Original Article

Incremental Predictive Value of ECG Burden Added to Clinical Risk for Appropriate ICD Therapy Selection: Model Development and Internal Validation

Muneera AlTaweel [1,3*], Maysan Almegbel [2,3], Abdulmohsen Almusaad [2,3]

[1] Cardiac Science, Department of Medicine, King Abdulaziz Hospital, Ministry of National Guard-Health Affairs, Al-Ahsa, Saudi Arabia.
[2] King Abdulaziz Cardiac Center, King Abdulaziz Medical City, Ministry of National Guard-Health Affairs, Riyadh, Saudi Arabia.
[3] King Abdullah International Medical Research Center, Ministry of National Guard-Health Affairs, Saudi Arabia.

Citation: AlTaweel M, Almegbel M, Almusaad A. Incremental Predictive Value of ECG Burden Added to Clinical Risk for Appropriate ICD Therapy Selection: Model Development and Internal Validation. Res Heart Yield Transl Med 2026; 21(3):197-202.
https://doi.org/10.18502/rhythm.v21i3.22562

Highlight

- A prespecified 6-component ECG burden score did not provide incremental predictive value beyond minimal clinical predictors for appropriate ICD therapy selection in internal validation.

Article info:
Received: 25 Feb. 2026
Revised: 10 May. 2026
Accepted: 25 May. 2026

*** Corresponding Author:**
Muneera AlTaweel
King Abdulaziz Hospital (KAH), MNGHA, Al-Ahsa, Saudi Arabia.
Email: taweelm@mngha.med.sa

ABSTRACT

Background: Risk stratification for selecting appropriate implantable cardioverter-defibrillator (ICD) therapy remains imperfect, and pragmatic electrocardiogram (ECG)-derived markers may provide incremental information beyond routine clinical variables.

Objective: We sought to evaluate whether a prespecified 6-component ECG burden score improves the prediction of appropriate ICD therapy selection beyond a minimal clinical model.

Methods: Retrospective prediction study using deidentified data from a parent ICD cohort at King Abdulaziz Medical City, Riyadh (n=236; 35 events; mean [SD] follow-up, 7.3 [2.3] years), analyzed under King Abdulaziz Hospital/KAIMRC governance (Protocol No. NRA 26/004/1). Minimal clinical predictors (age, sex, left ventricular ejection fraction, cardiomyopathy type, and ICD indication) were modeled with ridge logistic regression (clinical only) and with the addition of ECG burden (combined). Internal validation used 5-fold cross-validation with out-of-fold predictions.

Results: In complete-case modeling (n=231), out-of-fold discrimination was 0.502 for clinical only and 0.452 for combined (ΔAUC, −0.050; 95% CI, −0.097 to −0.008). In the imputed sensitivity analysis (n=236), discrimination was 0.514 vs 0.485 (ΔAUC, −0.028; 95% CI, −0.066 to 0.008). Decision curve analysis showed minimal separation between the models.

Conclusion: In internal validation, adding the ECG burden score did not improve discrimination beyond minimal clinical predictors and provided only minimal incremental clinical utility in predicting appropriate ICD therapy selection.

Keywords: Implantable Cardioverter-Defibrillator; Electrocardiogram; Risk Prediction; Logistic Regression; Cross-Validation; Decision Curve Analysis

RHYTHM
Research in Heart Yield and Translational Medicine

Introduction

Implantable cardioverter-defibrillators are a cornerstone therapy for preventing sudden cardiac death in selected patients at risk of life-threatening ventricular arrhythmias, with landmark randomized trials demonstrating survival benefits in both ischemic and nonischemic cardiomyopathy populations.[3,4] Contemporary guidelines recommend implantable cardioverter-defibrillator (ICD) implantation for primary and secondary prevention based on clinical characteristics and left ventricular systolic function, most commonly left ventricular ejection fraction (LVEF).[1,2] Nonetheless, relying on LVEF alone provides an incomplete assessment of arrhythmic risk. In large ICD registries, approximately 60% to 70% of ICD-eligible patients never receive appropriate therapy during long-term follow-up, while others experience recurrent ventricular arrhythmias despite guideline-based implantation.[1,2] This gap motivates pragmatic approaches to refine patient selection and risk stratification in routine practice.

Electrocardiographic (ECG) markers-including QRS duration, QT interval, QRS-T angle, and Tpeak-Tend interval-have been associated with arrhythmic risk in observational studies, reflecting underlying myocardial fibrosis, conduction disease, and repolarization heterogeneity.[1,2] These ECG features are, nevertheless, typically studied in isolation, and it remains unclear whether a composite ECG burden score derived from multiple binary abnormality flags provides incremental predictive value in a real-world ICD cohort.

