Consumer-grade wearable cardiac monitors: What they do well, and what needs work

ABSTRACT

Consumer-grade smart devices, including smartwatches and smartphones, are potentially valuable tools in detecting cardiac arrhythmias, particularly atrial fibrillation, and their use is increasing. These devices, which use photoplethysmography, show remarkably high sensitivity and specificity for detection of atrial fibrillation, with implications for stroke prevention and management in at-risk patients. The ability of the devices to detect atrial fibrillation is being compared with single-lead electrocardiography. Physicians will increasingly be asked to interpret data from these nonmedical-grade devices as they become more common. Limitations include high false-positive rates in certain populations and disparities in access.

KEY POINTS

Familiarity with the available devices and the data they generate will enhance patient care.

Many consumer devices have been validated against gold-standard medical-grade devices and have shown high sensitivity and specificity for heart rate and detection of atrial fibrillation.

There is a large gap between consumer-grade and medical-grade devices for detecting more complex arrhythmias.

Technological advances in consumer-grade wearable devices have increased the opportunity to diagnose and manage cardiac arrhythmias, especially atrial fibrillation. Devices that provide remote and long-term cardiac monitoring, such as smartphones, smartwatches, and handheld electrocardiography (ECG) devices, allow us to monitor high-risk patients outside the hospital.

As consumer wearables become more user-friendly, less costly, and more widely available, patients will expect physicians to be familiar with data generated from their devices. Therefore, knowledge of the available devices and their reliability compared with medical-grade devices will become increasingly important.

This article reviews common consumer-grade wearables, their accuracy compared with standard medical-grade devices, and our approach to patients with rate or rhythm abnormalities identified on at-home monitoring.

ATRIAL FIBRILLATION: A SIGNIFICANT RISK FACTOR

By 2030, an estimated 2.6 million people in the United States will have atrial fibrillation. Often asymptomatic, atrial fibrillation may remain undetected until a thromboembolic event such as an ischemic stroke occurs. Approximately 25% of patients with transient
ischemic attack or stroke are found to have atrial fibrillation, diagnosed only after the event. In more than 25% of strokes, the stroke itself is the initial manifestation of atrial fibrillation. Even subclinical atrial fibrillation is a significant risk factor. A recent meta-analysis found a 2.4-fold increase in annual stroke risk (95% confidence interval [CI] 1.8–3.3, \( P < .001 \)) in patients with subclinical atrial fibrillation compared to those without. Therefore, early recognition is critical.

Consumer wearables were validated primarily for detection of atrial fibrillation because of the ease of identifying irregular intervals. Most smartphones, smartwatches, and handheld single-lead ECG consumer devices use photoplethysmography (PPG) sensor technology, which measures changes in blood flow based on the intensity of reflected light. This produces pulse intervals known as tachograms, with the “peak-to-peak” interval representing the R-R interval, or the interval from 1 QRS complex to the next. Device-dependent algorithms can therefore be used to detect irregular rhythms based on variation in pulse intervals. Table 1 summarizes the available devices with their regulatory clearance and validation.

