Evaluating AI-ECG for Incident Atrial Fibrillation via Double Machine Learning and Nonlinear Risk Models

André R Ekengren*, Jiawei Li, Fabio Bonassi, Antônio H Ribeiro

Uppsala University, Uppsala, Sweden

Background: While AI-ECG derived biomarkers have been shown to improve prediction of atrial fibrillation (AF) when added to traditional risk score models such as CHARGE-AF, existing studies predominantly rely on linear combinations of an AI-ECG output and patient clinical variables in their analyses, potentially masking nonlinear relationships.

Aim: To determine whether an AI-ECG biomarker adds significant independent prognostic value for 5-year incident AF when accounting for complex nonlinear relationships and confounding by patient clinical variables.

Methods: We derived a retrospective cohort from the MIMIC-IV database to predict the outcome of incident AF within 5 years. The logit of a pretrained AI-ECG AF prognosis model and the CHARGE-AF variables were used as features. Several linear and nonlinear models were trained (e.g. Logistic Regression, XGBoost, Random Forest) with and without the AI-ECG biomarker, and performance was compared via AUC-ROC with bootstrapped 95% confidence intervals. Additionally, we applied the Double Machine Learning (DML) framework using an interactive regression model (IRM) to estimate the average predictive effect of a positive AI-ECG prognosis when controlling for nonlinear confounding.

Results: Validation showed improved discrimination when including the AI-ECG biomarker in both Logistic Regression and Random Forest models, with similar improvements observed in all other evaluated models. Further, DML analysis revealed a statistically significant average predictive effect of a positive AI-ECG prognosis, which was underestimated by a naive additive linear probability model.

Conclusion: An AI-ECG biomarker for incident AF provides significant independent prognostic value and is complementary to traditional risk models, even when accounting for nonlinear clinical interactions.

Table 1. Predictive utility and independent prognostic effect of the AI-ECG biomarker.

Model / Methodology	**Metric (95% CI)**
Panel A: AUC-ROC	
Logistic Regression (Clinical Baseline)	0.74 [0.71, 0.76]
Logistic Regression (+ AI-ECG)	0.76 [0.74, 0.79]
Random Forest (Clinical Baseline)	0.75 [0.73, 0.78]
Random Forest (+ AI-ECG)	0.78 [0.75, 0.80]
Panel B: Average Predictive Effect	
Linear Probability Model	0.33 [0.28, 0.38]
Double Machine Learning (IRM)	0.42 [0.35, 0.50]