Performance of an automated photoplethysmography-based artificial intelligence algorithm to detect atrial fibrillation

Check for updates

Daniel Mol, BSc,*[†] Robert K. Riezebos, MD, PhD,* Henk A. Marquering, PhD,[†]
Marije E. Werner, MD,* Trudie C.A. Lobban, MBE, FRCP,[‡]
Jonas S.S.G. de Jong, MD, PhD,*[1] Joris R. de Groot, MD, PhD[†1]

From the *Department of Cardiology, OLVG Hospital, Amsterdam, The Netherlands, [†]Department of Cardiology, Amsterdam University Medical Centres, University of Amsterdam, Amsterdam, The Netherlands, and [‡]Arrhythmia Alliance, Chipping Norton, United Kingdom.

Atrial fibrillation (AF) is the most common sustained arrhythmia. Symptoms may include palpitations, dyspnea, limited exercise tolerance, and fatigue. However, AF is asymptomatic in 12%–25% of cases and remains undetected in 1.4% of patients aged >65 years.[1,2] Early detection and adequate treatment of silent AF is essential, especially in patients who are at increased risk for stroke.

Photoplethysmography (PPG) using the camera of smartphones and smartwatches is a promising technology for heart rate and rhythm assessment. To retrieve PPG recordings, a photoemitter is connected to a photoreceiver. The amount of light absorbed or reflected by the blood is related to the cardiac cycle. Although PPG technology in smartphones or smartwatches is easy to use and publicly accessible, reliable PPG recording requires good signal quality, which may be affected by many factors, including poorly perfused tissue, tremors, ambient light, camera characteristics, and correct placement of the PPG sensor.

Recently, an artificial intelligence smartphone-based PPG algorithm for detection of AF was developed by Happitech (Amsterdam, The Netherlands). The algorithm was trained using 2560 selected recordings retrieved from a worldwide online data donation campaign (Heart for Heart) and consists of 3 main components: (1) peak detection to measure R-R intervals; (2) quality; and (3) rhythm classification. The most critical part of rhythm classification is peak detection. Proper signal and noise discrimination are essential, especially for AF detection, because of the irregular nature of the arrhythmia. Motion artifacts, variations in heart rate, and peak size increase the number of incorrectly assigned peaks. To overcome this issue, the algorithm uses a shallow

neural network (SNN). The SNN outperformed 3000 different settings of MATLAB (Natick, MA) peak findings functionality (Figure 1). PPG recordings of approximately 90 seconds were used. Each recording was divided into three 30-second windows for quality and rhythm classification. The support vector machine classified each window as low, medium, or high quality based on waveform and vibrations; and as sinus rhythm (SR), AF, or undetermined based on several rhythm and signal features, such as heart rate variability parameters, peak amplitude, and other signal characteristics. The complete recording was labeled as SR, AF, or undetermined if ≥2 segments were assigned to the same group (Figure 2).

We validated the algorithm in patients with AF who were admitted to OLVG Hospital (Amsterdam, The Netherlands) for elective electrical cardioversion (ECV). PPG recordings were obtained directly before and after ECV using an iPhone 8 (Apple Inc., Cupertino, CA). Continuous electrocardiography was monitored simultaneously with the PPG heart rhythm recording for verification. The study was approved by the local medical ethics committee, and all participants provided written informed consent.

In total, 161 patients were eligible for the validation study. Of these patients, 12 were excluded (8 had an atrial arrhythmia other than AF; 1 had a ventricular paced rhythm; 1 converted to SR before the recording; 1 withdrew informed consent; and 1 had been included in the study before) (Figure 3). Thus, the algorithm was validated in 149 patients between March 2018 and March 2019. Of these patients, 85 (57%) were male, mean age 69 ± 9 years, body mass index 27.2 ± 5.0 kg/m^2, and CHA$_2$DS$_2$-VASc score 2.0 ± 1.2. All patients performed a PPG recording during AF. After ECV, 41 patients were excluded (1 unable to perform PPG recording; 1 experienced a technical issue; 39 in whom ECV was unsuccessful or was not performed due to inadequate anticoagulation, congestive heart failure, or ongoing infection) (Figure 3). PPG recordings during SR were obtained in 108 patients. Two hundred sixteen recordings

https://doi.org/10.1016/j.cvdhj.2020.08.004

KEY FINDINGS

- The shallow neural network showed excellent performance for peak detection.

- High sensitivity and specificity for detection of atrial fibrillation were obtained using a new automated plethysmography (PPG) algorithm.