The present study is a secondary predictive analysis of a parent cohort of ICD recipients treated at King Abdulaziz Cardiac Center, King Abdulaziz Medical City, MNGHA, Riyadh. This study examined whether a prespecified ECG burden score, derived from 6 binary ECG abnormality flags, could provide incremental predictive value beyond a minimal clinical predictor set for identifying patients who could subsequently receive appropriate ICD therapy. Model performance was evaluated using discrimination, calibration, and clinical utility.[7–9]

Methods

This retrospective observational prediction study used deidentified data from a parent cohort of ICD recipients treated at King Abdulaziz Cardiac Center, King Abdulaziz Medical City, MNGHA, Riyadh, Saudi Arabia. The analysis was conducted under the governance of King Abdulaziz Hospital (KAH), MNGHA, Al-Ahsa, and is reported in accordance with the TRIPOD+AI recommendations.[5] Risk of bias was assessed using PROBAST+AI.[6]

Ethical approval was obtained from the Institutional Review Board of KAIMRC, MNGHA, Saudi Arabia (Protocol No. NRA 26/004/1; approved). The study drew upon retrospective, deidentified clinical data, and informed consent was waived in keeping with the IRB policy.

The analytic cohort comprised 236 ICD recipients with available outcome data. The mean (SD) follow-up was 7.3 (2.3) years (range, 4-16 years), measured from the date of baseline ECG extraction to the last available device interrogation or clinical contact. Complete-case data for model development were available for 231 participants.

The primary outcome was appropriate ICD therapy selection, defined as any delivered antitachycardia pacing (ATP) or shock for ventricular tachycardia (VT) or ventricular fibrillation (VF), confirmed by device electrograms and ICD interrogation. Therapies for supraventricular tachyarrhythmias, oversensing, or other nonventricular causes were excluded. The endpoint includes both ATP for slower VT and shocks for faster VT or VF, consistent with standard registry definitions.[1,2]

Predictors were selected a priori for clinical relevance and routine availability. The minimal clinical core included age, sex, LVEF, cardiomyopathy type, and ICD indication (primary vs secondary prevention). The ECG predictor was a prespecified ECG burden score, defined as the unweighted sum of 6 binary ECG flags (score range, 0-6). The binary thresholds for each ECG component are defined in (Table 2).

The primary analysis used a complete-case approach (n=231). A sensitivity analysis retained the full cohort (n=236) and used median or mode imputation for predictors; the outcome was not imputed.

Two prespecified logistic regression models were developed: (1) clinical-only and (2) combined (clinical core + ECG burden score). Models were fit using ridge (L2-penalized) logistic regression to reduce overfitting. Ridge regression was selected to retain all prespecified predictors with coefficient shrinkage, consistent with the prespecified analysis plan. A prespecified ElasticNet sensitivity analysis was also performed. (See the Supplementary.) Age and sex were retained following the prespecified analysis plan, irrespective of individual statistical significance, in accordance with TRIPOD reporting standards.[5]

Performance was assessed using 5-fold cross-validation with out-of-fold (OOF) predictions. Discrimination was quantified by area under the curve (AUC); incremental value by Δ*AUC* (combined minus clinical-only), with 95% bootstrap confidence intervals (CIs).[8,10] Calibration was evaluated both graphically and quantitatively using the calibration intercept and slope and the Brier score.[9] Clinical utility was assessed by decision curve analysis (DCA) across clinically relevant threshold probabilities.[7] Primary figures for Discrimination, Calibration, and Clinical utility are presented in (Figures 1 to 3) appear in the Results section.

Results

A total of 236 participants were included; 35 (14.8%) received appropriate ICD therapy selection during a mean (SD) follow-up of 7.3 (2.3) years. The primary complete-case modeling dataset included 231 participants (35 events; 15.2%). Baseline characteristics by outcome status are shown in (Table 1). ECG binary threshold definitions and component prevalence are summarized in (Table 2).

Internal validation utilizing OOF predictions showed modest discrimination, consistent with a null incremental finding. In the complete-case analysis, OOF AUC was 0.502 for clinical-only and 0.452 for combined (ΔAUC, −0.050; 95% CI, −0.097 to −0.008) (Figure 1). In the imputed sensitivity analysis (n=236), OOF AUC was 0.514 vs 0.485 (ΔAUC, −0.028; 95% CI, −0.066 to 0.008). Age and sex failed to constitute statistical significance in the multivariable model, concordant with their modest individual contributions to predicting appropriate ICD therapy. Both were, therefore, retained per the prespecified analysis plan.

