Smartphone-based Use of the Photoplethysmography Technique to Detect Atrial Fibrillation in Primary Care: A Diagnostic Accuracy Study of the FibriCheck App

Christophe Mortelmans, MD¹, Tine Proesmans, MSc², Ruth Van Haelst, MD¹, Frederik H. Verbrugge, MD, PhD², Pieter Vandervoort, MD², Bert Vaes, MD, PhD¹

¹Department of Public Health and Primary Care, University of Leuven (KU Leuven), Leuven, Belgium
²Department of Cardiology, Ziekenhuis Oost-Limburg, Genk, Belgium

Corresponding author: Vaes Bert
Kapucijnenvoer 33, Blok J, 3000 Leuven, Belgium
Tel: +32 16 337 468 Fax: +32 16 337 480
E-mail: bert.vaes@kuleuven.be
Abstract

Objectives
This study tested the diagnostic accuracy of the FibriCheck AF algorithm for the detection of atrial fibrillation (AF) based on smartphone photoplethysmography (PPG) and single-lead electrocardiography (ECG) signals.

Methods
A convenience sample of patients, aged 65 and older, with and without a known history of AF, was recruited from 17 primary care facilities. Patients with an active pacemaker rhythm were excluded. A PPG signal was obtained with the backside camera of an iPhone 5S. Simultaneously, a single-lead ECG was registered using a dermal patch with a wireless connection to the same smartphone. PPG and single-lead ECG signals were analysed using the FibriCheck AF algorithm. At the same time, a 12-lead ECG was obtained and interpreted off-line by independent cardiologists to determine the presence of AF.

Results
A total of 102/223 subjects (46%) were in AF. PPG signal quality was sufficient for analysis in 93%, and single-lead ECG quality was sufficient in 94% of the participants. After removing insufficient quality measurements, the sensitivity and specificity were 96% (95% CI 89-99%) and 97% (95% CI 91-99%) for the PPG signal versus 95% (95% CI 88-98%) and 97% (95% CI 91-99%) for the single-lead ECG, respectively. False-positive results were mainly due to premature ectopic beats. In 196 subjects where the signal quality of both techniques was adequate, PPG and single-lead ECG yielded a similar diagnosis in 192 subjects (98%).

Conclusions
The FibriCheck AF algorithm can accurately detect AF based on smartphone PPG and single-lead ECG signals in a primary care convenience sample.

Keywords
Atrial fibrillation; electrocardiography; photoplethysmography; smartphone; algorithm; diagnostic accuracy; FibriCheck
Introduction
Atrial fibrillation (AF) is the most common cardiac arrhythmia, affecting approximately 33.5 million people worldwide [1]. AF prevalence is estimated at 3% in adults aged >20 years, increasing in the elderly and patients with comorbid conditions, such as hypertension, heart failure, coronary artery disease, heart valve disease, obesity, diabetes, and chronic kidney disease [2]. Stroke remains the most fearsome complication of AF; its risk after a diagnosis of AF is multiplied five-fold [1]. Although effective anticoagulation therapy reduces this risk dramatically by 60%, initial AF episodes may frequently go undetected [3]. Indeed, contemporary studies on ischaemic stroke demonstrate that AF is regularly diagnosed during or immediately after an event [4]. Importantly, AF incidence is markedly influenced by the intensity of screening efforts [5]. Current guidelines of the European Society of Cardiology recommend opportunistic screening in people aged ≥65 years by pulse palpation and, if irregular, a 12-lead electrocardiogram (ECG) [6]. However, evolving technology may offer scalability to reach a general population at relatively low cost and with minimal logistic efforts, which may further lower the threshold for screening. Smartphones may offer an interesting modality to aid AF diagnosis, as their use has exponentially increased in recent years and is continuing to grow. By applying the photoplethysmography (PPG) technique through the smartphone camera, rhythm registration from the fingertip of a subject becomes a real possibility. Software has been developed to acquire PPG measurements with most common smartphones and to use these signals to analyse the heart rhythm. The aim of this study was to test the diagnostic accuracy of such an approach, using the FibriCheck smartphone app (Qompium, Hasselt, Belgium), in comparison with the gold standard method of AF diagnosis, the 12-lead ECG.
Methods

