Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECTAF PRO)

Noé Brasier1†, Christina J. Raichle1†, Marcus Dörr2,3, Adrian Becke2, Vivien Nohturfft2, Stefan Weber4, Fabienne Bulacher1, Lorena Salomon1, Thierry Noah1, Ralf Birkemeyer5, and Jens Eckstein1,6*

1CMIO office, University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland; 2Department of Internal Medicine B, University Medicine Greifswald, Greifswald, Germany; 3German Centre for Cardiovascular Research (DZHK), partner site Greifswald, Greifswald, Germany; 4Department of Internal Medicine, University Hospital Regensburg, Regensburg, Germany; 5Herzklinik Ulm, Ulm, Germany; and 6Department of Internal Medicine, University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland

Received 20 April 2018; editorial decision 1 July 2018; accepted 25 July 2018

Aims
Early detection of atrial fibrillation (AF) is essential for stroke prevention. Emerging technologies such as smartphone cameras using photoplethysmography (PPG) and mobile, internet-enabled electrocardiography (iECG) are effective for AF screening. This study compared a PPG-based algorithm against a cardiologist’s iECG diagnosis to distinguish between AF and sinus rhythm (SR).

Methods and results
In this prospective, two-centre, international, clinical validation study, we recruited in-house patients with presumed AF and matched controls in SR at two university hospitals in Switzerland and Germany. In each patient, a PPG recording on the index fingertip using a regular smartphone camera followed by iECG was obtained. Photoplethysmography recordings were analysed using an automated algorithm and compared with the blinded cardiologist’s iECG diagnosis. Of 672 patients recruited, 80 were excluded mainly due to insufficient PPG/iECG quality, leaving 592 patients (SR: n = 344, AF: n = 248). Based on 5 min of PPG heart rhythm analysis, the algorithm detected AF with a sensitivity of 91.5% (95% confidence interval 85.9–95.4) and specificity of 99.6% (97.8–100). By reducing analysis time to 1 min, sensitivity was reduced to 89.9% (85.5–93.4) and specificity to 99.1% (97.5–99.8). Correctly classified rate was 88.8% for 1-min PPG analysis and dropped to 60.9% when the threshold for the analysed file was set to 5 min of good signal quality.

Conclusion
This is the first prospective clinical two-centre study to demonstrate that detection of AF by using a smartphone camera alone is feasible, with high specificity and sensitivity. Photoplethysmography signal analysis appears to be suitable for extended AF screening.

Clinical trial registration

Keywords
Atrial fibrillation • Arrhythmia • Photoplethysmography • Smartphone • eHealth
What’s new?

- Photoplethysmography signal recorded by a smartphone camera alone and analysed by automated algorithms appears to be suitable for clinical use, given a sufficient signal quality.
- In this multicentre, international, prospective trial, specificity for detection of atrial fibrillation (AF) was 99.1%, which is important for a potential screening tool to avoid a high false positive rate.
- Validation was performed using a recent FDA-approved mobile internet-enabled electrocardiography interpreted by board certified cardiologists.
- The results of this trial strongly support the recent European Heart Rhythm Association (EHRA) recommendations for AF screening.
- The application has the potential to transform any smartphone into an easy-to-use AF screening device.
- Validation in a larger-scale population screening trial is the next step, as recommended by the AF-SCREEN international Collaboration.

Introduction

The prevalence of atrial fibrillation (AF) increases with age and reaches 24.2% in men and 16.1% in women older than 85 years of age. Early diagnosis of AF and, if indicated, anticoagulation therapy, are of utmost importance to prevent AF-related complications such as stroke. Despite extensive long-term electrocardiographic (ECG) monitoring, AF detection remains a challenge due to its paroxysmal nature and its weak temporal correlation with stroke occurrence. It has been shown that AF detection rates are higher if screening periods are prolonged up to 1 year. Since stroke remains a frequent initial manifestation of AF, there is a need for a reliable, cost-effective, convenient and easy-to-apply, long-term AF screening device.

