Deep learning based prediction of atrial fibrillation incidence from 1-lead ECGs: a model development and validation study

J. Bremer¹, M. Neyazi¹, M.S. Knorr³, M. Vollmer², S. Gross², J. Brederecke¹, F.M. Ojeda¹, M. Doerr¹, S. Blankenberg¹, R.B. Schnabel¹

¹University Medical Center Hamburg Eppendorf, Department of Cardiology, University Heart & Vascular Center Hamburg, Hamburg, Germany
²University Hospital of Greifswald, Greifswald, Germany

Funding Acknowledgements: Type of funding sources: Public grant(s) – EU funding. Main funding source(s): European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under the grant agreement No 847770

Introduction: Wearables capable of deriving biomarkers such as 1-lead ECGs become largely available, therefore offering the potential to screen for prevalent and incident cardiovascular disease within the population. Screening for prevalent atrial fibrillation (AF) from such devices is already widely applied. It has remained unclear whether predicting incident AF using 1-lead ECGs is possible. The assessment of AF risk before clinical manifestation may permit early detection and thus slow or limit the onset and progression of the disease by taking appropriate measures.

Purpose: Utilizing data from the UK-Biobank (UKB), we examined the possibility to predict three-year AF incidence from single-lead ECGs using deep learning-based methods.

Methods: For this study 100,551 1-lead ECGs from 91,898 unique participants were utilized, including 57,897 4-lead resting ECGs during ergometry from 2006-2010 and 9,826 at the first follow-up (FU) from 2010-2013 and 32,828 12-lead resting ECGs during the FU imaging study from 2014-2020. The 12-lead and 4-lead ECGs share information from lead I, which was used for all further analyses. We employ a 1D-Convolutional Neural Network (CNN) to predict the 3-year risk of incident AF. The CNN is compared with the clinically used CHARGE-AF Score. External validation was performed in the SHIP-START cohort including its three FU examinations (n=2,579; nFU1=3,101; nFU2=2,613; nFU3=1,482) and a median FU time of 5-years to FU and SHIP-TREND (n=2,931) with one FU after 7-years.

Results: The CNN provided a good ability to predict of 3-year incident AF at all three UKB FUs (AUC: 0.729, 0.727, 0.705), while the CHARGE-AF Score provides better results at Exam1 (0.773), worse at Exam2 (0.692) and equal at Exam3 (0.717). A combination of the CHARGE-AF Score and the CNN improved the individual predictions at all three FUs (AUC: 0.811, 0.741, 0.753). The CNN provides an equally good ability to predict incident AF at all four FUs in SHIP-START (AUC: 0.859, 0.785, 0.689, and 0.772) and in SHIP-TREND (0.780). Combined with the CHARGE-AF-Score score, the performance in SHIP-START increases to 0.797, 0.881, 0.765, and 0.812 and in SHIP-Trend to 0.826 respectively.

Conclusion: Using a CNN to predict incident AF from 1-lead ECGs, we provide further evidence for the added value of clinically applied deep learning-based methods. By combining the CNN predictions on ECG data with the well-established CHARGE-AF score, 3-year AF incidence can be predicted with high accuracy.