# **Review**

# Diagnostic Accuracy of Smartwatches for the Detection of Cardiac Arrhythmia: Systematic Review and Meta-analysis

Scarlet Nazarian<sup>\*</sup>, BSc, MBBS, MRCS; Kyle Lam<sup>\*</sup>, MA (Cantab), MB BChir, MRCS; Ara Darzi, PC, KBE, FRS, FMedSci, HonFREng; Hutan Ashrafian, PhD

Imperial College London, London, United Kingdom \*these authors contributed equally

#### **Corresponding Author:**

Hutan Ashrafian, PhD Imperial College London 10th Floor QEQM Building, St Mary's Hospital Praed Street London, W2 1NY United Kingdom Phone: 44 7799 871 597 Email: <u>hutan@researchtrials.net</u>

# Abstract

**Background:** Significant morbidity, mortality, and financial burden are associated with cardiac rhythm abnormalities. Conventional investigative tools are often unsuccessful in detecting cardiac arrhythmias because of their episodic nature. Smartwatches have gained popularity in recent years as a health tool for the detection of cardiac rhythms.

**Objective:** This study aims to systematically review and meta-analyze the diagnostic accuracy of smartwatches in the detection of cardiac arrhythmias.

**Methods:** A systematic literature search of the Embase, MEDLINE, and Cochrane Library databases was performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to identify studies reporting the use of a smartwatch for the detection of cardiac arrhythmia. Summary estimates of sensitivity, specificity, and area under the curve were attempted using a bivariate model for the diagnostic meta-analysis. Studies were examined for quality using the Quality Assessment of Diagnostic Accuracy Studies 2 tool.

**Results:** A total of 18 studies examining atrial fibrillation detection, bradyarrhythmias and tachyarrhythmias, and premature contractions were analyzed, measuring diagnostic accuracy in 424,371 subjects in total. The signals analyzed by smartwatches were based on photoplethysmography. The overall sensitivity, specificity, and accuracy of smartwatches for detecting cardiac arrhythmias were 100% (95% CI 0.99-1.00), 95% (95% CI 0.93-0.97), and 97% (95% CI 0.96-0.99), respectively. The pooled positive predictive value and negative predictive value for detecting cardiac arrhythmias were 85% (95% CI 0.79-0.90) and 100% (95% CI 1.0-1.0), respectively.

**Conclusions:** This review demonstrates the evolving field of digital disease detection. The current diagnostic accuracy of smartwatch technology for the detection of cardiac arrhythmias is high. Although the innovative drive of digital devices in health care will continue to gain momentum toward screening, the process of accurate evidence accrual and regulatory standards ready to accept their introduction is strongly needed.

**Trial Registration:** PROSPERO International Prospective Register of Systematic Reviews CRD42020213237; https://www.crd.york.ac.uk/prospero/display\_record.php?RecordID=213237.

(J Med Internet Res 2021;23(8):e28974) doi: 10.2196/28974

# **KEYWORDS**

wearables; smartwatch; cardiac arrhythmia; atrial fibrillation; cardiology; mHealth; wearable devices; screening; diagnostics; accuracy

# Introduction

# Background

Cardiac arrhythmia encompasses a group of conditions in which the heart beats too quickly, too slowly, or in an irregular pattern. Significant morbidity, mortality, and financial burden are associated with cardiac rhythm abnormalities [1]. Of these cardiac rhythm abnormalities, atrial fibrillation (AF) is the most common type of cardiac arrhythmia [2], and its prevalence increases sharply with age, reaching 17.8% in a European population for those aged >85 years [3,4]. The presence of AF increases the risk of ischemic stroke by five-fold [5] and can lead to other thromboembolic events. It is well recognized that AF often remains asymptomatic, and therefore, by the time of screening, the patient may have already suffered the consequences.

Although AF is the most common type of cardiac arrhythmia, other arrhythmias, such as premature cardiac contractions, are responsible for significant symptomatic burden. Premature atrial contractions have been shown to be an independent risk factor for all strokes in a longitudinal study [6]. Similarly, a cohort study found that having premature ventricular contractions resulted in a higher rate of ischemic stroke than those without contractions [7].

Conventional screening tools, in the form of 12-lead electrocardiograms (ECGs) and ambulatory electrocardiography monitors, are often unsuccessful in detecting AF or other cardiac arrhythmias, such as bradyarrhythmias or tachyarrhythmias, because of the transient nature of episodes. The episodic and infrequent nature of cardiac arrhythmias means that they are not captured within the investigation period, making diagnosis very difficult.