At the same time, medical-grade smartphone apps based on photoplethysmography (PPG) signal analysis constitute a promising, non-invasive and cost-effective option for AF screening. Without the need for additional devices and fully functional on a smartphone alone, PPG-based apps record signals from the patient’s fingertip, solely using the built-in camera. As shown in recent trials, the PPG-based apps can differentiate between AF and sinus rhythm (SR) with a sensitivity and specificity of ≥90%. The development of single-lead mobile, internet-enabled ECGs (iECGs) connected to smartphones (e.g., AliveCor® Kardia Monitor) show high potential for AF screening. The REHEARSE-AF study showed a three-fold higher AF detection rate using iECG in ambulatory patients over 12 months compared with routine care.

In this single-blinded, prospective, international, two-centre clinical validation study (DETECT AF PRO), we validated the performance of a PPG-based automated algorithm in differentiating between AF and SR, and compared its results to an iECG interpreted by at least two blinded cardiologists.

Methods

Participants

Between September 2016 and October 2017, patients were recruited for this prospective, clinical validation study at the University Hospital Basel, Switzerland, and at the Department of Cardiology and Pulmonology at the University Medicine Greifswald, Germany. Hospitalized patients aged 18 years or older, without pacemaker or implanted defibrillator and able to give written informed consent, were eligible. The study population was not limited to patients above 65 years of age, because the screening tools tested in this trial are also used by younger individuals when they experience palpitations or just want to monitor their health status. A high number of false positive results could then cause a significant cost burden for the health system.

Electronic patient records of hospitalized patients were screened for AF history. If a diagnosis of AF was present, the patients who consented were recruited. Age- and sex-matched patients without a history of AF in their medical records were recruited as potential matches for the SR group. It was taken into account that some of the patients with a history of AF were likely to be in SR at the time of recruitment. Because the ECGs were analysed in a blinded fashion, the final allocation of the patients to the respective groups was only possible after recruitment was closed and all data analysed. This design allowed a prospective study character as well as evaluation of test criteria.

The study protocol complied with the Declaration of Helsinki and was approved by the local ethics committees (EKNZ 2016-01176, BB 140/16) and was registered with ClinicalTrials.gov as NCT02949180. The Clinical Trial Unit Basel provided independent data monitoring. The authors declare that all supporting data are available within the article and Supplementary material online.

Test methods

In this standardized technical setup, we used regular smartphones (iPhone 4S, Apple®, Cupertino, CA, USA) and installed a study version (PreventicusRhythm, V1.0) of the commercially available Heartbeats app (Preventicus®, Jena, Germany). Compared with the commercially available version for Android and iOS, where a continuous and automated quality check is obtained, the study version included neither a signal quality check nor a readout of any results. Performance feature reduction was intentionally chosen to prevent any possible bias by study personnel.

After obtaining written informed consent, a single-lead iECG was recorded on a commercially available, mobile, iECG (AliveCor®, Mountain View, CA, USA). Patients were asked to put their left index and middle fingers on the left electrode of the iECG and their right index and middle fingers on the right electrode. Internet-enabled ECG was recorded for 1 min to obtain a long enough documentation to diagnose or exclude AF (Figure 1). The iECG was saved as a PDF document for later analysis by at least two board-certified cardiologists. The iECG device was chosen as reference standard as it is recommended for AF screening by the European Heart Rhythm Association (EHRA) and is FDA-approved for AF screening, as well as more comfortable than a standard 12-lead ECG.

The smartphone camera was placed on an index fingertip for 5-min pulse wave recording (Figure 2). Measurements were performed in a quiet surrounding and patients were instructed to remain in a comfortable sitting position to reduce movement artefacts.

After the predefined recruitment goal was achieved, PPG files were pseudonymised and sent to Preventicus® for blinded automated analysis. Cosehesive 1-, 3-, and 5-min segments were extracted from each patient’s 5-min PPG file, based on best signal-to-noise ratio. There were two
reasons to analyse different predefined interval lengths of PPG files. First, we aimed to determine whether a 1-min analysis would be suitable for AF-screening purposes, because a reduced procedure time most likely increases compliance. Second, we anticipated that not every PPG file recorded with the trial version of the app would contain 5 min of good recording quality as the trial app version lacks a signal-quality check. All three PPG data strips were then analysed with the automated Heartbeats algorithm.