Study design
This diagnostic accuracy study was performed between October 2015 and March 2016 in 17 general practitioner (GP) centres in Belgium. Participating GPs were asked to invite patients with known, paroxysmal or persistent AF to participate in the study. By searching electronic medical records, patients aged ≥65 years with a diagnosis of AF were identified. This convenience sample was supplemented with subjects without a history of AF. The presence of an active pacemaker rhythm was an exclusion criterion, since this could impact the diagnostic results obtained during the subsequent measurements. With a probability of finding a false-positive result of 5% or less (α = 0.05), an estimated AF prevalence of 50% in the study population, an expected sensitivity and specificity of 95%, and a confidence interval of 4%, a sample size of 160 subjects was calculated. The study complies with the declaration of Helsinki and was approved by the ethical review board of the Medical Faculty of the KU Leuven, Belgium (no MP 05256). All study subjects provided written informed consent before participation. For all participants, researchers (CM, RVH) registered demographics, vital parameters, medication use and components of the CHA2DS2-VASc score to determine the stroke risk (i.e. congestive heart failure, hypertension, age, diabetes mellitus, prior stroke, vascular disease, and sex category).

Photoplethysmography (PPG) and the FibriCheck® app
In every subject, a smartphone-based assessment of the cardiac rhythm using the CE-approved FibriCheck app was performed by a single researcher (CM or RVH) who was not blinded for the medical history of the patient. For this purpose, a PPG signal was acquired with the backside camera of an iPhone 5S (Apple Inc., CA, USA). PPG is a technique whereby a volumetric measurement is optically obtained. A classic application of the PPG technique is the pulse oximeter, which illuminates the skin and measures changes in light intensity with blood volume pulse variation in the local arterioles, using this information to determine arterial oxygen saturation and pulse frequency. The same principle can be applied by using the camera of a smartphone and measuring the amount of reflected light. In this way, each heartbeat is recorded, and the rhythm can be determined based on the intervals between heartbeats (i.e. RR-intervals). The FibriCheck app provides software to obtain and
analyse such measurements with most common smartphones. To obtain a high-quality PPG signal, subjects were asked to adopt a sitting position with both arms resting on a table, holding the iPhone 5S in a vertical position with their right hand. Subsequently, they were asked to cover the flashlight and the backside camera horizontally with their left index finger (Figure 1). The measurement time to acquire the PPG signal with the FibriCheck app is 1 min, visualized by a countdown clock on the smartphone screen. To minimalize motion artefacts, subjects were instructed not to speak or move during the registration process. Subjects were asked to independently perform three consecutive measurements. To avoid evoking a reaction following the result of a measurement, researchers and participants were blinded for the PPG-signal during the measurements and the automated interpretations after the measurements. The researcher performing the measurements scored every study subject on a scale from 1 to 4 according to their experience with and handling of the smartphone (1, optimal handling; 2, subject has good knowledge of the smartphone and only requires minor input or corrections on how to perform handlings; 3, subject has some knowledge of the smartphone but needs substantial corrections on how to perform handlings; 4, subject has never held a smartphone before or has many issues in holding it correctly and performing the handlings).

Figure 1. Smartphone-based assessment of the cardiac rhythm using the FibriCheck® app. The ECG-bone, attached to a subject’s chest, for obtaining a single-lead electrocardiogram wirelessly connected to the smartphone by the FibriCheck® app.
**Single-lead electrocardiogram using the ECG-bone**

Simultaneously with the PPG measurement, a synchronized single-lead ECG was obtained using the ECG-bone (IMEC, Heverlee, Belgium) [7]. This module was attached with a patch on the left side of the subject’s chest above ribs 2-3 (Figure 1) and was wirelessly connected to the iPhone 5S with the help of the FibriCheck app. This procedure was performed by the same researcher who helped with the operation of the FibriCheck app.

**Data processing**

After simultaneous collection of both the PPG and single-lead ECG signal, data were transferred to a secured online data platform for analysis. First, raw signals were analysed by a recurrent neural network algorithm to classify them based on quality metrics. The PPG signal quality judgement was based on the capacity to detect and differentiate heartbeats. If heartbeat detection was compromised with noise, or if heartbeats were absent, these measurements were filtered out as insufficient quality. QRS-complexes in the single-lead ECG signals were detected using the Pan-Tompkins algorithm based on slope, amplitude and width analysis of the waveform [8]. The reliable measurements were evaluated by the FibriCheck AF algorithm based on RR-interval variability analysis (Figure 2).

