaccuracies. Studies have shown that the PCE, developed using American cohorts, tends to overestimate risk in Asian and European populations [26–29], which points to potential challenges in directly interpreting ECG-CAC model-based risk classification following the PCE in these populations. Third, in the SH data, cardiovascular event occurrences extracted from EMR may be incompletely captured, as events diagnosed at other hospitals could be missed. Fourth, in the UKB cohort, the time gap between initial clinical data collection and subsequent ECG recording may have influenced results. Fifth, vessel- or segment-level analyses were not feasible because the radiology reports used in this study did not include vessel-specific CAC scores. Future studies with standardized vessel-specific annotations may allow more granular assessment of ECG–coronary vessel relationships. Sixth, although we proposed a potential mechanism whereby CAC may lead to structural remodeling and subsequent subtle ECG alterations, this interpretation remains hypothesis-generating. The saliency findings provide indirect rather than causal evidence, and further multimodal studies are warranted to confirm the underlying electrophysiologic mechanisms. Finally, although the ECG-CAC model demonstrated consistent discriminative performance across multiple external validation datasets, calibration performance varied across institutions. Calibration was suboptimal in the AUMC external validation dataset, where predicted probabilities tended to exceed observed CAC prevalence, likely reflecting the younger, healthier population with lower baseline CAC prevalence, in which class imbalance and a narrower disease spectrum can adversely affect calibration and AUPRC. Residual institutional heterogeneity, including differences in ECG acquisition hardware, signal characteristics, and CAC measurement and reporting practices—which are difficult to fully characterize or disentangle in retrospective multicenter data—may have further contributed to miscalibration despite identical preprocessing. These findings highlight the need for site-specific recalibration or local threshold adjustment prior to clinical deployment.

Conclusions

In conclusion, our ECG-CAC model not only serves as an indicator for CAC but also demonstrates potential as a biomarker representing coronary-cerebrovascular atherosclerotic burden. It has potential to impact clinical workflows by guiding targeted CAC screening, optimizing primary prevention strategies, or being incorporated into cardiovascular disease risk prediction tools. It could have particularly important implications in health checkups, providing enhanced risk stratification to the general population.

Abbreviations

ACC/AHA	American college of cardiology/ american heart association
AI	artificial intelligence
AUMC	Ajou university medical center
AUROC	Area under the receiver operating characteristic curve
AUPRC	Area under the precision–recall curve
CAC	Coronary artery calcium
CI	Confidence interval
CT	Computed tomography
CVD	Cardiovascular disease
ECE	Expected calibration error
ECG	Electrocardiogram
ECG-CAC	ECG-based deep learning model for CAC prediction
EMR	Electronic medical record
HDL	High-density lipoprotein
ICD-10	International classification of diseases, 10th edition
IG	Integrated gradients
IR	Incidence rate
IQR	Interquartile range
LV	Left ventricular

MACE	Major adverse cardiovascular event
NPV	Negative predictive value)
NRI	Net reclassification improvement)
PCE	Pooled cohort equations
PPV	Positive predictive value
PY	Person-years
SH	Severance hospital
UKB	United Kingdom Biobank
YSH	Yongin severance hospital

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-026-03400-9.

Supplementary Material 1

Acknowledgements
This study was supported by Division of Digital Health of Yonsei University Health System, Seoul, Korea, with regard to data curation.

Author contributions
CH conceived and designed the overall study. CH, CP and HSL extracted the initial data. HNP and OSK conducted the data linkage with KOSTAT mortality data. CH analyzed the data, implemented the codes, and performed statistical analyses. CP conducted the external validation. CH interpreted the results and drafted the manuscript. SCY, JSK, DY, HCL, JYP, HSL, HNP and SK critically revised the manuscript and improved the study conception and design. DY and JSK supervised this study. CH, DY, JSK and HNP obtained funding. All authors revised and approved the final version of the manuscript.