**Heartbeats algorithm**
The Heartbeats algorithm (Version 20171120) analyses PPG signals recorded by a standard smartphone camera as described elsewhere.\(^{10}\) The algorithm discriminates between SR and absolute arrhythmia consistent with AF, using a complex non-linear combination analysis comprising beat-to-beat changes of pulse wave time intervals and pulse wave morphology.\(^{10,11}\) Prior to rhythm analysis, the algorithm performs an automatic data quality check by analysing signal-to-noise ratio and accelerometer data. Noisy data segments (e.g., movement artefacts) are detected and excluded automatically from analysis. If the percentage of noisy or disturbed data segments compared with good quality data exceeds the threshold of 10%, analysis is rejected per se. For example, a 5-min analysis is rejected if more than 30 s were noisy or disturbed. If less than 30 s were noisy or disturbed, those data segments were detected and removed from analysis automatically. For the 3-min segment analysis, the algorithm screened the 5-min files for a 3-min signal segment with less than 10% noisy signal. For the 1 min analysis, the algorithm did the same for a segment of 1 min.

**Kardia Monitor algorithm**
The Kardia Monitor algorithm (Version 4.2.0.1487) provides automated diagnostic information about a patient’s heart rhythm based on an iECG recorded by a mobile device. It automatically generates a PDF document with the original iECG tracing and diagnosis. Possible diagnoses are
shown earlier by Krivoshei et al. The assumption that the app sensitivity and specificity are each 95%, as shown by McGhie et al., was blinded for the diagnosis and any other patient data. Results of the 1-, 3-, and 5-min analysis, just as Preventicus was blinded for the iECG diagnosis and other patient data. Results of the 1-, 3-, and 5-min PPG segment analysis and iECG diagnosis were returned to University Hospital Basel for final analysis.

**Analysis**

Sample size was calculated according to the method of Blaker, based on the assumption that the app sensitivity and specificity are each 95%, as shown earlier by Krivoshei et al. Aiming at a specificity and sensitivity of 95%, confidence intervals above 90%, and a power of 0.8, a sample size of >660 patients was calculated.

The Heartbeats diagnosis (AF or SR) was validated against iECGs analysed by two blinded cardiologists. In case of uncertainty, a third cardiologist was consulted and blinded for the previous diagnosis. All cardiologists were blinded for the results of PPG analysis, just as Preventicus was blinded for the iECG diagnosis and other patient data. Results of the 1-, 3-, and 5-min PPG segment analysis and iECG diagnosis were returned to University Hospital Basel for final analysis.

In an additional analysis, the automated diagnosis of the Kardia Monitor algorithm was validated against the cardiologists’ diagnosis. Results were analysed using the χ² test and Fisher’s exact test. Sensitivity and specificity, for each 1-, 3-, and 5-min analysis were obtained using a contingency table. Calculations were performed using SPSS version 22.

Examples of PPG and iECGs recordings in SR, AF, and SR with premature beats are depicted in Figures 1 and 2 in the Supplementary material online, Figures S1–S6.

**Results**

Between September 2016 and October 2017, a total of 672 patients were recruited, and 80 were excluded from final analysis. The most frequent reasons for exclusion were insufficient PPG signal quality (44) and insufficient iECG signal quality (18) (Figure 3). A total of 592 patients (SR: n = 344, AF: n = 248) with at least 1 min of sufficient PPG signal quality (cohesive 1-, 3-, or 5-min analysis) and interpretable iECG were included for final analysis (Figure 3).

In the two groups, median (interquartile range) age was 78 (13) years, and 84 patients (14.2%) were below 65 years of age. Sex ratio was malefemale 324:268, 191:153, and 133:115.

