AI-PPG for pre-screening of reduced left ventricular ejection fraction: development and prospective external validation

Mikołaj Basza[1], Mateusz Soliński[2], Damian Kucharski[3], Damian Waląg[4], Cezary Maciejewski[1], Weronika Kowalczyk[1], Arkadiusz Czerwiński[3], Agata Wijata[5], Jakub Nalepa[3], Łukasz Kołtowski[1], Marcin Grabowski[1]

[1]1st Chair and Department of Cardiology, Medical University of Warsaw, Warsaw, Poland, [2]Computational Physics Laboratory, Tampere University, Tampere, Finland, Department of [3]Algorithmics and Software, Silesian University of Technology, Gliwice, Poland, Center For Digital Medicine, [4]National Institute of Cardiology, Warsaw, Poland, [5]Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland

Heart failure (HF) affects 55–64 million people worldwide, and asymptomatic pre-HF is estimated to be three to four times more common than clinically overt disease. Early detection of reduced left ventricular ejection fraction (LVEF) is important, yet scalable non-invasive screening tools remain lacking. This study developed and externally validated a machine learning model based on photoplethysmography (PPG) to identify reduced LVEF.

The model was trained on the retrospective MIMIC-III critical care database and externally validated in an independent prospective cohort from the Medical University of Warsaw. Adults with high-quality 10-minute PPG recordings and paired echocardiography were included, while those with acute haemodynamic instability, septic or cardiogenic shock, or recent vasoactive therapy were excluded. The development dataset (split into 80% train, 20% internal test set) included 2,072 paired recordings from 1,834 patients, and external validation included 103 patients. After signal quality assessment and amplitude normalization, 206 morphological and HRV features were extracted. An ensemble model combining logistic regression, random forest, gradient boosting, and support vector machine was developed using cross-validated hyperparameter tuning. Subgroup discrimination was evaluated using AUC.

In the internal test set, the area under the receiver operating characteristic curve (AUC) was 0.722 (95% confidence interval [CI] 0.670–0.778) for LVEF below 40% and 0.684 (95% CI 0.640–0.724) for LVEF below 50%. In the external validation cohort, AUC was 0.710 (95% CI 0.582–0.811) for LVEF below 40% and 0.708 (95% CI 0.598–0.805) for LVEF below 50%. The negative predictive value for LVEF below 40% was 87.7% in the external cohort. Performance was broadly consistent across major clinical subgroups, although discrimination decreased at extreme heart rates.

The PPG-based model showed, moderate and stable discrimination for reduced LVEF across both cohorts, representing an important step toward non-invasive HF pre-screening with wearable sensors.

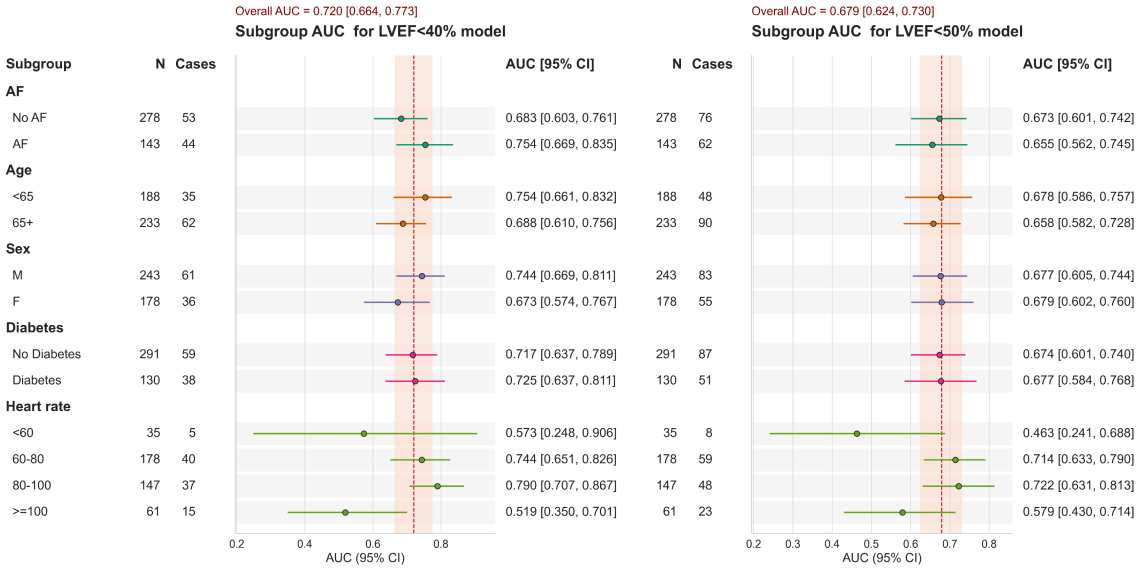

Figure 1. Subgroup AUCs with 95% confidence intervals for the models identifying LVEF <40% and LVEF <50% across atrial fibrillation, age, sex, diabetes, and heart rate categories.