- Predefined exclusion of recordings with low confidence boosted the diagnostic performance of the algorithm, resulting in 1.8% increase in sensitivity and 4.6% increase in specificity.

---

were analyzed. In 20 recordings, the first attempt was of low quality and a second recording obtained immediately after the first recording was used for analysis. High signal quality was observed in 201 recordings (SR 100, AF 101), medium quality in 12 recordings (SR 7, AF 5), and low signal quality in 3 recordings (SR 1, AF 2) ($P = .72$) (Figure 4). Mean heart rate was 78.6 ± 21.7 bpm (SR 64.6 ± 10.5 bpm vs AF 92.4 ± 21 bpm; $P <.001$). The algorithm correctly classified AF in 104 patients and incorrectly classified AF in 2 patients. The recording was labeled undetermined in 2 patients. Among the PPG recordings during SR, 101 were correctly classified, 2 were incorrectly classified, and in 5 the algorithm outcome was undetermined. The signal quality of the undetermined recordings was high in 6 and medium in 1. By excluding the undetermined recordings, sensitivity improved from 96.3% (95% confidence interval [CI] 90.8%–99.0%) to

98.1% (95% CI 93.4%–99.8%), and specificity improved from 93.5% (95% CI 87.1%–97.4%) to 98.1% (95% CI 93.2%–99.8%) (Figure 4). In patients who only performed a pre-ECV recording, PPG recording was true positive in 35, false positive in 1, and in 5 recordings the outcomes remained undetermined. Sensitivity with and without the "undetermined" outcome was 85.4% (95% CI 70.8%–94.4%) and 97.2% (95% CI 85.5%–99.9%), respectively.

Smartphone-based PPG technology is able to identify AF with reported sensitivity of 89.9%–95.3% and specificity of 90.9%–99.7%.[3–7] Unfortunately, PPG analysis alone results in a significant number of recordings that are insufficient for definitive classification. In our study, signal quality was high in 93%, medium in 5.6%, and low 1.4%. Of note, high signal quality was also observed in 86% of the undetermined recordings. Therefore, we should not only eliminate the recordings with low signal quality due to incorrectly performed recording or poorly perfused tissue; rather, we should exclude the recordings with low confidence. To improve sensitivity and specificity, other investigators have excluded only recordings without a diagnosis.[5]

In most studies, PPG recordings were supervised in a clinical setting. PPG recordings obtained in nonsupervised environments could potentially result in a higher number of insufficient recordings. Verbrugge et al[8] studied participants who performed unsupervised recordings. They observed an increase in signal quality when recordings were obtained for 7 days. Therefore, it should be accepted that, in some cases, multiple attempts are needed to provide an acceptable
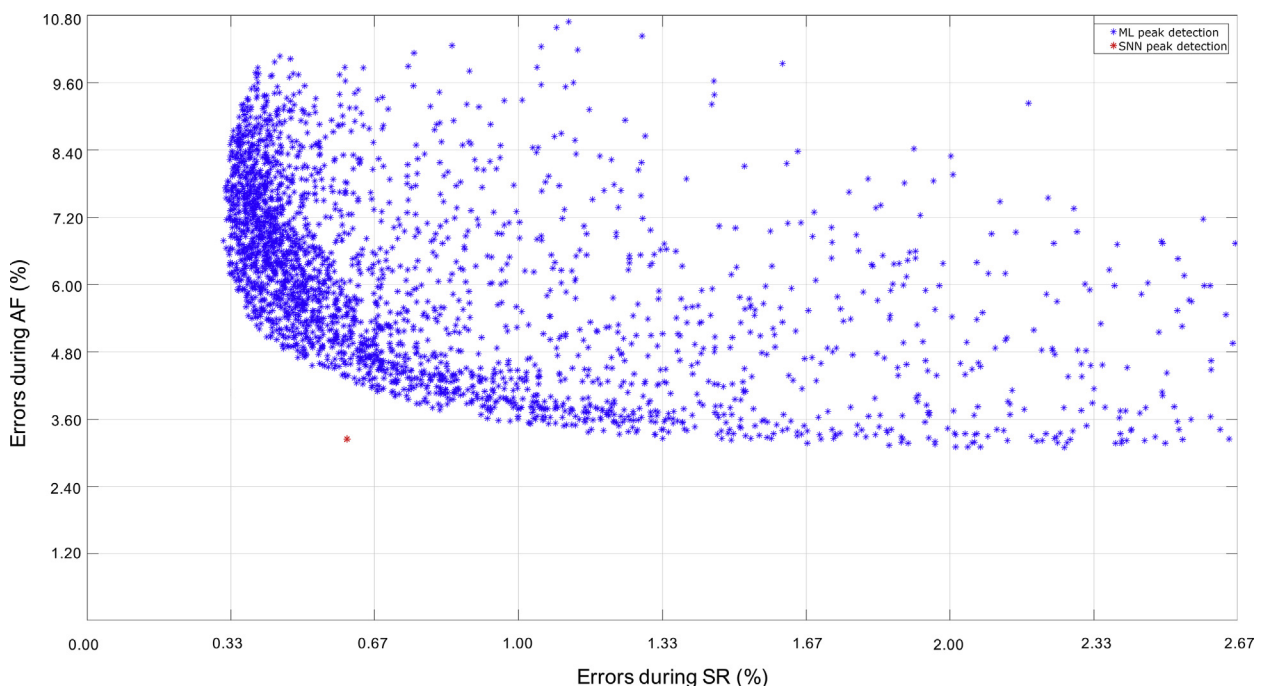

