ECG BASED UNSUPERVISED MODEL PREDICTS HEART FAILURE AND MAJOR ADVERSE CARDIO-VASCULAR EVENTS IN THE GENERAL POPULATION

Josseline Madrid*, Stefan van Duijvenboden, Patricia B Munroe, Ana Mincholé, Julia Ramírez,

University of Zaragoza, Zaragoza, Spain

Aims: Early detection of cardiovascular risk in the general population is still a challenge. The electrocardiogram (ECG) is a cost-effective screening test for large populations. We hypothesize that individuals with subclinical cardiovascular disease (CVD) might share similar ECG features that can be used for non-invasive risk stratification. This study aimed to identify distinct clusters of individuals without prevalent CVDs based on their ECG morphology using unsupervised learning, and investigate their association with incident CVD risk.

Methods: A median heartbeat was derived from 15-seconds resting ECGs (leads I, II and the 6 precordial leads) from 61,489 individuals without prevalent CVDs in the UK Biobank study. ECG parameters related to time intervals, amplitudes, and morphology were calculated for each lead's median heartbeat. Using k-means clustering, individuals were classified into k distinct clusters. Survival analysis assessed the association of each cluster with incident atrial fibrillation, heart failure (HF), myocardial infarction, major adverse cardiovascular events (MACE) and ventricular arrhythmias risk (9-years follow-up).

Results: The model distinguished 3 clusters with varying morphological features, which significantly differed in terms of HF and MACE events rate (Table 1). Cluster 2 (N = 2,562) included the highest rate of HF (1.5%, $p<0.001$) and MACE events (2.4%, $p<0.001$). These individuals exhibited lower QRS amplitudes (in leads II, V4 and V6), wider QRS duration, longer QTc and Tpec intervals, greater T wave morphological variations with respect to a normal reference and major variability in ST segment deviation than clusters 1 and 3.

Conclusions: Our analysis has identified in an unsupervised manner a group of individuals at risk of HF and MACE using 15-seconds ECGs enabling fast and noninvasive risk assessment in large populations.

Table 1. Results of 3-means Clustering

	Cluster 1 (N=30,103)	Cluster 2 (N=2,562)	Cluster 3 (N= 28,824)	*P-value*
HF ratio	0.4 %	1.5%	0.3%	<0.001
MACE ratio	1.0 %	2.4%	1.1%	<0.001