Calibration plots are shown in (Figure 2). The calibration intercept was 0.116, and the apparent calibration slope was 0.20, indicating moderate miscalibration of the predicted probabilities—a finding expected given the low events-per-variable ratio (approximately 3.2) and the absence of external validation. Decision curve analysis showed minimal separation in net benefit between the clinical-only and combined models across clinically relevant thresholds (Figure 3).

Univariable associations for the ECG burden score and its components are summarized in (Figure 4).

Table 1. Baseline characteristics by outcome status

Variable	Entire Cohort (n=236)	With Events (n=35)	Without Events (n=201)
Age, y (mean [SD])	61.7 (15.6)	62.5 (14.3)	61.6 (15.8)
Male sex, n (%)	187 (79.2%)	28 (80.0%)	159 (79.1%)
LVEF ≤ 35%, n (%)	198 (83.9%)	31 (88.6%)	167 (83.1%)
LVEF, % (mean [SD])	33.3 (11.4)	30.8 (10.1)	33.7 (11.6)
Ischemic cardiomyopathy, n (%)	232 (98.3%)	35 (100%)	197 (98.0%)
Primary prevention ICD, n (%)	196 (83.1%)	25 (71.4%)	171 (85.1%)
Diabetes mellitus, n (%)	146 (61.9%)	22 (62.9%)	124 (61.7%)
Hypertension, n (%)	157 (66.5%)	23 (65.7%)	134 (66.7%)
Follow-up, y (mean [SD])	7.3 (2.3)	7.1 (2.1)	7.3 (2.4)
ECG burden score (mean [SD])	2.8 (1.7)	3.0 (1.8)	2.8 (1.6)
QRS prolongation, n (%)	112 (47.9%)	18 (51.4%)	94 (46.8%)

Obtuse QRS-T angle, n (%)	134 (56.8%)	22 (62.9%)	112 (55.7%)
Delayed QRS transition, n (%)	143 (60.6%)	22 (62.9%)	121 (60.2%)
QT prolongation, n (%)	133 (56.6%)	21 (60.0%)	112 (55.7%)
Tpe prolongation, n (%)	59 (25.0%)	9 (25.7%)	50 (24.9%)
Intraventricular conduction delay, n (%)	90 (38.1%)	15 (42.9%)	75 (37.3%)

LVEF: left ventricular ejection fraction; ICD: implantable cardioverter-defibrillator

Table 2. ECG burden score: binary component definitions and cohort prevalence (n = 236)

ECG Component	Binary Threshold (1=abnormal)	Prevalence in the Cohort
QRS duration	≥120 ms	112 (47.9%)
QTc interval	≥450 ms (male) / ≥470 ms (female)	133 (56.6%)
Tpeak-Tend (Tpe)	≥90 ms	59 (25.0%)
QRS-T angle	Obtuse (≥90°)	134 (56.8%)
QRS transition	Delayed (transition zone ≥V5)	143 (60.6%)
Intraventricular conduction delay	IVCD pattern not meeting LBBB/RBBB criteria	90 (38.1%)

QTc: corrected QT interval using the Bazett formula; Tpe: Tpeak-Tend interval; IVCD: intraventricular conduction delay

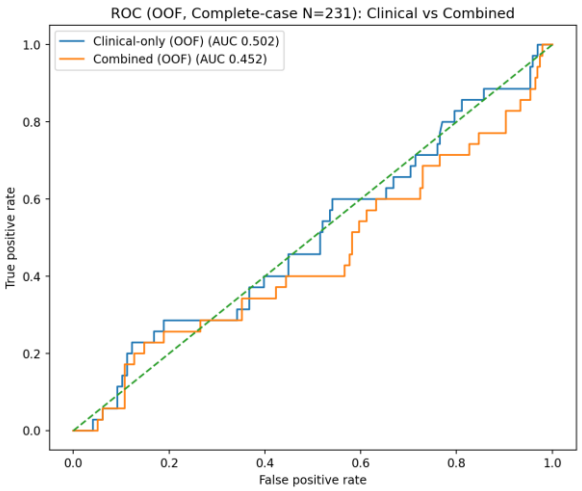


Figure 1. Discrimination (out-of-fold [OOF] receiver operating characteristic): clinical-only vs combined models (complete-case n=231)