Figure 1 Peak detection errors (%) during sinus rhythm (SR) (x-axis) and atrial fibrillation (AF) (y-axis) using MATLAB findpeaks (ML) *(blue dots)* and shallow neural network (SNN) *(red dot)*. Errors included false and missed peaks.
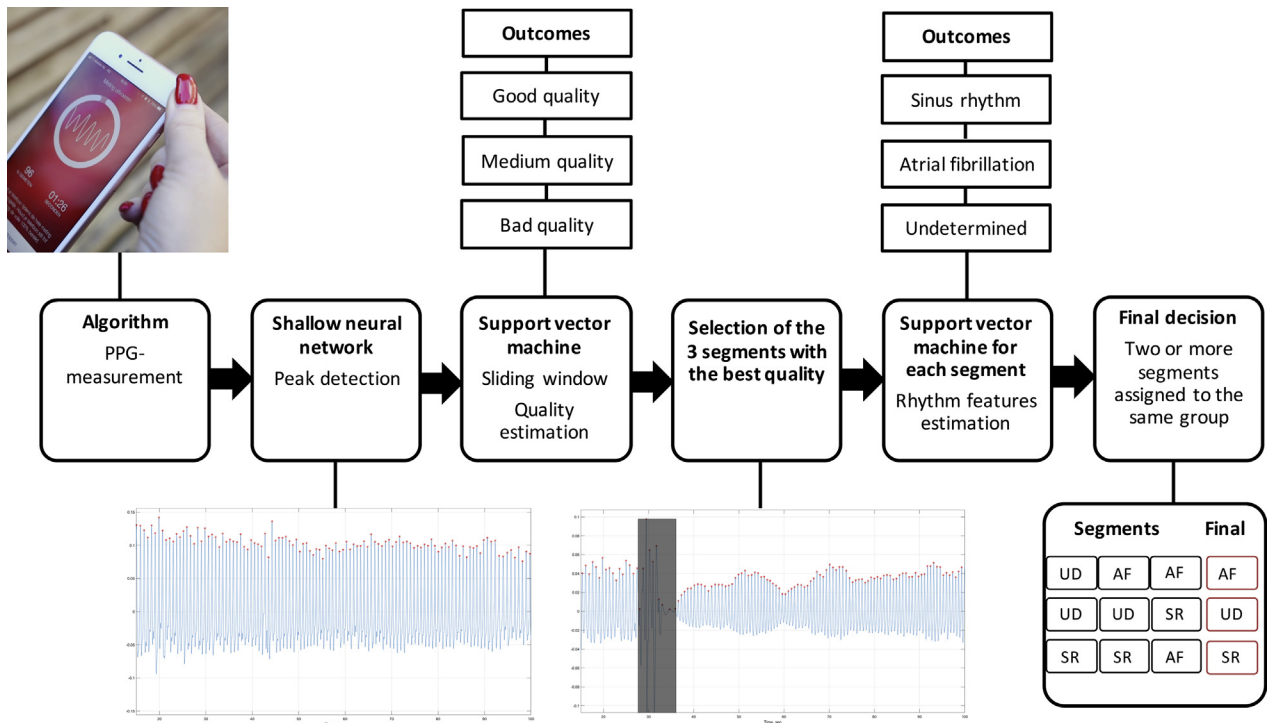
Figure 2   Steps taken by the photoplethysmography (PPG) algorithm to provide heart rhythm outcomes. First is detection of peaks using a shallow neural network; second is quality estimation using the support vector machine. After selection of 3 segments with the best quality in the third step, each segment is, based on rhythm features, classified as sinus rhythm (SR), atrial fibrillation (AF), or undetermined (UD). The final decision was made if ≥ 2 segments were classified in the same group.

outcome and that, despite these multiple recordings, it still may be impossible to produce acceptable data for interpretation. The specific advantage of our AF detection algorithm lies in the strategy of excluding the low-confidence rather than the low-quality recordings.