### Consumer-grade ‘smart devices’ for detecting cardiac arrhythmias

<table>
<thead>
<tr>
<th>Device</th>
<th>CE and FDA clearance</th>
<th>Validation</th>
<th>PPG monitoring frequency</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>FibriCheck (^6) smartphone camera app</td>
<td>Atrial fibrillation</td>
<td>Validated vs standard 12-lead ECG</td>
<td>Not applicable</td>
<td>95.6</td>
<td>96.6</td>
</tr>
<tr>
<td>KardiaMobile (^7) ECG monitor</td>
<td>Single-lead and 6-lead ECG to detect bradycardia, tachycardia, and atrial fibrillation</td>
<td>Validated vs standard 12-lead ECG</td>
<td>Not applicable</td>
<td>96.6</td>
<td>94.1</td>
</tr>
<tr>
<td>Apple Watch Series 6 (^8)</td>
<td>Irregular heart rhythm notification and ECG monitoring</td>
<td>Validated vs standard 12-lead ECG (^9)</td>
<td>Intermittent (every 5 minutes)</td>
<td>85</td>
<td>75</td>
</tr>
<tr>
<td>Garmin smartwatch (^8)</td>
<td>Garmin Venu 2 Plus model with ECG capability</td>
<td>Garmin Forerunner 945 model validated vs Holter monitoring</td>
<td>Continuous</td>
<td>96.9</td>
<td>99.3</td>
</tr>
<tr>
<td>Samsung smartwatch (^7)</td>
<td>ECG capability</td>
<td>Active 2 model validated vs BioTech ECG patch</td>
<td>Intermittent or continuous (user defined)</td>
<td>96.9</td>
<td>99.3</td>
</tr>
<tr>
<td>Fitbit (^9)</td>
<td>Detecting atrial fibrillation, with ECG capability</td>
<td>Fitbit Sense model validated vs standard 12-lead ECG (^2)</td>
<td>Continuous in some models (eg, Fitbit Charge 5)</td>
<td>66</td>
<td>79</td>
</tr>
<tr>
<td>Withings ScanWatch (^9)</td>
<td>Detecting atrial fibrillation using ECG functionality and measuring blood oxygen saturation</td>
<td>Validated vs standard 12-lead ECG (^2)</td>
<td>Intermittent (every 10 minutes)</td>
<td>58</td>
<td>75</td>
</tr>
</tbody>
</table>

\( ^{*} \) The BASEL Wearable Study (reference 9) also validated Samsung Galaxy Watch 3 and KardiaMobile against standard 12-lead ECG and demonstrated closely comparable sensitivity and specificity to the Apple Watch, Fitbit, and Withings ScanWatch.

**CE** = Conformité Européenne; **ECG** = electrocardiography; **FDA** = US Food and Drug Administration; **PPG** = photoplethysmography

### Consumer-grade vs medical-grade devices: validation comparisons

Overall, the sensitivity of smart devices for atrial fibrillation detection is remarkably high. A recent meta-analysis found that smartphones detected atrial fibrillation with a sensitivity of 94% and a specificity of 96%, and there was no difference in atrial fibrillation detection between devices that use PPG and single-lead ECG. \(^1\) Another meta-analysis showed that smartwatches were noninferior to medical-grade devices for detecting atrial fibrillation. \(^12\)
Camera applications
Using smartphone camera applications to detect atrial fibrillation is convenient, as it is easily accessible and requires no additional hardware. FibriCheck is the only smartphone-based application with US Food and Drug Administration (FDA) clearance for rhythm monitoring. It uses the light-emitting diode smartphone flash reflected from the finger (the index finger is placed on the smartphone’s camera) or from facial video recordings. A validation study that compared the FibriCheck atrial fibrillation algorithm with a standard 12-lead ECG found that the application’s sensitivity and specificity for atrial fibrillation detection were 95.6% and 96.6%, respectively. A meta-analysis of 3,852 participants found that smartphone camera applications for diagnosing atrial fibrillation (Cardio Rhythm Mobile, PULSESMART, FibriCheck, and Preventicus) were highly successful in detecting atrial fibrillation (combined sensitivity 94.2%, specificity 95.8%). The negative predictive value was high (99.8%) in all analyses, but the positive predictive value was very low (19.3%–37.5%) in asymptomatic individuals age 65 or older.

Smartphone-paired devices
Handheld ECG devices are comparable in ease of use with the standard single-lead devices such as Zio patch but have the benefit of real-time monitoring. However, data from the Zio patch can be seen only after it is mailed in.

KardiaMobile is a small, portable handheld ECG device that can provide a 30-second single-lead ECG. The user places 1 finger of each hand on the electrodes and the device wirelessly transmits the ECG to a connected smartphone. It has multiple forms, including a small handheld device, phone case, watchband, and card. KardiaMobile 6L has the ability to record all 6 limb leads. A single-center study examined whether KardiaMobile could accurately detect atrial fibrillation. When blinded electrophysiologists compared the KardiaBand ECGs with standard 12-lead ECGs, the sensitivity and specificity of KardiaBand ECG recordings for detecting atrial fibrillation were 93% and 84%, respectively, with a K coefficient of 0.77. In a similar study, patients with KardiaMobile were instructed to record their ECG 3 times daily or if they had palpitations. The KardiaMobile detection rate was superior to 24-hour ECG monitoring (9.5% vs 2.0%, respectively). Monitoring with KardiaMobile also seemed to increase atrial fibrillation detection. The REHEARSE-AF study (Assessment of Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation) randomized 1,001 participants over age 65 with no history of atrial fibrillation to standard care vs twice-weekly monitoring with AliveCor Kardia. Atrial fibrillation was noted in 3.8% of patients in the handheld ECG arm compared with less than 1% in the standard-care arm.