**Table 1** Baseline characteristics of the study population

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>SR</th>
<th>AF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), median (IQR)</td>
<td>78 (13)</td>
<td>77 (14)</td>
<td>79 (13)</td>
</tr>
<tr>
<td>Range (low:high)</td>
<td>31:95</td>
<td>31:95</td>
<td>35:95</td>
</tr>
<tr>
<td>Sex male:female</td>
<td>324:268</td>
<td>191:153</td>
<td>133:115</td>
</tr>
<tr>
<td>Study centre Basel: Greifswald</td>
<td>331:261</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECG rhythm SR:AF</td>
<td>344:248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>427 (72.1)</td>
<td>241 (70.1)</td>
<td>186 (75.0)</td>
</tr>
<tr>
<td>Heart failure (%)</td>
<td>222 (37.5)</td>
<td>95 (27.6)</td>
<td>127 (51.2)</td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>185 (31.3)</td>
<td>101 (29.4)</td>
<td>84 (33.9)</td>
</tr>
<tr>
<td>Stroke (%)</td>
<td>94 (15.9)</td>
<td>44 (12.8)</td>
<td>50 (20.2)</td>
</tr>
<tr>
<td>CHA2DS2-VASc mean points</td>
<td>4.02</td>
<td>3.79</td>
<td>4.26</td>
</tr>
<tr>
<td>OACs (%)</td>
<td>288 (48.6)</td>
<td>116 (33.7)</td>
<td>172 (69.4)</td>
</tr>
<tr>
<td>Antiplatelet therapy (%)</td>
<td>226 (38.2)</td>
<td>161 (46.8)</td>
<td>65 (26.2)</td>
</tr>
<tr>
<td>Aspirin (%)</td>
<td>207 (35)</td>
<td>151 (43.9)</td>
<td>56 (22.6)</td>
</tr>
<tr>
<td>Clopidogrel (%)</td>
<td>57 (9.6)</td>
<td>32 (9.3)</td>
<td>25 (10.1)</td>
</tr>
<tr>
<td>Antiarrhythmic therapy (%)</td>
<td>81 (13.7)</td>
<td>36 (10.5)</td>
<td>45 (18.1)</td>
</tr>
<tr>
<td>Amiodarone (%)</td>
<td>29 (4.9)</td>
<td>19 (5.5)</td>
<td>10 (4.0)</td>
</tr>
<tr>
<td>Others (%)</td>
<td>52 (8.8)</td>
<td>17 (4.9)</td>
<td>35 (14.1)</td>
</tr>
<tr>
<td>ACE inhibitors (%)</td>
<td>200 (33.8)</td>
<td>107 (31.1)</td>
<td>93 (37.5)</td>
</tr>
<tr>
<td>ARBs (%)</td>
<td>170 (28.7)</td>
<td>99 (28.8)</td>
<td>71 (28.6)</td>
</tr>
<tr>
<td>CCB (%)</td>
<td>151 (25.5)</td>
<td>97 (28.2)</td>
<td>54 (21.8)</td>
</tr>
<tr>
<td>Beta blockers (%)</td>
<td>414 (69.9)</td>
<td>218 (63.4)</td>
<td>196 (79)</td>
</tr>
<tr>
<td>Diuretics (%)</td>
<td>370 (62.5)</td>
<td>180 (52.3)</td>
<td>190 (76.1)</td>
</tr>
</tbody>
</table>

AF, atrial fibrillation; ACE, angiotensin converting enzyme; ARBs, angiotensin II receptor blockers; CCBs, calcium channel blockers; ECG, electrocardiogram; IQR, interquartile range; OAC, oral anticoagulant; SR, sinus rhythm.

**Table 2** Accuracy of the Heartbeats algorithm

<table>
<thead>
<tr>
<th>Test criterion</th>
<th>1-min analysis</th>
<th>3-min analysis</th>
<th>5-min analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value (%)</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>89.9 (85.5–93.4)</td>
<td>91.3 (86.5–94.7)</td>
<td>91.5 (85.9–95.4)</td>
</tr>
<tr>
<td>Specificity</td>
<td>99.1 (97.5–99.8)</td>
<td>98.7 (96.7–99.6)</td>
<td>99.6 (97.8–100)</td>
</tr>
<tr>
<td>No diagnosis</td>
<td>6.7</td>
<td>13</td>
<td>32.2</td>
</tr>
<tr>
<td>CCR (%)</td>
<td>88.8</td>
<td>77.6</td>
<td>60.9</td>
</tr>
</tbody>
</table>

CCR, correctly classified rate; CI, confidence interval.
Second sentence... is 4.1 min, out of the 5 min obtained.