![Figure 2. Synchronized photoplethysmography signal and single-lead electrocardiogram in a patient with sinus rhythm (left) and atrial fibrillation (right). Arrows indicate the RR-interval.](image)

AF, atrial fibrillation; ECG, electrocardiogram; PPG, photoplethysmography

**12-lead electrocardiogram**

The same researcher obtained a 12-lead electrocardiogram (gold standard). The ECGs were taken using digital machines (a CardiMax FCP-7101 (Fukuda Denshi, Tokyo, Japan), a CP 50 (Welch Allyn,
New York, USA), a Universal ECG (QRS Diagnostic, Plymouth MN, USA) and an ECG-1150 (Nihon Kohden Corporation, Tokyo, Japan)), and the data were immediately printed. All 12-lead ECGs were analysed off-line based on the Minnesota Code Classification System for Electrocardiographic Findings (code 8-3-1) by 2 experienced, independent cardiologists blinded to all other data. In case of disagreement, a third cardiologist was consulted to interpret the rhythm.

**Statistical analysis**
Continuous variables are expressed as the means ± standard deviations if normally distributed or otherwise by medians (interquartile ranges (IQR)). Categorical variables are expressed as percentages. A Mann-Whitney U test was used to compare smartphone handling between patients with versus without AF. The levels of diagnostic accuracy of the PPG and single-lead ECG signals analysed by the FibriCheck AF algorithm were tested against the gold standard using 2x2 tables (MedCalc Software, Mariakerke, Belgium). Data analysis was performed both on measurement level, including the results of all three measurements, and on participant level, using a majority rule to determine the overall result. For both approaches, data analysis was performed (i) after exclusion of insufficient quality measurements, (ii) with insufficient quality measurements categorized as ‘sinus rhythm’, and (iii) with insufficient quality measurements categorized as ‘possible AF’. If two insufficient quality measurements were present, the majority rule did not uphold and decision was made based on severity. The positive and negative predictive values (PPV and NPV) were also estimated based on an expected AF prevalence of 6% in the population aged ≥65 years [9]. Finally, the results of PPG versus single-lead ECG were compared case by case for inconsistencies, with beat-to-beat analysis of the raw data to reveal the underlying reasons for any differences.

**Results**

**Study population**
A total of 241 patients participated in the study. The study flowchart is presented in Figure 3. 18 pacemaker patients had to be excluded because of active pacing during the measurements. Therefore, the final study population consisted of 223 subjects. Their characteristics are presented in Table 1. Overall, the mean age was 77 ± 8 years (range: 59-95 years), with 104 (46.6%) male. AF was present in 102 patients (45.7%). Patients in AF had a mean CHA2DS2-VASc score of 5 ± 2. Smartphone
handling was significantly different between patients in AF (4 [IQR 3 – 4]) versus not (3 [IQR 2 – 4]; $P = 0.001$).

**Figure 3.** Study flowchart. AF, atrial fibrillation.

**Table 1.** Characteristics of the study population (n=223)

<table>
<thead>
<tr>
<th></th>
<th>Healthy patients (n = 79)</th>
<th>AF patients in sinus rhythm (n = 42)</th>
<th>AF patients in AF (n = 102)</th>
<th>Total population (n = 223)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean ± SD</td>
<td>75 ± 8</td>
<td>78 ± 8</td>
<td>79 ± 8</td>
<td>77 ± 8</td>
</tr>
<tr>
<td>Male gender</td>
<td>32 (41%)</td>
<td>21 (50%)</td>
<td>51 (50%)</td>
<td>104 (46.6%)</td>
</tr>
<tr>
<td>Resting heart rate</td>
<td>71 ± 14</td>
<td>70 ± 18</td>
<td>83 ± 20</td>
<td>77 ± 19</td>
</tr>
</tbody>
</table>
(bpm), mean ± SD

| Systolic BP (mmHG), mean ± SD | 130 ± 16 | 129 ± 14 | 129 ± 17 | 129 ± 16 |
| Diastolic BP (mmHG), mean ± SD | 73 ± 8 | 74 ± 7 | 74 ± 11 | 74 ± 9 |

**Risk factors**

| Congestive heart failure | 12 (15%) | 10 (24%) | 42 (41%) | 64 (29%) |
| Diabetes | 9 (11%) | 9 (21%) | 27 (26%) | 45 (20%) |
| Stroke or transient ischaemic attack | 9 (11%) | 9 (21%) | 32 (31%) | 50 (22%) |
| Atherosclerotic disease | 19 (24%) | 22 (52%) | 48 (47%) | 89 (40%) |