Smartwatches
The Apple Watch uses PPG technology to periodically measure heart rate and rhythm over 1-minute intervals while the user is stationary. It can also continuously monitor every 6 seconds during workout mode. Earlier models of the Apple Watch (Series 1 to 3) had only PPG technology. Newer models have incorporated single-lead 30-second ECG, which can be recorded on demand through electrodes on the back of the watch and the watch crown. Of note, the International Trade Commission recently ruled that most Apple Watch models contain technology that infringes on patents held by Masimo Corporation. A cease-and-desist order on sales of the Apple Watch is scheduled to take effect December 26, 2023.

Similarly, Garmin watches also use PPG technology. The Garmin Venu 2 Plus has ECG capability and FDA clearance for detecting arrhythmias. The Samsung smartwatches, including the Galaxy Watch 3 and Galaxy Watch Active 2, and the Withings ScanWatch have PPG and ECG technology.

The growth of the smartwatch market makes it easier to conduct studies with large sample sizes. The Apple Heart Study recruited 419,297 participants without atrial fibrillation over 8 months. Participants who received a notification of irregular pulse through their smartwatch would get a telemedicine visit and have an ECG patch mailed to them to monitor for up to 7 days. More than 2,000 participants (0.52%) had irregular pulse notifications; 450 returned their ECG patches with analyzable data. The positive predictive value for irregular pulse notifications for atrial fibrillation was 84% (95% CI 0.76–0.92).

The Fitbit device, with 37 million active users as of 2022, is a wrist-worn device used primarily as a fitness tracker, but it is also equipped with PPG technology. Fitbit models such as Fitbit Sense and Charge 5 can also record a single-lead ECG. The Fitbit Heart Study is a large prospective remote clinical trial that enrolled Fitbit users. Compared with the Apple Heart Study, it showed a better positive predictive value at 98.2% (95% CI 95.5%–99.5%).

While the Apple Heart Study used smartwatch PPG technology, it only monitored 1-minute intervals every 2 hours. Other studies assessed the ability of smartwatches using continuous PPG monitoring to detect atrial fibrillation and quantify atrial fibrillation burden.
in a daily-living setting. In Avram et al,7 when the Samsung Galaxy Watch Active 2 was compared with a 28-day Holter monitor, it was found to have moderate ability to detect atrial fibrillation with PPG (sensitivity 87.8%, specificity 97.4%). Sensitivity improved to 96.9% and specificity improved to 99.3% with the addition of on-demand ECG for rhythm confirmation.7 In another study, the Garmin smartwatch also had high sensitivity, specificity, and positive predictive value for atrial fibrillation detection.8

Mannhart et al9 assessed the accuracy of 5 wearable smart devices in detecting atrial fibrillation and found that the sensitivity and specificity for atrial fibrillation detection were comparable between devices. A manual review was required in about one-fourth of the cases due to inconclusive tracings.

**Concerns: Duration, Risk Reduction, Other Arrhythmias**

**Duration of monitoring**

Since rhythm monitoring with PPG is usually intermittent and of short duration (typically less than 5 minutes at a time), there is a theoretical concern that it may have a lower detection rate than longer-duration sampling. However, longer sampling frequency did not improve atrial fibrillation detection in the Watch AF trial (Smartwatches for the Detection of Atrial Fibrillation),22 which compared a smartwatch-based algorithm using PPG signals vs a single-lead handheld ECG analyzed by 2 cardiologists. The smartwatch algorithm detected atrial fibrillation based on 1-minute PPG recordings with 96.1% accuracy, and the diagnostic accuracy did not improve significantly with 3-minute or 5-minute recording durations.22

**Does increased atrial fibrillation detection reduce stroke risk?**

It is crucial to determine whether increased detection with smart devices leads to increased use of therapeutic anticoagulation and reduced stroke risk. The Heartline Study,21 an ongoing randomized app-based trial with more than 26,000 participants age 65 and older, addresses this uncertainty. Patients were randomized to 2 cohorts based on the presence of atrial fibrillation and were further randomized to a digital engagement program with or without the Apple Watch. The key outcomes are the detection of atrial fibrillation in patients with no prior history of atrial fibrillation and improved adherence to direct oral anticoagulation in patients previously diagnosed with atrial fibrillation.21

