Screening for Atrial Fibrillation Using a Smartphone: Is There an App for That?

Ben Freedman, MBBS, PhD

Atrial fibrillation (AF) is a growing problem in cardiovascular disease and is associated with an increased risk of severe stroke, heart failure, and death. Indeed, recent stroke registries indicate that AF is associated with a third of all ischemic strokes. Anticoagulation with warfarin or the non-vitamin K–dependent oral anticoagulants is extremely effective in reducing stroke and mortality, but between a quarter and a third of all AF-related strokes occur in patients with stroke as the first manifestation of AF. Because unknown asymptomatic AF is common, occurring in 1.4% of those aged 65 or older on a single time point check for presence of AF, 3% with patient-activated recordings over 2 weeks, and even more if long-term continuous recordings are used, it is intuitive that screening for AF and subsequent anticoagulant treatment should reduce the stroke burden related to AF. This is the basis of guideline recommendations to screen for AF in those aged ≥65.

Guidelines for screening for AF are based on a large-cluster randomized controlled trial in general practice, of opportunistic pulse taking versus systematic screening with 12-lead ECG or standard care. This study found that both opportunistic pulse taking and systematic screening detected similar numbers of unknown AF, and more than conventional care. But opportunistic pulse taking was less expensive than systematic 12-lead electrocardiography, so opportunistic pulse taking followed by a 12-lead ECG is now recommended as being cost-effective. Unfortunately, pulse taking is not commonly performed in general practice, and if the heart rate is recorded during a visit, it is usually performed by an automated sphygmomanometer rather than by auscultation over the brachial artery, so an irregular pulse will go undetected.

The advent of small devices that can diagnose AF automatically, based on either pulse irregularity or rhythm analysis of a single-lead handheld ECG, could change the cost-benefit equation in favor of a more systematic approach to screening for AF either in the clinic or in the community. In this context it is interesting to see the contribution of Chan et al in this issue of the Journal of the American Heart Association evaluating the diagnostic performance of a smartphone app. The app uses the smartphone flash/lightsource and camera to obtain a photoplethysmographic (PPG) recording of the pulse wave. This is the principle behind a number of popular mobile devices and apps used to measure heart rate, primarily for exercise and fitness.

The additional step in this case is an algorithm to analyze the regularity of the pulse waves and diagnose AF or normal rhythm based on it. Similar algorithms are used in oscillometric sphygmomanometers, which produce a pulse wave signal and diagnose AF from the pattern using a similar algorithm. Even with the devices that use a hand-held ECG to produce a single-lead (lead I) rhythm strip, regularity of the RR intervals is one of the features that algorithms use to diagnose AF. ECG devices do have the advantage of also being able to examine P waves as well as QRS morphology to aid in discrimination from sinus rhythm with ventricular or atrial premature contractions, which also produce an irregular pulse. The QRS complex may also be a more precise signal than the pulse wave for rhythm algorithms. Unfortunately all of these devices are subject to noise and artifacts produced by movement of hands, fingers, and arms or electrical noise/interference from the smartphone or adjacent appliances. Some algorithms do attempt to compensate, but a noise-free signal is crucial to provide optimal diagnostic accuracy, irrespective of the device or technology used. This is particularly important when devices are used by people with minimal or no training.

Use of a smartphone PPG to diagnose AF is not new. It was first reported in 2012 and 2013 by a Massachusetts group in
Atrial Fibrillation Screening Using a Smartphone

Follow-up study in 2016 used the same data plus some cardioversion. The algorithm calculated root mean square of algorithm required that 2 out of 3 17-second pulse wave patterns in 17.1 seconds of collected PPG pulse waveforms. The study population included just over 1000 ambulatory patients with AF: readings were taken before and after cardioversion patients, and a relatively small number of additional patients selected on the basis of rhythm, pre- and postcardioversion patients, and a relatively small number of additional patients with diabetes, a more relevant population for screening than the general outpatient clinic aged ≥65 with hypertension and/or diabetes, a more relevant population for screening than the pre- and postcardioversion patients, and a relatively small number of additional patients selected on the basis of rhythm, used in the previous PPG studies. The authors of the current study also used the AliveCor smartphone-based handheld ECG as both a diagnostic standard (cardiologist diagnosis) and a comparator using the AliveCor AF smartphone.