Funding
This research was supported by a grant of the MD-Phd/Medical Scientist Training Program through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (RS-2023-00276320). This work was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: 1711174471, RS-2022-00141473). The funders had no role in the study design, data collection, analysis, interpretation, manuscript preparation, or decision to submit for publication.

Data availability
Data and other materials from Severance Hospital, Yongin Severance Hospital and Ajou University Medical Center contain potentially identifying or sensitive patient information and are therefore only accessible to researchers who fulfill the confidentiality requirements stipulated by the respective Institutional Review Boards. United Kingdom Biobank data are available to researchers at https://www.ukbiobank.ac.uk/use-our-data/, and our study was conducted under application number 85037. Our code is available at https://github.com/CMI-Laboratory/AI_ECG_CACS/. For further details, please contact the corresponding authors.

Declarations

Ethics approval and consent to participate
This study was conducted in accordance with the Declaration of Helsinki. The Institutional Review Boards (IRB) of SH, YSH and AUMC approved this study and waived the requirement for informed consent because only anonymized data were used retrospectively (IRB no. 4-2022-1299 and 4-2022-1506 [SH], 9-2024-0032 [YSH], AJOUIRB-DB-2024-207 [AUMC]). The UKB obtained Research Tissue Bank (RTB) approval (21/NW/0157) from the North West Multicentre Research Ethics Committee, which covers ethical clearance, and all participants provided informed consent.

Consent for publication
Not applicable.

Competing interests
JSK reports being a chief executive officer of CPEC inc; and grants from Samjin, Yuhan, Daiichi Sankyo, Biosensors, DIO medical, Qualitech Korea has received consultant fees from Abbott Vascular, Philips and Genoss. SCY reports being a chief executive officer of PHI Digital Healthcare; and grants from Daiichi Sankyo. He is a coinventor of granted Korea Patent DP-2023-1223 and DP-2023-0920, and pending Patent Applications DP-2024-0909, DP-2024-0908, DP-2022-1658, DP-2022-1478, and DP-2022-1365, unrelated to current work. All other authors declare no competing interests.

Author details
[1]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea
[2]Institute for Innovation in Digital Healthcare, Yonsei University, Seoul, Republic of Korea
[3]Department of Anesthesiology and Pain Medicine, Seoul National University College of Medicine, Seoul National University Hospital, Seoul, Republic of Korea
[4]Department of Psychiatry, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Korea
[5]Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Seoul, Korea
[6]Center for Digital Health, Yongin Severance Hospital, Yonsei University Health System, Yongin, Republic of Korea
[7]Department of Cardiology, Ajou University School of Medicine, Suwon, Republic of Korea
[8]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Republic of Korea
[9]Yonsei University College of Medicine, Yonsei University Health System, Seoul, Republic of Korea
[10]Division of Cardiology, Department of Internal Medicine, Yonsei University College of Medicine, 50-1, Yonsei-ro, Seodaemun-gu, Seoul 16995, Republic of Korea
[11]Dongbaekjukjeon-daero, Giheung-gu, Yongin-si, Gyeonggi-do 16995, Republic of Korea

Received: 25 August 2025 / Accepted: 13 February 2026
Published online: 21 February 2026

References

1. Greenland P, Lloyd-Jones DM. Role of coronary artery calcium testing for risk assessment in primary prevention of atherosclerotic cardiovascular disease: a review. JAMA Cardiol. 2022;7:219–24.
2. Onnis C, Virmani R, Kawai K, Nardi V, Lerman A, Cademartiri F, et al. Coronary artery calcification: current concepts and clinical implications. Circulation. 2024;149:251–66.
3. Detrano R, Guerci AD, Carr JJ, Bild DE, Burke G, Folsom AR, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N Engl J Med. 2008;358:1336–45.
4. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2019;139(25):e1082-143.
5. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: A report of the American college of Cardiology/American heart association task force on clinical practice guidelines. Circulation. 2019;140:e596–646.
6. Siontis KC, Noseworthy PA, Attia ZI, Friedman PA. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. Nat Rev Cardiol. 2021;18:465–78.
7. Attia ZI, Harmon DM, Behr ER, Friedman PA. Application of artificial intelligence to the electrocardiogram. Eur Heart J. 2021;42:4717–30.
8. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet. 2019;394:861–7.
9. Han C, Kang K-W, Kim TY, Uhm J-S, Park J-W, Jung IH, et al. Artificial Intelligence-Enabled ECG algorithm for the prediction of coronary artery calcification. Front Cardiovasc Med. 2022;9:849223.