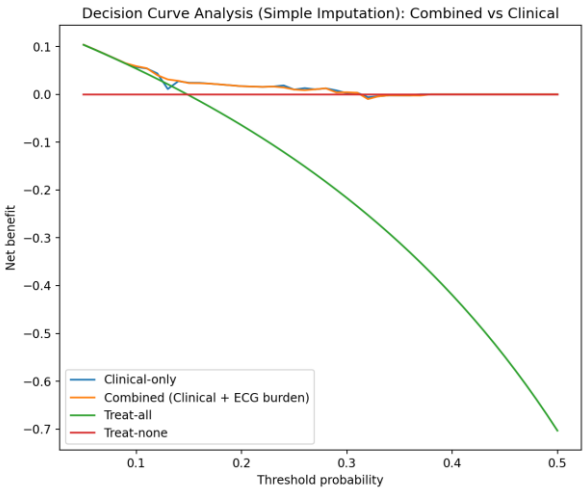


Figure 3. Clinical utility (decision curve analysis): combined vs clinical-only models (imputed n=236)

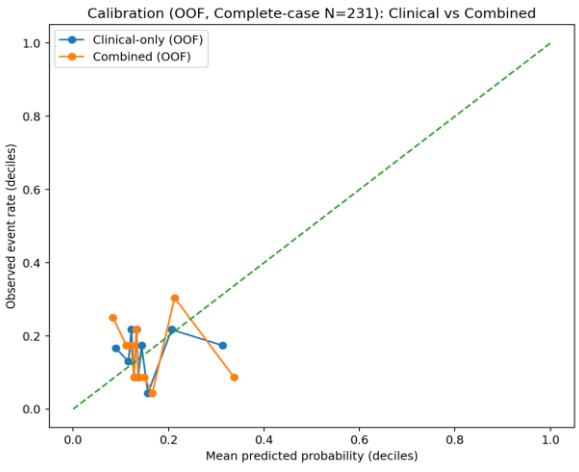
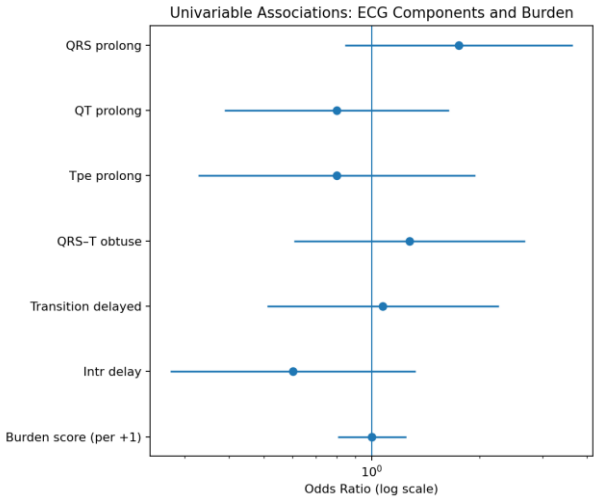


Figure 2. Calibration out-of-fold (OOF): clinical-only vs combined models (complete-case n=231)



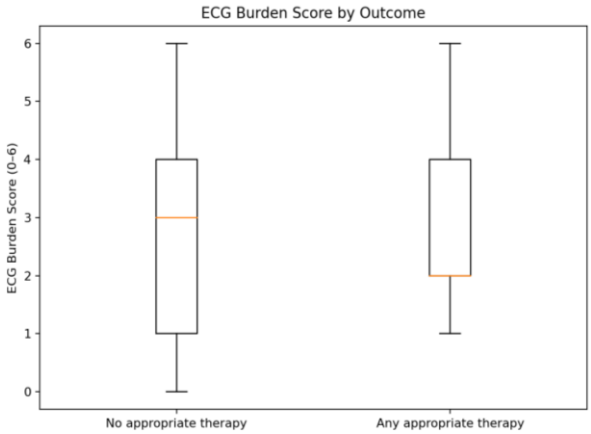


Figure 4. ECG burden associations and distribution: (A) univariable odds ratios; (B) ECG burden score distribution by outcome

Discussion

In this retrospective prediction study of ICD recipients, a prespecified ECG burden score did not confer incremental predictive value beyond a minimal clinical model for selecting appropriate ICD therapy. Across both complete-case and imputed sensitivity analyses, discrimination was low to modest and did not improve with the ECG burden score (Figure 1). Calibration assessment and decision curve analysis showed no compensatory gain from the ECG burden score (Figures 2 and 3).