Recent advances in mobile health technology and wearable electronic devices allow heart rhythm monitoring to be undertaken in real time with greater comfort, ease, and engagement [8]. Wearable devices such as smartwatches show great potential for the detection of cardiac arrhythmias. Timely diagnosis of AF ensures that management is commenced early to prevent ensuing events that impact the quality of life while also relieving the burden that this poses on the health care system.

Smartwatches have gained popularity in recent years, especially as a health tool for the detection of heart rhythms. Patients with a smartwatch can self-diagnose their heart rhythm within 30 seconds using one finger [<mark>9</mark>]. These apps use photoplethysmography (PPG) from an optical sensor to analyze the pulse rate from the wrist [10]. However, adoption of the technology by both clinicians and patients requires that these devices are accurate and provide clinically applicable information in a manner that is compatible with workflow in the health setting.

# Objectives

This study aims to systematically review and meta-analyze the diagnostic accuracy of smartwatches in the detection of cardiac arrhythmias.

```
Nazarian et al
```

# Methods

# Overview

This review was carried out and reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [11]. The review was registered at the International Prospective Register of Systematic Reviews (PROSPERO ID: CRD42020213237).

# Search Strategy

A thorough literature search was performed using the Embase, MEDLINE, and Cochrane Library databases. All articles published until February 2021 were included in the study. The appropriate MeSH (Medical Subject Headings) terms and free text all field searches were performed and combined with appropriate Boolean operator terms for *arrhythmias, cardiac OR irregular pulse*\* *OR atrial fibrillation, wearable electronic devices OR smartwatch*\* *OR wristband*\*, *diagnosis, computer-assisted OR diagnos*\*, and *detect*\* in Embase and Ovid in MEDLINE. Search terms in the Cochrane Library *included arrhythmias, cardiac OR atrial fibrillation OR irregular pulse*\* *OR arrhythmia*\*, *smartwatch*\* *OR wearable electronic device*\*, and *diagnosis, computer-assisted OR detect*\* *OR diagnos*\*. The full search strategy is provided in Multimedia Appendix 1.

# **Inclusion and Exclusion Criteria**

Inclusion criteria were as follows:

- studies reporting detection of cardiac arrhythmias using smartwatches;
- studies reporting sensitivity, specificity and diagnostic accuracy; or studies with adequate information to calculate these data; and
- studies published or translated into English.

Exclusion criteria were as follows:

- studies with no original data present (eg, review article, letter);
- studies with no full text available;
- studies >20 years; and
- studies without adequate data to calculate sensitivity, specificity and diagnostic accuracy data.

# **Study Selection**

Studies obtained from the literature search were analyzed, and duplicates were removed. Title, abstract, and full-text review were performed by 2 reviewers independently, and irrelevant studies were excluded. Disagreements were settled by consensus among the reviewers.

#### **Data Extraction**

Data were extracted onto a standard spreadsheet template. Information regarding the journal, author, study design, type of smartwatch, number of subjects, and diagnostic accuracy data (sensitivity, specificity, accuracy, positive predictive value [PPV], and negative predictive value [NPV]) was selected from each paper.

#### **Study Quality Assessment**

The Quality Assessment of Diagnostic Accuracy Studies 2 tool was used to assess the risk of bias of the included studies [12]. Each domain was classified as low risk, high risk, or unclear risk of bias.

### **Statistical Analysis**

Summary estimates of sensitivity, specificity, and area under the curve data were attempted using a bivariate model for diagnostic meta-analysis. Independent proportions and their differences were calculated and pooled using DerSimonian and Laird random effects modeling [13]. This considered both between-study and within-study variances, which contributed to study weighting. Study-specific estimates and 95% CIs were computed and represented in forest plots. Statistical heterogeneity was determined by the  $I^2$  statistic, where <30% was low, 30%-60% was moderate, and >60% was high. Analyses were performed using Stata version 15 (StataCorp). *P* values of ≤.05 were considered statistically significant.

# Results

# Search Results and Characteristics

The database searches identified 292 studies that matched the criteria. Duplicates were removed, and 215 studies were eligible for title and abstract screening. Following this, a full-text review

was undertaken, and a total of 18 studies were included in this review. Studies that failed to satisfy the inclusion criteria were excluded, and the reasons for exclusion of these articles included wrong intervention (such as the lack of use of a smartwatch) or wrong outcomes (such as studies that did not involve the detection of cardiac arrhythmias or reports on diagnostic accuracy). The study screening and selection process is shown in Figure 1.