The results indicated a sensitivity of 92.9% and specificity of 97.7% for AF diagnosis, which would be adequate for using the smartphone PPG app for screening. Given the widespread availability of smartphones, the authors conclude that the app would have the potential to enable widespread screening for AF in the community. Of course there are a number of caveats. Crucially, the study was supervised by trained personnel, who directed the patients to ensure high-quality PPG and ECG recordings. This would not necessarily be the case if the app were used widely in the community. One of the main issues causing false-positive Cardiio Rhythm results in the study, was movement artifact. Artifact is likely to be a much greater problem if large numbers of unsupervised recordings are made at home or work in each subject. Moreover, the final diagnosis in the study was checked by a near-simultaneous handheld ECG followed by a 12-lead ECG if required, but neither would be available if a pulse-based PPG were used for AF screening. Transmission of the PPG waveform would not really help confirm an AF diagnosis, unlike ECG-based devices, which can all store and transmit a diagnostic quality ECG rhythm strip, as used by the expert readers in the current study. Although motion and noise are also problems for algorithms in handheld ECG devices, part of the collected ECG is often analyzable by trained ECG readers who are able to view the transmitted waveforms.

The Chan et al study used ambulatory patients from a general outpatient clinic aged ≥65 with hypertension and/or diabetes, a more relevant population for screening than the pre- and postcardioversion patients, and a relatively small number of additional patients selected on the basis of rhythm, used in the previous PPG studies. The authors of the current study also used the AliveCor smartphone-based handheld ECG as both a diagnostic standard (cardiologist diagnosis) and a comparator using the AliveCor AF smartphone.

Figure. Other devices to screen for atrial fibrillation. Top left, MyDiagnostick, which is a bar activated when held by left and right hands (lead 1 ECG). After recording for 1 minute, a green light indicates normal rhythm, a red light AF. The device can be connected by a cable to a PC to view the recordings, which have a date and time stamp only. Top right, Zenicor handheld ECG. Thumbs of each hand placed on the gray button electrodes on each side for 30 seconds generate a lead 1 ECG. There is no diagnosis or ECG trace on the device, which transmits wirelessly via the mobile network to a central data base with arrhythmia detection of ECG traces, viewable by the physician. Bottom left, Microlife sphygmomanometer. This is a standard oscillometric BP device that uses pulse wave irregularity to diagnose AF. An immediate diagnosis is given, but AF diagnosis on 2 of 3 readings is required before a diagnosis is confirmed. Bottom center and right, AliveCor (now Kardia) miniaturized ECG device with 2 silver electrodes held by fingers of right and left hands. It is usually stuck to the back of an android or iOS smartphone. Recording is activated by holding the device, which transmits to the smartphone via the microphone and app. The ECG is displayed on the smartphone, with automated AF diagnosis after 30 seconds of recording, and also transmitted to a central website or accessed by or emailed to the physician.

The Chan et al study used ambulatory patients from a general outpatient clinic aged ≥65 with hypertension and/or diabetes, a more relevant population for screening than the pre- and postcardioversion patients, and a relatively small number of additional patients selected on the basis of rhythm, used in the previous PPG studies. The authors of the current study also used the AliveCor smartphone-based handheld ECG as both a diagnostic standard (cardiologist diagnosis) and a comparator using the AliveCor AF smartphone.

76 patients with AF: readings were taken before and after cardioversion. The algorithm calculated root mean square of successive pulse wave differences and Shannon entropy. A follow-up study in 2016 used the same data plus some additional patients with ectopic beats and addition of a Poincaré plot adaptation to differentiate premature contractions, which pose problems for purely pulse-based methods. They used 2 minutes of PPG recording and a post-hoc algorithm analysis and reported high sensitivity and specificity for AF detection.

In this journal the smartphone “Cardiio Rhythm” app was used to diagnose AF. The algorithm in the app utilized a support vector machine term to analyze lack of repeating patterns in 17.1 seconds of collected PPG pulse waveforms. The study population included just over 1000 ambulatory outpatients with testing performed in real time in a clinic. The algorithm required that 2 out of 3 17-second pulse wave recordings were coded as AF before an AF diagnosis had been made, so the total time required was at least 1 minute of PPG recording (a similar requirement for 2 of 3 readings to show AF is used in the oscillometric BP devices). The trained observers who supervised the recording in the study ensured that both the PPG recording and the sequentially recorded hand-held single-lead AliveCor ECG were of high quality. Those with an abnormality detected on either the PPG Cardiio Rhythm app or the ECG AliveCor app had a 12-lead ECG recorded within 15 minutes, which was the standard for diagnosis of AF. Two cardiologists blinded to the data read the 30-second ECG rhythm strips to provide the diagnosis standard for the remainder. Not having a 12-lead ECG available for every PPG result is a potential study weakness, but those with no 12-lead ECG required for diagnosis were negative for AF by both PPG and single-lead ECG algorithms. Their ECG rhythm strips are therefore likely to be relatively easily diagnosed as sinus rhythm by cardiologists.