PPG technology has low costs and does not require additional hardware; therefore, it can be used as a screening tool in a wide population. However, a screening tool should provide reliable outcomes, as a higher number of unreliable outcomes would lead to unnecessary medical or clinic visits and treatments. PPG-based detection of AF ideally should be performed in a population with a sufficiently increased risk for AF. For example, if pretest probability is 5%, a positive test would result in a posttest probability of AF of only
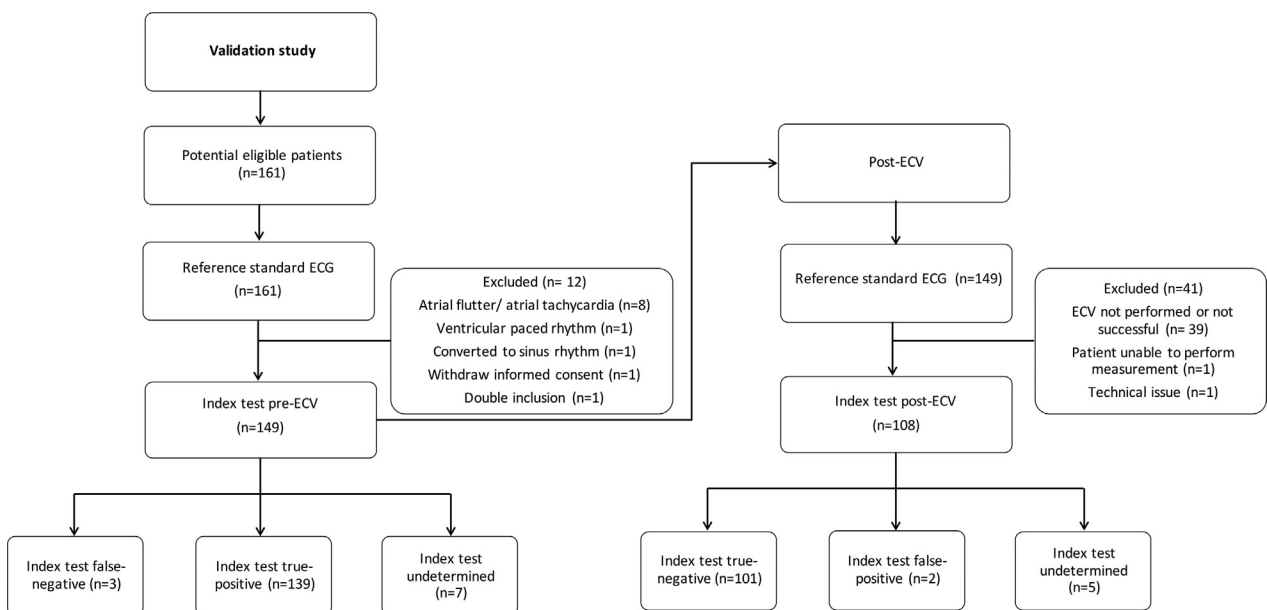