10. McEvoy JW, Blaha MJ, Defilippis AP, Budoff MJ, Nasir K, Blumenthal RS, et al. Coronary artery calcium progression: an important clinical measurement? A review of published reports. J Am Coll Cardiol. 2010;56:1613–22.
11. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12:e1001779.
12. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. arXiv [cs.LG]. 2019. https://doi.org/10.48550/ARXIV.1905.11946
13. Einthoven W. Weiteres über Das elektrokardiogramm. Pflugers Arch. 1908;122:517–84.
14. Goldberger E. A simple, indifferent, electrocardiographic electrode of zero potential and a technique of obtaining augmented, unipolar, extremity leads. Am Heart J. 1942;23:483–92.
15. Hecht HS. Coronary artery calcium scanning: past, present, and future. JACC Cardiovasc Imaging. 2015;8:579–96.
16. Bosco E, Hsueh L, McConeghy KW, Gravenstein S, Saade E. Major adverse cardiovascular event definitions used in observational analysis of administrative databases: a systematic review. BMC Med Res Methodol. 2021;21:241.
17. Sundararajan M, Taly A, Yan Q. Axiomatic attribution for deep networks. arXiv [cs.LG]. 2017.
18. Makowski D, Pham T, Lau ZJ, Brammer JC, Lespinasse F, Pham H, et al. NeuroKit2: A python toolbox for neurophysiological signal processing. Behav Res Methods. 2021;53:1689–96.
19. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44:837–45.
20. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378.
21. Castro-Diehl C, Song RJ, Mitchell GF, McManus D, Cheng S, Vasan RS, et al. Association of subclinical atherosclerosis with echocardiographic indices of cardiac remodeling: the Framingham study. PLoS ONE. 2020;15:e0233321.
22. Yared GS, Moreira HT, Ambale-Venkatesh B, Vasconcellos HD, Nwabuo CC, Ostovaneh MR, et al. Coronary artery calcium from early adulthood to middle age and left ventricular structure and function. Circ Cardiovasc Imaging. 2019;12:e009228.
23. Yoon D, Jang J-H, Choi BJ, Kim TY, Han CH. Discovering hidden information in biosignals from patients using artificial intelligence. Korean J Anesthesiol. 2020;73:275–84.
24. Tat E, Bhatt DL, Rabbat MG. Addressing bias: artificial intelligence in cardiovascular medicine. Lancet Digit Health. 2020;2:e635–6.
25. Ying X. An overview of overfitting and its solutions. J Phys Conf Ser. 2019;1168:022022.
26. Cook NR, Ridker PM. Calibration of the pooled cohort equations for atherosclerotic cardiovascular disease: an update. Ann Intern Med. 2016;165:786–94.
27. Damen JA, Pajouheshnia R, Heus P, Moons KGM, Reitsma JB, Scholten RJPM, et al. Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: a systematic review and meta-analysis. BMC Med. 2019;17:109.
28. Jung KJ, Jang Y, Oh DJ, Oh B-H, Lee SH, Park S-W, et al. The ACC/AHA 2013 pooled cohort equations compared to a Korean risk prediction model for atherosclerotic cardiovascular disease. Atherosclerosis. 2015;242:367–75.
29. Cho YG. Cardiovascular risk prediction in Korean adults. Korean J Fam Med. 2018;39:135–6.

Publisher's note