The studies included in this systematic review were all published between 2017 and 2021. The outcome measure in the studies was mainly AF detection but also included bradyarrhythmias, tachyarrhythmias, and premature contractions. The studies measured diagnostic accuracy using smartwatches in 424,371 subjects in total. The Apple watch was used in 7 studies, Samsung smartwatches were used in 5 studies, and the remaining studies used a Huawei, Huami, or Empatica smartwatch. One study used the Wavelet wristband. Three different types of Huawei smartwatches were used in 2 studies to assess the diagnostic accuracy [14,15].

The reference standard was an ECG in most studies in the form of a 12-lead ECG, a Holter monitor, an ECG patch, telemetry, or an internet-enabled mobile ECG. In one study, an implantable cardiac monitor was used as the standard [16]. Almost all studies, except for 2 that did not specify, used PPG-based sensors to assess pulse rate. Table 1 provides the characteristics of the included studies.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for study selection.



RenderX

 Table 1. Characteristics of included studies on detection of cardiac arrhythmias.

Authors (year)	Primary outcome	Study design	Type of sensor	Reference standard	Research or re- al-life setting	Type of smart- watch	Number of subjects
Corino et al (2017) [17]	AF <sup>a</sup> detection	Prospective	PPG <sup>b</sup>		Research	Empatica E4	70
Bumgarner et al (2018) [18]	AF detection	Prospective, nonran- domized, adjudicator blinded	_	12-lead ECG <sup>d</sup> (physician re- viewed)	Research	Apple watch	100
Tison et al (2018) [19]	AF detection	Multinational, cohort	PPG	12-lead ECG	Research	Apple watch	1617
Wasserlauf et al (2019) [16]	AF detection	Prospective	PPG	Insertable cardiac monitor	Research	Apple watch	24
Perez et al (2019) [20]	AF detection	Prospective, single group, open label, site less, pragmatic	PPG	ECG patch	Real life	Apple watch	419,297
Zhang et al (2019) [14]	AF detection	Pilot, cohort	PPG	12-lead ECG and physical examina- tion	Real life	Huawei Watch GT	263
						The Honor Watch (Huawei)	263
						The Honor Band4 (Huawei)	209
Ding et al (2019) [21]	AF detection	Observational	PPG	Holter monitor ECG	Research	Samsung Simband 2	40
Dorr et al (2019) [22]	AF detection	Prospective, two cen- ter, case-control	PPG	Internet-enabled mobile ECG	Research	Samsung GearFit 2	508
Bashar et al (2019) [23]	AF detection	Prospective	PPG	Holter monitor ECG	Research	Samsung Simband	20
Bashar et al (2019) [24]	AF detection	Prospective	PPG	Holter monitor ECG	Research	Samsung Simband	37
Valiaho et al (2019) [25]	AF detection	Multicenter prospec- tive case-control	PPG	Three-lead ECG	Research	Empatica E4	213
Guo et al (2019) [15]	AF detection	Prospective	PPG	Clinical evaluation, ECG, or 24-hour Holter monitoring	Real life	Huawei Watch GT	212
						The Honor Watch (Huawei)	265
						The Honor Band4 (Huawei)	264
Chen et al (2020) [26]	AF detection	Prospective	PPG	12-lead ECG (physician re- viewed)	Research	Amazfit Health Band 1S (Huami)	401
Rajakariar et al (2020) [27]	AF detection	Prospective, multicen- ter validation	PPG	12-lead ECG	Research	Apple watch	200
Seshadri et al (2020) [28]	AF detection	Prospective	_	Telemetry	Research	Apple watch	50
Selder et al (2020) [29]	AF detection	Observational, prospective cohort	PPG	One-lead ECG	Research	Wavelet wristband	60
Han et al (2020) [30]	Premature atrial contraction or pre- mature ventricular contraction	Prospective	PPG	ECG patch	Research	Samsung Gear S3	2

Nazarian et al

Authors (year)	Primary outcome	Study design	Type of sensor	Reference standard	Research or re- al-life setting	Type of smart- watch	Number of subjects
Caillol et al (2021) [31]	AF, atrial flutter, brady arrhythmias, and tachyarrhyth- mias	Prospective	PPG	12-lead ECG	Research	Apple watch	256

<sup>a</sup>AF: atrial fibrillation.