**Medication use**

| Anticoagulation | 2 (3%) | 30 (71%) | 92 (90%) | 124 (55.6%) |
| ACE inhibitor | 11 (14%) | 13 (31%) | 32 (31%) | 56 (25%) |
| Angiotensin receptor blocker | 11 (14%) | 10 (24%) | 21 (21%) | 42 (19%) |
| Beta blocker | 33 (42%) | 22 (52%) | 71 (70%) | 126 (56.5%) |
| Diuretics | 16 (20%) | 11 (26%) | 53 (52%) | 80 (36%) |
Photoplethysmography measurements

On participant level
PPG measurements were recorded for a total of 223 participants. After exclusion of measurements of insufficient quality (n = 16; 7%), a PPG signal suitable for analysis was obtained for 207/223 subjects (92.8%). Positive results were found in 91 subjects and negative results in 116 subjects. PPG results matched the diagnosis made by cardiologists based on the 12-lead ECG in 199/207 subjects (96.1%), resulting in an overall sensitivity of 95.6% (95% CI 89.1-98.8%) and a specificity of 96.6% (95% CI 91.4-99.1%) (Table 2). From the 8 inconsistent results, 4 were false-positive and 4 false-negative. False-positive results were caused by atrial premature beats (n = 4). False-negative results were caused by peak wave under-sensing (n = 1) and misinterpretation of an atrial flutter as sinus rhythm (n = 3).

Based on an expected prevalence of 6% in the population aged ≥ 65 years, a PPV of 63% (95% CI 61.3-64.8%) and an NPV of 99.7% (95% CI 99.6-99.8%) were estimated.

Using the same approach, but classifying insufficient quality measurements as ‘sinus rhythm’, a sensitivity of 87% (95% CI 78.80-92.89%) and a specificity of 96.75% (95% CI 91.88-99.11%) were obtained (Table 2). In this scenario, PPG results matched the cardiologist interpretation of the 12-lead ECG in 206/223 subjects (92.4%). The amount of false-negatives in this scenario increased to 13. Classifying insufficient quality measurements as ‘possible AF’ yielded a sensitivity of 96% (95% CI 90.07-98.90%) and a specificity of 91.06% (95% CI 84.56-95.45%) (Table 2). Here, PPG results

<table>
<thead>
<tr>
<th>Smartphone handling</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Smartphone ownership</td>
<td>19 (24%)</td>
<td>6 (14%)</td>
<td>11 (11%)</td>
<td>36 (16%)</td>
</tr>
</tbody>
</table>

AF, atrial fibrillation; SD, standard deviation; bpm, beats per minute; BP, blood pressure; IQR, interquartile range; ACE, angiotensin-converting enzyme
matched the diagnosis of 12-lead ECG in 208/223 subjects (93.2%) and the number of false-positives increased to 11.

Table 2. Diagnostic accuracy of PPG and single-lead ECG signal analysis on participant level, based on a majority rule, compared to the reference gold standard, 12-lead ECG.

<table>
<thead>
<tr>
<th>Diagnostic test</th>
<th>Insufficient quality excluded</th>
<th>Insufficient quality categorized as ‘sinus rhythm’</th>
<th>Insufficient quality categorized as ‘possible AF’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PPG (n = 207)</td>
<td>ECG (n = 210)</td>
<td>PPG (n = 223)</td>
</tr>
<tr>
<td>Prevalence</td>
<td>43.96%</td>
<td>45.02%</td>
<td>44.84%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>95.60%</td>
<td>94.74%</td>
<td>87.00%</td>
</tr>
<tr>
<td>Specificity</td>
<td>96.55%</td>
<td>96.55%</td>
<td>96.75%</td>
</tr>
<tr>
<td>PPV</td>
<td>95.60%</td>
<td>95.74%</td>
<td>95.60%</td>
</tr>
<tr>
<td>NPV</td>
<td>96.55%</td>
<td>95.73%</td>
<td>90.15%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>96.14%</td>
<td>95.73%</td>
<td>92.38%</td>
</tr>
</tbody>
</table>

AF, atrial fibrillation; PPG, photoplethysmography; ECG, electrocardiogram; PPV, positive predictive value; NPV, negative predictive value