**Table 3**  Accuracy of the Kardia algorithm

<table>
<thead>
<tr>
<th>Test criterion</th>
<th>Value % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>99.6 (97.9–100)</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>97.8 (95.3–99.2)</td>
</tr>
<tr>
<td>No diagnosis (%)</td>
<td>18.8</td>
</tr>
<tr>
<td>CCR (%)</td>
<td>82.2</td>
</tr>
</tbody>
</table>

CCR, correctly classified rate; CI, confidence interval.

Discussion

This study evaluated the ability of the Heartbeats algorithm to differentiate between SR and AF based on analysis of PPG signals recorded by a standard smartphone camera. The algorithm achieved a specificity of 99.1% and a sensitivity of 89.9% for cohesive 1-min PPG segments. Prolonging the analysed file from 1 to 5 min did not greatly improve test results, and led to a significant signal quality decrease in analysable files from 93.3% (1-min test) to 67.7% (5-min test).

The AliveCor® iECG device was used as a reference to document the heart rhythm at the time of recruitment. This device is widely used and is FDA approved. The device provided satisfactory comfort for our patients and good quality iECGs in the majority of recordings. For highest validation, at least two blinded cardiologists validated the performance of the Heartbeats algorithm and the Kardia Monitor algorithm by interpreting standard iECGs. The Kardia Monitor algorithm reached a specificity of 99.6% and a sensitivity of 97.8% for the files deemed suitable for automated interpretation. Notably, in 18.8% of iECG files, the Kardia Monitor algorithm did not perform an automated interpretation, although cardiologists were still able to identify the cardiac rhythm.

Although the trial was not designed to compare the two methods, it was a predefined aim to validate the performance of the PPG algorithm on AF screening added mobile screening technologies using PPG or iECG signal analysis to their screening recommendations. Although both technologies reach the requirements for the intended use, the PPG signal-based approach offers the advantage that it can be used without the need for additional devices other than widely available smartphones.

However, both methods tested depend on sufficient signal quality to achieve a high enough sensitivity and specificity. To address this, an automated signal quality check has been implemented in the commercial version of the PPG algorithm, but was deactivated in the trial to prevent bias.

To our knowledge, this study is to date the largest blinded, prospective, international, two-centre study to successfully validate a PPG-based algorithm that differentiates between AF and SR using a smartphone camera alone. Despite these promising results, final diagnosis of AF in patients remains ECG-based. The current standard for AF screening is a Holter ECG device. Modern Holter devices are easy to use, lightweight, and small. They permit digital data export and frequently provide automated AF detection software. Major shortcomings of Holter ECG devices are equipment costs and lack of comfort for patients relating to skin irritation caused by adhesive gel from electrodes. PPG-based screening avoids these significant limitations. Therefore, PPG-based, hardware-independent algorithms for AF detection represent a promising alternative to other established technologies, and enable convenient AF screening given that smartphones and health apps are already widely used around the globe. Even though no population-based data for PPG AF screening is available yet, a consumer market price of less than 50 USD/year for an app could represent a cost-effective AF screening solution. A PPG-based algorithm has been validated in a trial using wrist-worn devices and preliminary data are promising. A wrist-worn solution would enable continuous AF screening during routine activities but is limited by the availability of only a few suitable devices, compared with that of suitable standard smartphones.
In the next step, this algorithm should be tested in a larger population-based setting to validate its potential for AF screening, as recently recommended by the AF-SCREEN International Collaboration.17

The Cardiogram® algorithm used in the Health eHeart study detects sudden changes in heart rate averaged over 5 s, thereby can potentially identify AF episodes if they are related to higher ventricular rates. This algorithm was tested in the Apple Watch® and Garmin® devices.18,19 The same hardware dependence is present for a novel algorithm now used in the recently launched Apple Heart Study (NCT03335800). This algorithm analyses the PPG signal of the pulse wave, and detects cardiac arrhythmias with a specific app to be installed on an Apple Watch.