The STROKESTOP trial24 (Systematic ECG Screening for Atrial Fibrillation Among 75-year-old Subjects in the Region of Stockholm and Halland, Sweden) randomly assigned 27,993 participants residing in the region of Halland and Stockholm, age 75 to 76, to a control group or to the use of a handheld single-lead Zenicor-ECG device twice daily for 2 weeks. At 6.9 years of follow-up, the screening group had a lower incidence of the combined end point of ischemic or hemorrhagic stroke, systemic embolism, bleeding leading to hospitalization, and all-cause mortality, although the effect was small (5.45 vs 5.68 events per 100 patient-years), with a hazard ratio of 0.96 (95% CI 0.92–1.00, P = .045).24

**Detection of other rate or rhythm abnormalities**

A small pilot study assessed the feasibility of measuring the corrected QT interval with KardiaMobile vs standard 12-lead ECG. The handheld single-lead ECG was noninferior to standard 12-lead ECG and was accurate within a range of plus or minus 20 ms.25 The new 6-lead KardiaMobile device has interval measurements comparable to a standard 12-lead ECG.26 There is currently no commercially available QT interval measurement algorithm, though preliminary data show that the Apple Watch can reliably assess the corrected QT interval.27

Validation of consumer-grade devices has been less promising in detecting supraventricular tachycardia, in part because of the regular ventricular rhythm and lack of variation of the R-R interval.8 A prospective multicenter validation study of 50 patients aimed to improve the detection of atrial flutter using KardiaMobile.27 After KardiaMobile recorded lead I, the device was repositioned by holding the panel in the right hand and placing the opposite electrode onto the left leg to generate a lead II. Two independent blinded electrophysiologists analyzed the recordings. The sensitivity of lead I alone for detecting atrial flutter was poor for both electrophysiologists at 27.3%, but sensitivity improved to 72.7% and 54.6% with the incorporation of the additional lead.28

Detection of other forms of supraventricular tachycardia, pathologic Q waves, and heart blocks has received limited study. One study of Apple Watch 2, Samsung Galaxy Gear S3, and Fitbit Charge 2 found excellent accuracy in diagnosing the heart rate of supraventricular tachycardia, but the rhythm was not analyzed.29 The sensitivity of KardiaMobile in detecting pathologic Q waves was found to be 20.6% in a study by Koltowski et al.30 Limited data suggest that Apple Watch’s single-lead ECG may help recognize first- and second-degree atrioventricular block.31
LIMITATIONS OF CONSUMER-BASED DEVICES

The availability of consumer-grade heart rhythm monitors comes with limitations as well as potential for future research, including the following:
- A high false-positive rate
- Disparities in access
- An influx of consumer-grade data on a strained provider workforce
- Potential for improvements in technology and data.

False-positive results and pretest probability

The high false-positive rate for detecting atrial fibrillation in young, otherwise healthy populations is a significant limitation of consumer-based devices that may lead to increased anxiety and unnecessary healthcare utilization. However, although false alerts have been shown to reduce perceived physical well-being, the financial impact of false-positive detections is not well understood.32

As with all medical tests, the positive predictive value varies significantly based on the patient population. An important tenet of Bayesian reasoning is that the posttest probability depends on the pretest probability. In other words, atrial fibrillation detected on a smartwatch in a young, healthy patient (low pretest probability) is unlikely to be atrial fibrillation. In contrast, atrial fibrillation detected in an elderly hypertensive patient with obstructive sleep apnea (high pretest probability) is highly likely to be atrial fibrillation. Atrial fibrillation incidence increases with age, from 1.5% at age 55 to 59 to 23.5% at age 80 to 89.33 Both the Apple Heart Study19 and the Fitbit Heart Study20 noted higher rates of detection and diagnosis of atrial fibrillation in participants age 65 and older. The VITAL-AF Study (Screening for Atrial Fibrillation in Older Adults at Primary Care Visits)34 evaluated more than 30,000 participants age 65 or older without atrial fibrillation. The study compared KardiaMobile vs usual care and found no difference in the incidence of atrial fibrillation diagnosis between the screening and the control groups. However, in a prespecified analysis of patients over age 85, atrial fibrillation was more likely to be detected in the screening group than in the control group (3.56% vs 3.76%).34