On measurement level
A total of 657 PPG measurements were recorded, 110 (16.7%) of which were labelled as ‘insufficient quality’ by the algorithm quality filter. Analysing solely high quality PPG measurements resulted in a sensitivity of 95.28% (95% CI 91.71-97.62%) and a specificity of 96.18% (95% CI 93.42-98.01%) (Table 3). For 524/547 PPG measurements (95.8%), the diagnosis matched the diagnosis based on 12-lead ECG. The 23 inconsistent results were caused by 12 false-positives and 11 false-negatives. When categorizing insufficient quality as ‘sinus rhythm’, the sensitivity dropped to 76.03% (95% CI 70.71-80.81%) with a specificity of 96.71% (95% CI 94.33-98.29%) (Table 3). This resulted in an agreement between PPG and 12-lead ECG for 575/657 measurements (87.5%) and an increase of false-negatives to 70. Interpreting insufficient quality as ‘possible AF’ resulted in a sensitivity and specificity of
96.23% (95% CI 93.36-98.10%) and 82.74% (95% CI 78.46-86.47%), respectively (Table 3). 583/657 PPG measurements (88.7%) had the same diagnosis compared to 12-lead ECG. 11 measurements were false-negative and 63 measurements were false-positive.

Table 3. Diagnostic accuracy of PPG and single-lead ECG signal analysis on measurement level compared to reference gold standard, 12-lead ECG.

<table>
<thead>
<tr>
<th>Diagnostic test</th>
<th>Insufficient quality excluded</th>
<th>Insufficient quality categorized as ‘sinus rhythm’</th>
<th>Insufficient quality categorized as ‘possible AF’</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPG (n = 547)</td>
<td>PPG (n = 657)</td>
<td>PPG (n = 657)</td>
<td>ECG (n = 584)</td>
</tr>
<tr>
<td>Prevalence</td>
<td>42.60%</td>
<td>44.44%</td>
<td>44.44%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>95.28%</td>
<td>76.03%</td>
<td>96.23%</td>
</tr>
<tr>
<td>Specificity</td>
<td>96.18%</td>
<td>96.71%</td>
<td>82.74%</td>
</tr>
<tr>
<td>PPV</td>
<td>94.87%</td>
<td>94.87%</td>
<td>81.69%</td>
</tr>
<tr>
<td>NPV</td>
<td>96.49%</td>
<td>83.45%</td>
<td>96.49%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>95.80%</td>
<td>87.52%</td>
<td>88.74%</td>
</tr>
</tbody>
</table>

AF, atrial fibrillation; PPG, photoplethysmography; ECG, electrocardiogram; PPV, positive predictive value; NPV, negative predictive value

**Insufficient quality**
The Chi-squared statistic was used to identify causes or correlations between comorbidities and insufficient PPG-measurements.

<table>
<thead>
<tr>
<th>Comorbidity</th>
<th>P-value on measurement level</th>
<th>P-value on subject level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td>P = 0.18</td>
<td>P = 0.15</td>
</tr>
<tr>
<td>Heart failure</td>
<td>P = 0.32</td>
<td>P = 0.73</td>
</tr>
<tr>
<td>Gender</td>
<td>P = 0.02</td>
<td>P = 0.44</td>
</tr>
<tr>
<td>BMI (&gt; 25)</td>
<td>P = 0.02</td>
<td>P = 0.41</td>
</tr>
<tr>
<td>Age (&gt; 75y)</td>
<td>P = 0.06</td>
<td>P = 0.58</td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Vascular disease</td>
<td>P &lt; 0.001</td>
<td>P = 0.86</td>
</tr>
</tbody>
</table>

Additionally, based on the Chi-square statistic, there is no association between smartphone handling and insufficient quality (P = 0.4324).

**Single-lead ECG by the ECG-bone**

*On participant level*

Single-lead ECG recordings were collected from a total of 223 participants. After eliminating insufficient quality measurements, a single-lead ECG signal suitable for analysis was obtained for 210/223 subjects (94.2%). Positive results were found in 90 subjects and negative results in 111 subjects. Single-lead ECG results matched the diagnosis made by cardiologists based on the 12-lead ECG in 201/210 cases (95.7%), yielding a sensitivity and specificity of 94.74% (95% CI 88.14-98.27%) and 96.55% (95% CI 91.33-99.04%), respectively (Table 2). From the 9 inconsistent results, 4 were false-positive and 5 false-negative. False-positive results were caused by atrial (n = 3) and ventricular (n = 1) premature beats. False-negative results were caused by misinterpretation of an atrial flutter as sinus rhythm (n = 5).