Such algorithms, once validated in state-of-the-art trials, could as well be integrated in existing disease management apps for AF patients. The rhythm prevalent at the time of symptoms is important information, as patients with paroxysmal AF often relate non-specific symptoms of arrhythmia, even though arrhythmia is not present at that time. On the other hand, patients can be completely asymptomatic during an arrhythmia episode, and therefore can be unaware that they still have AF episodes, even after pulmonary vein isolation.20 The integration of such an objective documentation in the recently launched myAF-app of the European Society of Cardiology could support management of AF patients as well as empower them through feedback and education.16

Limitations

There are some limitations to be pointed out with this trial. Patient allocation to the AF or SR group was not randomized, but instead was preselected based on the medical history, and was then finalized, based on the reference ECG recorded together with the PPG signals. Patients with implantable cardiac devices were excluded. These methods of patient selection therefore did not reflect real-life conditions. As expected, a significant amount of the data had to be excluded due to insufficient signal quality. This is partly because the study used a trial version of the algorithm with a deactivated automated signal quality check to prevent bias. Even though this led to over-representation of insufficient signals, this relevant issue had to be addressed in the context of PPG signal-based analysis. In addition, the trial was designed to reproduce the sensitivity and specificity of 5-min PPG recordings and calculate the performance for 3 min and 1 min as a predefined secondary aim. To do so, 3- and 1-min fragments suitable for analysis were cut out of the 5-min files. To test this in a fully-independent manner, additional recordings with 1- and 3-min durations will be needed.

Smartphones in this trial were limited to iOS devices to reduce the number of factors that might influence the results. Based on real-life data, 4.28% of 105 664 PPG files recorded with iPhones were rejected due to insufficient signal quality, whereas 6.64% of 83 048 PPG files recorded with Android phones were rejected. This shows that there was only a small difference between the two operating systems and that the automated signal quality check reduces the amount of data with insufficient signal quality.

One might assess intense skin pigmentation as an additional limitation because PPG pulse wave tissue penetration is reduced in darker skin.21 This will be an important consideration for PPG recordings obtained by smart devices from the wrist, but palmar fingertips are not pigmented so we did not see a significant impact. Further, we observed no influence of calluses on signal-to-noise ratio, but their influence should be further investigated. Another limitation concerns patients with tremor, which can lead to a worse signal-to-noise ratio. These limitations were already addressed by the automated signal quality check in the commercial app version.

The PPG-based method was compared against a mobile iECG and not against the gold standard 12-lead ECG or Holter ECG. The researchers chose this design because the iECG is recommended in the EHRA guidelines for AF screening and has FDA approval for this purpose. It seemed of relevance to directly compare these two technologies. Nevertheless, an additional 12-lead ECG would have reduced the amount of data that had to be excluded.

A last limitation can be seen in the fact that recordings of PPG signals and iECG were not performed simultaneously, but successively. Although very unlikely, a spontaneous conversion of the heart rhythm between the recordings might lead to a false discordance of PPG and iECG diagnosis. This was considered negligible, as oppositional conversion was equally likely.

Conclusion

In conclusion, PPG based algorithms appear to be suitable for AF screening and should be tested in population-based, large-scale AF screening studies as a next step.

Supplementary material

Supplementary material is available at Europace online.

Acknowledgements

The authors would like to thank A. Winterhalder and U. Schmitt for their support.

Funding

This work was supported by Preventicus®, which provided the iPhones and software and performed the analysis of PPG data based on the Heartbeats algorithm. This work was further funded by a research grant of the University Hospital Basel.

Conflict of interest: Jens Eckstein, Marcus Dörr, Ralf Birkemeyer, and Stefan Weber hold 0.5% virtual shares in Preventicus®. Jens Eckstein received a travel grant from Preventicus®. All other authors declared no conflict of interest.

References


N. Brasier et al.