The overall discrimination of the clinical-only model (AUC, 0.502) was modest, reflecting the inherent difficulty of predicting appropriate ICD therapy from baseline clinical variables alone. This finding is not unique to this dataset. Indeed, predicting arrhythmic outcomes in ICD populations using clinical risk factors is a recognized methodological challenge, as appropriate therapy reflects a complex interplay of substrate, trigger, and programming factors not fully captured by admission-level variables.[1,2]

Several factors may explain the absence of incremental ECG value. First, the composite end point of appropriate ICD therapy includes both ATP events for slower VT and shocks for faster VT or VF. This heterogeneity may dilute prognostic associations because the substrates and predictors for these event types differ. Future studies should consider analyzing ATP and shock endpoints separately to improve signal detection. Second, the ECG burden score prioritizes simplicity and reproducibility; however, dichotomizing ECG features discards prognostic information from continuous measurements and assumes equal weighting of each component without clinical or statistical justification. Future studies should evaluate continuous ECG representations and data-driven weighting, which may recover prognostic signal lost in the binary construction. Third, arrhythmic outcomes and ICD therapies are influenced by device programming, follow-up intensity, and competing clinical risks, which may attenuate baseline ECG associations.

With respect to the choice of ridge regression, this study selected ridge (L2) penalization to retain all prespecified predictors with coefficient shrinkage, consistent with the analysis plan. We acknowledge that Elastic Net penalization-combining L1 and L2 penalties--may be preferable in the low EPV setting because it performs implicit feature selection. A prespecified Elastic Net sensitivity analysis yielded consistent results (see the Supplement), supporting the robustness of the primary null finding.

The findings do not negate the clinical relevance of ECG abnormalities or the established role of ICD therapy in appropriate populations.[1–4] Instead, the results suggest that the current 6-flag ECG burden construct, when combined with readily available clinical predictors, may be insufficient to improve prediction of appropriate ICD therapy selection in this cohort. Future work should evaluate alternative ECG representations (eg, continuous features, automated ECG-derived metrics, and deep learning), incorporate device-related factors, and apply time-to-event modeling. External validation in independent cohorts is essential before clinical adoption.[5]

The strengths of the current study include the prespecified ECG burden definition, penalized regression, and internal validation using OOF predictions, with complementary assessment of discrimination, calibration, and clinical utility.[5, 7–9] Limitations include the retrospective single-center design; the low events-per-variable ratio (approximately 3.2), which limits model stability; the composite end point heterogeneity discussed above; and the absence of external validation. The low calibration slope (0.20) indicates that predicted probabilities are poorly calibrated, which is expected in this underpowered exploratory model and reinforces the need for external validation before any clinical application.

Conclusion

In internal validation, adding the prespecified ECG burden score failed to improve discrimination beyond minimal clinical predictors and provided only minimal incremental clinical utility in predicting appropriate ICD therapy selection. Future studies should evaluate richer ECG representations and validate the findings in high-volume, independent ICD cohorts before clinical adoption.

Declarations:

Ethical Approval

Institutional Review Board, KAIMRC, MNGHA (Protocol No. NRA 26/004/1). Informed consent waived - retrospective de-identified data.

Funding

According to the authors, this article has no financial support.

Conflict of Interest

The authors report no conflict of interest.

Acknowledgments

The authors acknowledge the clinical and data support provided by the King Abdulaziz Cardiac Center, King Abdulaziz Medical City, Ministry of National Guard–Health Affairs (MNGHA), Riyadh, Saudi Arabia. Research governance and IRB oversight were provided by the King Abdullah International Medical Research Center (KAIMRC), MNGHA, Saudi Arabia.

References

1. Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB, et al. 2017 AHA/ACC/HRS Guideline for Management of Patients with Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death. J Am Coll Cardiol. 2018;72(14):e91-e220.
2. Zeppenfeld K, Tfelt-Hansen J, de Riva M, Winkel BG, Behr ER, Blom NA, et al. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. Eur Heart J. 2022;43(40):3997-4126.
3. Moss AJ, Zareba W, Hall WJ, Klein H, Wilber DJ, Cannom DS, et al. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. N Engl J Med. 2002;346(12):877-83.
4. Bardy GH, Lee KL, Mark DB, Poole JE, Packer DL, Boineau R, et al. Amiodarone or an implantable cardioverter-defibrillator for congestive heart failure. N Engl J Med. 2005;352(3):225-37.
5. 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. BMJ. 2024;385:e078378.
6. Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models. BMJ. 2025;388:e082505.
7. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565-74.
8. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001;54(8):774-81.
9. Brier GW. Verification of forecasts expressed in terms of probability. Mon Weather Rev. 1950;78(1):1-3.
10. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated ROC curves: a nonparametric approach. Biometrics. 1988;44(3):837-45.