Including the insufficient quality measurements as ‘sinus rhythm’ resulted in a sensitivity of 90% (95% CI 82.38-95.10%) and a specificity of 96.75% (95% CI 91.88-99.11%), while including these measurements as ‘possible AF’ resulted in a sensitivity of 95% (95% CI 88.72-98.36%) and a specificity of 91.06% (95% CI 83.58-94.86%) (Table 2). In the first scenario, the amount of false-negatives increased to 10. In the latter, the amount of false-positives increased to 12.

*On measurements level*

A total of 657 single-lead ECG measurements were recorded, of which 45 (7%) were identified as insufficient quality. Analysis of solely high quality measurements yielded a sensitivity and specificity of 91.97% (88.10-94.90%) and 96.45% (95% CI 93.88-98.15%), respectively (Table 3). There was an agreement between the diagnosis based on single- and 12-lead ECG for 578/612 measurements (94.4%). 12 measurements were false-positive and 22 measurements false-negative. Categorizing the insufficient quality measurements as ‘sinus rhythm’ resulted in 86.60% (95% CI 82.14-90.29%)
sensitivity and 96.72% (95% CI 94.34-98.29%) specificity (Table 3). The diagnosis based on single-lead ECG matched the diagnosis based on 12-lead ECG for 606/657 measurements (92.2%). The amount of false-negative measurements increased to 39. Interpreting insufficient quality as ‘possible AF’ resulted in a sensitivity of 92.44% (95% CI 88.87-95.20%) and a specificity of 89.07% (95% CI 85.42-92.08%) (Table 3). Here, there was an agreement for 595/657 measurements (90.6%). The amount of false-positives increased to 40.

Consistency between photoplethysmography and the single-lead ECG signals
In 196/223 subjects (87.9%), the quality of both the PPG and single-lead ECG signal were reliable for analysis. Both signals resulted in similar diagnoses in 192/196 subjects (98.0%). On measurement level, 516/656 (78.7%) PPG and single-lead ECG paired measurements had a sufficient quality of reliable analysis. This resulted in similar diagnosis in 506/516 (98.1%) measurements.
Discussion

Principal findings
This diagnostic accuracy study in a primary care convenience sample revealed that cardiac rhythm analysis through a smartphone-based PPG signal with the FibriCheck AF algorithm had very good sensitivity and specificity to detect AF. False-positive results were mainly due to the presence of extrasystoles. Furthermore, the FibriCheck AF algorithm accurately diagnosed AF based on a single-lead ECG, with a similar sensitivity and specificity compared to the PPG signal. Both sensitivity and specificity were affected when including insufficient quality measurements as either ‘sinus rhythm’ or ‘possible AF’, leading to a decrease in accuracy from 96.14% to 92.38% and 93.27%, respectively. Beat-to-beat analysis showed a strong agreement between the PPG and the single-lead ECG signal.

The diagnostic accuracy of the FibriCheck AF algorithm was comparable to other screening methods and devices. A recent systematic review and meta-analysis found the greatest accuracy for blood pressure monitors and non-12-lead ECGs [10]. The modified sphygmomanometers had a pooled sensitivity of 98% and a specificity of 92%. Non-12-lead ECGs scored a sensitivity of 91% and a specificity of 95%. However, when focusing on the primary care setting, a lower specificity of 89% was obtained. Smartphone apps also showed a good pooled accuracy, with 97% sensitivity and 95% specificity. The AliveCor showed a sensitivity of 94% and a specificity of 99% in cardiology clinic patients [11], but showed a low sensitivity ranging from 55-79% and a specificity between 97.5-97.9% in hospitalized patients [12]. This was later contributed to several defects which impaired diagnostic accuracy and necessitated a product recall in the United States during the course of the study. The commercial algorithm has been biased for enhanced specificity, while the version of the AF detection algorithm used in published screening studies was biased for enhanced sensitivity. The defects, together with the enhanced specificity biasing, resulted in the reported low sensitivity [13].

