

12-Lead ECG Detection of Digoxin Exposure: An Initial Step Toward Toxicity Stratification

Background: Digoxin remains clinically relevant in atrial fibrillation and heart failure, but ECG-based toxicity assessment is challenging because disease-related abnormalities may overlap with digoxin-associated changes. As an initial step toward toxicity-oriented stratification, we first tested whether a twelve-lead ECG model could distinguish normal controls from digoxin-exposed patients, and then classified high versus non-high digoxin levels as an initial surrogate of toxicity risk.

Methods: ECGResNet1D models were trained on 12-lead, 10-second ECGs from three tertiary hospitals (A train, B validation/internal, C external). The balanced normal-plus-digoxin cohort included 3,368 training, 1,628 validation, 700 internal-test, and 666 external-test ECGs; digoxin-high cases numbered 67, 39, 8, and 38. Metadata included age, sex, heart rate, PR interval, QRS duration, QT interval, corrected QT, and RR interval. Normal was defined as digoxin value 0 with no history; exposure as digoxin value >0; high level as ≥2.0.

Results: On the external test set, normal-versus-digoxin discrimination was excellent for ECG-only and ECG-plus-metadata models (AUROC 0.982 and 0.985; AUPRC 0.981 and 0.987). Within digoxin-exposed patients, high-versus-non-high classification was evaluated under marked class imbalance (positive prevalence 2.3-11.4%). In this setting, ECG plus metadata outperformed ECG alone on the external set (AUROC 0.709 vs 0.639; AUPRC 0.339 vs 0.250), with sensitivity 0.553 and specificity 0.820. The external prevalence baseline was 0.114, indicating about threefold precision-recall enrichment.

Conclusions: Twelve-lead ECG robustly captures digoxin exposure signatures. Under low-prevalence conditions, ECG plus metadata showed meaningful enrichment for identifying high digoxin levels, supporting next-stage toxicity-oriented modeling.

Table 1. Sample size by split in the balanced normal-plus-digoxin cohort

Split	Site	Total N	Digoxin N	Normal N	Digoxin High (≥2.0)	Digoxin Non-high (0<y<2.0)
Train	A	3,368	1,684	1,684	67	1,617
Validation	B 70%	1,628	814	814	39	775
Internal test	B 30%	700	350	350	8	342
External test	C	666	333	333	38	295

Table 2. Metadata used in the ECG+meta model

Category	Variables
Demographics	Age, Sex
ECG measurements	HR, PR interval, QRS duration, QT interval, corrected QT, RR interval

Table 3. Key external test results

Task	Model	N	AUROC	AUPRC	Sensitivity	Specificity	F1
Normal vs Digoxin	ECG only	666	0.982	0.981	0.934	0.943	0.938
Normal vs Digoxin	ECG+meta	666	0.985	0.987	0.916	0.970	0.941
Digoxin-only High vs Non-high	ECG only	333	0.639	0.250	0.395	0.847	0.306
Digoxin-only High vs Non-high	ECG+meta	333	0.709	0.339	0.553	0.820	0.375