Figure 3   Flow diagram validation study. ECG = electrocardiography; ECV = electrical cardioversion.
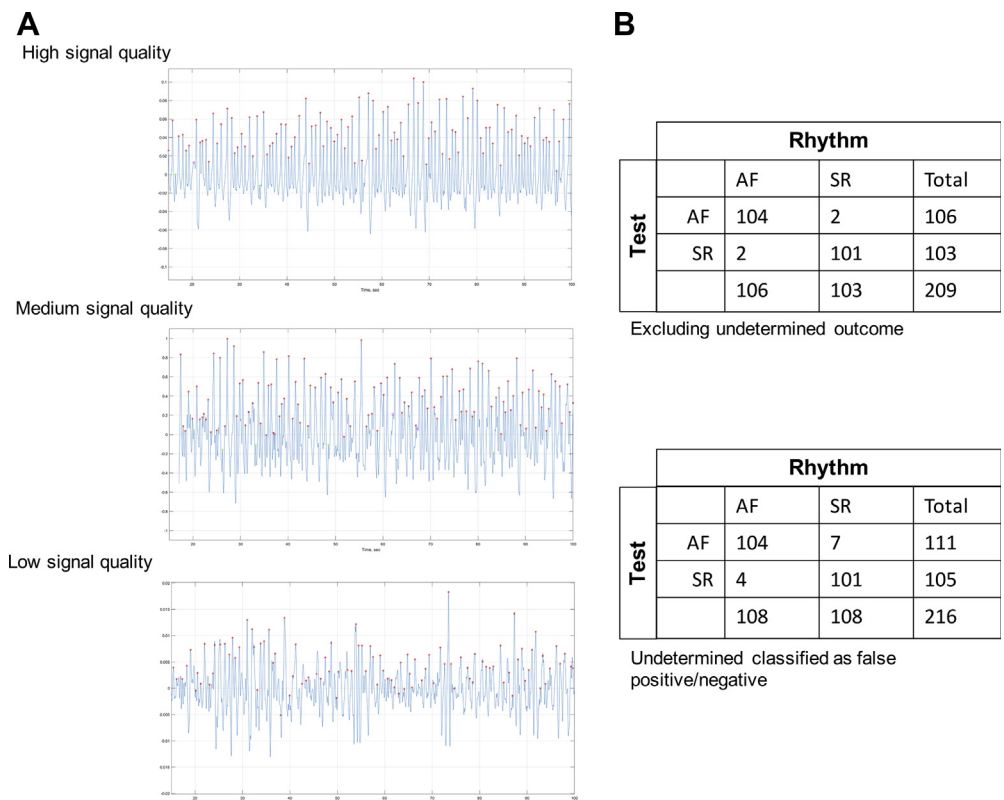
A

B

Figure 4   Signal quality and 2×2 tables. **A:** Example of high, medium, and low signal quality plethysmography recordings during atrial fibrillation (AF). **B:** 2× 2 tables with algorithm outcome. Top table does not include recordings classified as undetermined (sensitivity 98.1%, 95% confidence interval [CI] 93.4%–99.8%; specificity 98.1%, 95% CI 93.2%–99.8%). In the bottom table, undetermined recordings are classified as false positive or false negative (sensitivity 96.3%, 95% CI 90.8%–99.0%; specificity 93.5%, 95% CI 87.1%–97.4%). SR = sinus rhythm.

73% with the characteristics of this algorithm. Therefore, as with all diagnostic tests, the pretest probability in the tested population should be sufficient.

This study has several limitations. The algorithm was validated in patients with AF, and whether the algorithm is able to detect atrial tachycardias is unclear. All recordings in this study were supervised by a local investigator. The performance of this algorithm in a nonclinical environment and in a population with a lower pretest probability needs further investigation.

Funding Sources

Disclosures

Dr. Riezebos and Dr. de Jong are minor shareholders of Happitech. Dr. de Groot reports grants from AtriCure, Abbott, Boston Scientific, and Medtronic; and personal fees from AtriCure, Bayer, Daiichi Sankyo, Servier, Johnson & Johnson, Novartis, and Medtronic, outside the submitted work. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

References

1. Zoni-Berisso M, Lercari F, Carazza T, Domenicucci S. Epidemiology of atrial fibrillation: European perspective. Clin Epidemiol 2014;6:213–220.
2. Lowres N, Neubeck L, Redfern J, Freedman SB. Screening to identify unknown atrial fibrillation. A systematic review. Thromb Haemost 2013;110:213–222.
3. Chan PH, Wong CK, Poh YC, et al. Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting. J Am Heart Assoc 2016;5:e003428.
4. Brasier N, Raichle CJ, Dorr M, et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). Europace 2019;21:41–47.
5. Proesmans T, Mortelmans C, Van Haelst R, Verbrugge F, Vandervoort P, Vaes B. Mobile phone-based use of the photoplethysmography technique to detect atrial fibrillation in primary care: diagnostic accuracy study of the FibriCheck app. JMIR Mhealth Uhealth 2019;7:e12284.
6. Rozen G, Vaid J, Hosseini SM, et al. Diagnostic accuracy of a novel mobile phone application for the detection and monitoring of atrial fibrillation. Am J Cardiol 2018;121:1187–1191.
7. Fan YY, Li YG, Li J, et al. Diagnostic performance of a smart device with photoplethysmography technology for atrial fibrillation detection: pilot study (Pre-mAFA II Registry). JMIR Mhealth Uhealth 2019;7:e11437.
8. Verbrugge FH, Proesmans T, Vijgen J, et al. Atrial fibrillation screening with photoplethysmography through a smartphone camera. Europace 2019;21:1167–1175.