The exploratory clinical trial of AI-enabled Holter electrocardiogram for diagnosis of paroxysmal atrial fibrillation from sinus rhythm waveforms

Y. Tamura¹, H. Taniguchi¹, J. Iwasawa¹, E. Watanabe², S. Ogawa³, M. Takechi⁴, T. Takata⁴, T. Taniguchi⁵

¹IUHW Mita Hospital, Tokyo, Japan
²Fujita Health University, Nagoya, Japan
³Ogawa Satoshi Clinic, Tokyo, Japan
⁴Cardiointelligence Inc., Tokyo, Japan
⁵Ritsumeikan University - Faculty of Sport and Health Science, Shiga, Japan

Funding Acknowledgements: Type of funding sources: Public grant(s) – National budget only. Main funding source(s): Japan Agency for Medical Research and Development

Backgrounds: Machine learning (ML) models using convolutional neural networks (CNNs) have been successfully applied to arrhythmia detection and prediction. The current study aimed to develop a CNN-based AI algorithm for the prediction of paroxysmal atrial fibrillation (PAF) patients during sinus rhythm waveforms from Holter electrocardiograms.

Methods: A ML model was developed by using 2-leads from Holter ECG waveforms divided into slots of 30 seconds each. As training data, we prepared > 300000 slots with sinus rhythm and they were subsequently converted to image data for application to the CNN-based AF detection algorithm. The performance metrics obtained for the ML model for PAF detection were as follows: sensitivity, 85.2 %; specificity, 64.0 % and AUC for PAF prediction was 0.84. To validate the performance of the ML model, multicenter clinical trial was conducted. The clinical trial was conducted at three independent centers and included 24 patients with PAF and 20 patients in true sinus rhythm. Five factors were matched between the PAF and sinus groups: age, gender, hypertension under treatment, diabetes under treatment, and smoking. Waveform data from validation datasets to Grad-CAM were applied to obtain explainable AI (XAI) images, which allowed for a heat map over the waveform to be produced.

Results: For the primary endpoint, the estimated sensitivity and specificity [95% CI] were 91.7% [73.0-99.0%] and 65.0% [40.8-84.6%], respectively (Table 1). The sensitivity and specificity of the secondary endpoints using 10 time-independent slots were the same as for the primary endpoint. Regarding safety endpoints, no adverse events attributable to the test device itself were observed. Based on the regions of interest highlighted via Grad-CAM, the most contributing featured regions to predict PAF were indicated by areas on P waves and QT area (Figure 1). Featured regions to rule out the existence of PAF indicated areas just on QRS waves were also present. It was therefore determined that the proposed CNN-based ML model focuses on the P-wave area and atrial repolarization area in the case of PAF.

Conclusion: The study suggested that 2-lead Holter ECG with a 30 second slot could detect PAF patients from their sinus rhythm waveforms and the performance was more effectively than using 12-lead ECGs. The results of this clinical trial are expected to contribute to the screening of patients with PAF that could not be detected by a single time Holter ECG.

Table 1. The table below shows the estimated values for the sensitivity and specificity of the PPS tests in the target population.

<table>
<thead>
<tr>
<th>PPS</th>
<th>N</th>
<th>Frequency</th>
<th>Estimated Value (%)</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>24</td>
<td>22</td>
<td>91.7</td>
<td>73</td>
<td>98.97</td>
</tr>
<tr>
<td>Specificity</td>
<td>20</td>
<td>13</td>
<td>65</td>
<td>40.78</td>
<td>84.61</td>
</tr>
</tbody>
</table>

Table 1
Figure 1: Grad-CAM images from PAF cases

The figures above show the contributing features for the PAF prediction. The heatmap with red area shows key diagnostic features for the prediction of PAF.