Disparities

There are disparities in device access and utilization. Only one-third of US adults and 18% of patients with cardiovascular disease have smart devices. Further, patients over age 65 and those with lower education
and socioeconomic status have less access to smart devices and a higher risk of atrial fibrillation.10

Burden to healthcare system
The influx of data from consumer-grade devices will increase the burden on an already strained healthcare system. In addition to more data, automated rhythm readings may be deemed inconclusive despite producing readable single-lead ECGs, as was shown in the BASEL Wearable Study.9 However, a manual review of the tracings by a cardiologist reduced the rate of inconclusive tracings from 26% to around 1%.9 There are no well-established best practices for physician notifications, documentation, reimbursement protocols, and care coordination with detection of atrial fibrillation from consumer-grade devices.

Improved technology and security
Large, high-quality, randomized controlled trials demonstrating that wearable atrial fibrillation detection improves hard clinical outcomes are still lacking, and it is hoped that the randomized Heartline Study23 will address some of these gaps. Future trials and observational studies are needed to determine whether earlier atrial fibrillation diagnosis from consumer-grade wearable devices increases adherence to appropriate anticoagulation and reduces adverse events. Further studies on cost-effectiveness are also needed.

Advances in sensors like improved PPG and multilead ECG may enhance accuracy of detection. More sophisticated algorithms that utilize deep learning on large ECG datasets could also improve performance and decrease false-positive results.

OUR APPROACH
Our approach to atrial fibrillation identified on consumer devices is summarized in Figure 1. Given the high sensitivity of each device, lack of detection on device interrogation makes atrial fibrillation unlikely regardless of pretest probability. We consider atrial fibrillation “unlikely” rather than “ruled out,” given that consumer-grade devices are not truly continuous (they sample PPG or ECG only intermittently), are not always worn, and may need to be removed for charging.

If atrial fibrillation is detected, we review the tracings from the device, if available. It is common for a manual review to demonstrate normal rhythm with ectopy or sinus arrhythmia, in which case reassurance and continued consumer-grade monitoring are appropriate.

If no tracings are available or the tracings suggest atrial fibrillation, we move on to medical-grade cardiac monitoring because of a slightly higher specificity in medical-grade devices. If the medical-grade monitor also shows atrial fibrillation, we diagnose atrial fibrillation and engage in shared decision-making with the patient about the risks and benefits of treatment.

If the consumer-grade device suggests atrial fibrillation and the medical-grade device shows none, we assess the pretest probability of atrial fibrillation. If the pretest probability is low and there are clear alternate causes of an irregular rhythm on the medical-grade monitoring (such as sinus rhythm with frequent ectopy or sinus arrhythmia), we consider atrial fibrillation unlikely.

If the medical-grade device shows no atrial fibrillation and we think the pretest probability is high, we typically increase the monitoring duration via longer Holter monitoring or, if the arrhythmia is infrequent, an implantable loop recorder.

OUTLOOK: BETTER DETECTION, BETTER TREATMENT
With comparable sensitivity to medical-grade devices, wearable consumer-grade devices show promise in detecting cardiac arrhythmias, particularly atrial fibrillation. These increasingly common devices can potentially improve the detection of atrial fibrillation and the prescription of therapeutic anticoagulation in appropriate cases, leading to improved patient outcomes. Given the high sensitivity and lower specificity of these devices, absence of atrial fibrillation should be reassuring, while detected atrial fibrillation should prompt further testing with medical-grade devices and referral to an experienced ECG reader. As with any diagnostic test, the result needs to be contextualized with an understanding of the pretest probability of atrial fibrillation. Ongoing research will address the effectiveness of these devices in detecting other cardiac pathologies and their impact on long-term outcomes, such as stroke risk and therapeutic anticoagulation.

DISCLOSURES
Dr. Jensen has disclosed being an advisor or review panel participant for, having ownership interest in, and being an executive level board member for Dose Health. The other authors report no relevant financial relationships which, in the context of their contributions, could be perceived as a potential conflict of interest.
REFERENCES