A smartphone app is quick, inexpensive and practical without the need for special infrastructure or external hardware. The patient does not require any experience or medical education and can be easily trained to use the app. Physicians can remotely review the transferred data, which enables optimal patient follow-up in a less time-consuming manner. Furthermore, the high accessibility of smartphone
apps and the increasing smartphone usage among the elderly are important assets [14, 15]. However, only 17% of our study population owned a smartphone compared with the 27% reported in recent Austrian [13] and American [15] senior surveys. Recent Belgian and Dutch surveys reported smartphone use in 54% of the 65-75 population and 29% in the 75+ population [16]. Furthermore, a relatively high difficulty in smartphone handling was observed (Table 1). It is however expected that, together with AF prevalence, the smartphone usage in the senior population will continue to rise and the lack of familiarity will partially fade. Moreover, a recent study demonstrated an increasing willingness and capacity to use mobile health devices by older persons [17].

The current phase II diagnostic study demonstrates that great opportunities lie in AF screening through PPG measurements. However, the place of the FibriCheck app in future screening or case-finding programmes for AF remains to be determined. The FibriCheck app could be a good candidate for implementation as a case-finding or event-recording solution for paroxysmal AF in high-risk patients in primary care, or patients with paroxysmal palpitations without a clear diagnosis. Furthermore, this mobile technology also allows follow-up of patients after resynchronization or ablation. Indeed, it has been demonstrated that intermittent measurements over a longer time period, as made possible by a smartphone app, has a great chance of increasing the diagnostic yield.

Further research, such as validation studies and cluster randomized trials, is needed to investigate the effects of these implementation strategies and the performance in a population with a lower incidence of AF.

**Limitations**

This study is the first investigating the diagnostic accuracy of the FibriCheck app in a realistic primary care population. The simultaneous measurement of PPG and single-lead ECG offered the opportunity for beat-to-beat comparisons of the two measurement methods to reveal the underlying reasons for inconsistencies in diagnosis using the FibriCheck AF algorithm. However, a few limitations should be noted. First, different digital 12-lead ECG devices were used as the reference standard instead of one standardized device. Second, there was a gap of a few minutes between the simultaneous PPG and single-lead ECG measurements and the subsequent 12-lead ECG measurement, and the subject’s heart
rhythm might have changed in that short time period. Third, to calculate the PPV and the NPV in a population over 65 years old, we assumed an AF prevalence of 6%. However, due to the heterogeneity between conducted studies, various values were found for AF prevalence in the literature [9, 18]. Fourth, because the study population was a convenience sample, extrapolation of these results to the general population should be made with caution. In addition, all measurements were performed under medical supervision. Although participants and researchers were blinded for all notifications and results, and were thereby prevented to attempt to improve the measurement results, it remains unclear whether such apps would achieve the same accuracy in an unsupervised (real-world) situation. Another important aspect that should be considered is the accuracy of the algorithm to screen patients who may have uncontrolled high heart rates. Since this study was positioned as a validation study and not as a screening study, further research to assess real-life accuracy is warranted. Finally, some false-positive results with the FibriCheck AF algorithm were caused by atrial or ventricular extrasystoles, which is a known issue in AF screening using RR-interval variability analysis. However, since confirmation with 12-lead ECG or single-lead ECG documenting P-waves is required and recommended by several guidelines [7, 19, 20], this limitation does not jeopardize the potential of FibriCheck as screening tool.

**Conclusions**

To conclude, the FibriCheck app is an accessible standalone smartphone app that showed promising results for AF detection in a primary care convenience sample. The FibriCheck AF algorithm showed a very good sensitivity and specificity. These findings confirm the FibriCheck app to be a possible candidate to implement in future screening, case-finding programmes for AF or monitoring programmes in a home-setting. However, further research is needed to determine the place of the FibriCheck app in such a strategy.
Acknowledgments
Qompium (Hasselt, Belgium) provided the smartphone and free use of the FibriCheck® app. IMEC (Leuven, Heverlee, Belgium) offered the ECG-bone device without cost. Both companies had the opportunity to check the final version of the manuscript and to make recommendations, but were not involved in the data collection, analysis or decision to submit the report for publication.

Conflicts of interest
None declared.
References

13. Albert DE. Performance of hand-held electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology. 2017;19(8):1408. PMID: 27789561
17. Seifert A, Christen M, Martin M. Willingness of Older Adults to Share Mobile Health Data with Researchers2018. DOI: 10.1024/1662-9647/a000181


List of abbreviations
AF Atrial fibrillation
ECG Electrocardiogram
PPG Photoplethysmography
CE Conformité Européenne
IQR Interquartile range
PPV Positive predictive value
NPV Negative predictive value
CI Confidence interval