The Chan et al study used ambulatory patients from a general outpatient clinic aged ≥65 with hypertension and/or diabetes, a more relevant population for screening than the pre- and postcardioversion patients, and a relatively small number of additional patients selected on the basis of rhythm, used in the previous PPG studies. The authors of the current study also used the AliveCor smartphone-based handheld ECG as both a diagnostic standard (cardiologist diagnosis) and a comparator using the AliveCor AF smartphone.

DOI: 10.1161/JAHA.116.004000

Journal of the American Heart Association

2
Atrial Fibrillation Screening Using a Smartphone

Freedman

the initial clinic study and then tested it in community use (sensitivity 98% in initial studies that validated an AF detection algorithm in the clinic, 10 91.4% community15 in the original AliveCor ECG studies, and 97.9% in the Desteghe et al study16). The latter yielded a lower sensitivity but higher specificity than the PPG app in this study. In fact, the sensitivity of 71.4% was substantially lower than that in the initial studies that validated an AF detection algorithm in the clinic and then tested it in community use (sensitivity 98% in the initial clinic study10 and 98.5% in the larger community study conducted in pharmacies15) (Table). However, a similar lower sensitivity of 78.9% was reported in a recent study in hospitalized geriatric ward patients after exclusion of patients with implanted pacemaker devices, and even lower in cardiology ward inpatients.16 In the current study the Cardio Rhythm PPG specificity was 97.7%, and the contemporaneous AliveCor ECG specificity was very high (99.4%, Table). Specificity was lower in previous AliveCor studies (97% clinic, 10 91.4% community15 in the original AliveCor ECG studies, and 97.9% in the Desteghe et al study16). The principal reason for the difference in sensitivity between the original and more recent AliveCor studies is an intentional algorithm change down the receiver operator curve by the company favoring specificity over sensitivity.17 The rationale for this change is that the device is sold primarily to the public, who may not seek a physician or specialist overread. Additionally, for the Desteghe et al study,16 the AliveCor algorithm version contained a bug that necessitated a product recall in the United States during the course of the study but was not announced in Belgium, where the device was neither sold nor approved.17

Table. Summary of Sensitivity and Specificity of CardiioRhythm PPG, AliveCor ECG, and MyDiagnostick ECG

<table>
<thead>
<tr>
<th>Study, Device/Algorithm, Population</th>
<th>n</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chan et al,13 Cardiio Rhythm PPG, outpatient</td>
<td>1013</td>
<td>92.9%</td>
<td>97.7%</td>
</tr>
<tr>
<td>Chan et al,13 AliveCor ECG, outpatient</td>
<td>1013</td>
<td>71.4%</td>
<td>99.4%</td>
</tr>
<tr>
<td>Lau et al,10 AliveCor ECG, outpatient and inpatient</td>
<td>204</td>
<td>98%</td>
<td>97%</td>
</tr>
<tr>
<td>Lowres et al,15 AliveCor ECG, community pharmacy</td>
<td>1000</td>
<td>98.5%</td>
<td>91.4%</td>
</tr>
<tr>
<td>Desteghe et al,16 AliveCor ECG, inpatient geriatric</td>
<td>113</td>
<td>78.9%</td>
<td>97.9%</td>
</tr>
<tr>
<td>Desteghe et al,16 MyDiagnostick ECG, inpatient geriatric</td>
<td>113</td>
<td>89.5%</td>
<td>95.7%</td>
</tr>
</tbody>
</table>

algorithm. The latter yielded a lower sensitivity but higher specificity than the PPG app in this study. In fact, the sensitivity of 71.4% was substantially lower than that in the initial studies that validated an AF detection algorithm in the clinic and then tested it in community use (sensitivity 98% in the initial clinic study10 and 98.5% in the larger community study conducted in pharmacies15) (Table). However, a similar lower sensitivity of 78.9% was reported in a recent study in hospitalized geriatric ward patients after exclusion of patients with implanted pacemaker devices, and even lower in cardiology ward inpatients.16 In the current study the Cardio Rhythm PPG specificity was 97.7%, and the contemporaneous AliveCor ECG specificity was very high (99.4%, Table). Specificity was lower in previous AliveCor studies (97% clinic, 10 91.4% community15 in the original AliveCor ECG studies, and 97.9% in the Desteghe et al study16). The principal reason for the difference in sensitivity between the original and more recent AliveCor studies is an intentional algorithm change down the receiver operator curve by the company favoring specificity over sensitivity.17 The rationale for this change is that the device is sold primarily to the public, who may not seek a physician or specialist overread. Additionally, for the Desteghe et al study,16 the AliveCor algorithm version contained a bug that necessitated a product recall in the United States during the course of the study but was not announced in Belgium, where the device was neither sold nor approved.17

There is always a trade-off between sensitivity and specificity in setting algorithm diagnostic criteria. For physician or health practitioner use in screening for AF, very high sensitivity is required to pick up every person with unknown AF. This will of necessity reduce specificity. For an ECG-based device with an immediate high-quality ECG image, instant verification is possible, whereas for a PPG device, verification requires a sequential ECG. For health professional use of ECG-based devices for screening, higher sensitivity will be required than documented in the current study13 for the AliveCor algorithm. This would necessitate reversion to the original algorithm with very high sensitivity but lower specificity. Essentially a health professional screening version of the AliveCor app is required for this device if it is to be used for AF screening.

For screening in the community or for personal use, a drop in sensitivity may be a necessary trade-off to achieve very high specificity. This is particularly important if large numbers of people are screened or when multiple recordings are requested in each person, as might be the case with a PPG app. In this situation even small reductions in specificity could lead to very large numbers of false-positive results, which would require verification with a separate ECG.

Ultimately for the diagnosis of AF, an ECG will be required. This is the reason for the class 1 recommendation in guidelines: “Electrocardiographic documentation is recommended to establish the diagnosis of AF.” Therefore, AF screening workflow using any automated pulse-detection method, be it smartphone PPG or BP oscillometry, needs to factor in a mechanism for ECG verification. Verification will take time, involve logistic difficulties, and of course cost more. If hand-held ECGs are used to screen for AF, the cost of expert verification of the ECG signal must also be factored in, although having an ECG trace available immediately is an advantage. The Chan et al study13 examined the net reclassification index (NRI), of the Cardio Rhythm PPG algorithm over the AliveCor ECG algorithm, which would be appropriate only if the PPG were used as a second test following a handheld ECG test, which is highly unlikely and not the sequence of testing used in the study itself.

The best device for the purpose of AF detection for screening or diagnosis will need to be individualized to the situation and the healthcare system. It will be important to optimize sensitivity for AF detection while preserving specificity, to reduce the workload and expense caused by false-positive readings, and better noise detection and cancellation algorithms to reduce false positives. Adequate instruction in use of devices or apps is crucial to optimal performance. There cannot be a “one size fits all” solution for AF screening, and it is useful to have an addition to the diagnostic armamentarium, provided all are aware of the pitfalls as well as the advantages of the technique. Importantly, when deciding on devices for performing mass screening for AF, caveat emptor.

Disclosures

B.F. received research grants to conduct investigator-initiated studies from BMS/Pfizer, Bayer Pharma, and Boehringer-
Ingelheim and was a consultant for Bayer Pharma, BMS/Pfizer, Boehringer-Ingelheim, Servier, Astra-Zeneca, and Gilead, and speaker for Bayer Pharma, BMS/Pfizer, AstraZeneca. AliveCor provided ECG cases for screening studies, including some before the devices had been approved in any jurisdiction, but no funding.

References


Key Words: Editorials • atrial fibrillation • photoplethysmography • screening • smartphone
Screening for Atrial Fibrillation Using a Smartphone: Is There an App for That?
Ben Freedman

*J Am Heart Assoc.* 2016;5:e004000; originally published July 21, 2016;
doi: 10.1161/JAHA.116.004000
The *Journal of the American Heart Association* is published by the American Heart Association, 7272 Greenville Avenue,
Dallas, TX 75231
Online ISSN: 2047-9980

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://jaha.ahajournals.org/content/5/7/e004000