<sup>b</sup>PPG: photoplethysmography.

<sup>c</sup>Not available.

<sup>d</sup>ECG: electrocardiogram.

# Sensitivity, Specificity, and Accuracy

The diagnostic accuracy of the smartwatch in detecting cardiac arrhythmias was analyzed, reporting a pooled sensitivity of 100% (95% CI 0.99-1.00; Figure 2) in 17 studies with 5074 subjects and a pooled specificity of 95% (95% CI 0.93-0.97; Figure 3) in 16 studies with 5050 subjects. The sensitivity ranged

from 25% (95% CI 0.14-0.36) to 100% (95% CI 1.00-1.00), whereas the specificity ranged from 68% (95% CI 0.65-0.70) to 100% (95% CI 1.00-1.00).

Of the 18 studies, 7 (39%) reported data on accuracy. Among the 1769 subjects, the pooled accuracy for arrhythmia detection was 97% (95% CI 0.96-0.99; Figure 4).

Figure 2. Pooled analysis for sensitivity of cardiac arrhythmia detection by smartwatches. Effect sizes are shown with 95% CIs. A random effects model was used. ES: effect sizes.

Study – sensitivity of detection			ES (95% CI)	% Weight		
Corino et al.		-	0.75 (0.65, 0.85)	0.03		
Bumgarner et al.		+	0.93 (0.88, 0.98)	0.13		
Tison et al.		•	0.68 (0.65, 0.70)	0.60		
Wasserlauf et al.		+	0.98 (0.91, 1.04)	0.08		
Zhang et al.		+	1.00 (1.00, 1.00)	10.81		
Zhang et al.		+	1.00 (1.00, 1.00)	10.81		
Zhang et al.		•	1.00 (1.00, 1.00)	10.80		
Ding et al.		+	0.98 (0.94, 1.02)	0.19		
Dorr et al.		•	0.94 (0.92, 0.96)	0.69		
Bashar et al.		-+ <u> </u>	0.96 (0.88, 1.05)	0.05		
Bashar et al.		+	0.98 (0.94, 1.02)	0.17		
Valiaho et al.		+	0.96 (0.94, 0.99)	0.48		
Guo et al.		•	1.00 (1.00, 1.00)	10.80		
Guo et al.		+	1.00 (1.00, 1.00)	10.81		
Guo et al.		•	1.00 (1.00, 1.00)	10.81		
Chen et al.		+	0.80 (0.76, 0.84)	0.21		
Rajakariar et al.			0.94 (0.91, 0.98)	0.31		
Seshadri et al.		+	0.96 (0.91, 1.01)	0.11		
Selder et al.		•	1.00 (1.00, 1.00)	10.64		
Han et al.		•	- 0.93 (0.56, 1.29)	0.00		
Caillol et al.a		÷	0.96 (0.91, 1.01)	0.14		
Caillol et al.b			0.25 (0.14, 0.36)	0.03		
Caillol et al.c		•	1.00 (1.00, 1.00)	10.65		
Caillol et al.d		•	1.00 (1.00, 1.00)	10.65		
Overall (I-squared = 98.0%, p = 0.000)			1.00 (0.99, 1.00)	100.00		
NOTE: Weights are from random effects analysis						
-1.29	ò	1	.29			

XSL•FO RenderX

Figure 3. Pooled analysis for specificity of cardiac arrhythmia detection by smartwatches. Effect sizes are shown with 95% CIs. A random effects model was used. ES: effect sizes.

Study – specificity of detection	ES (95% CI)	% Weight
Corino et al.	0.96 (0.92, 1.01)	4.50
Bumgarner et al.	🛨 ¦ 0.84 (0.77, 0.91)	3.31
Tison et al.	• 0.68 (0.65, 0.70)	5.37
Zhang et al.a	• 0.99 (0.98, 1.00)	5.67
Zhang et al.b	• 0.99 (0.98, 1.00)	5.67
Zhang et al.c	• 0.99 (0.98, 1.00)	5.61
Ding et al.	• 0.98 (0.94, 1.02)	4.59
Dorr et al.	• 0.98 (0.97, 0.99)	5.66
Bashar et al.	0.97 (0.90, 1.04)	3.38
Bashar et al.	★ 0.97 (0.92, 1.03)	4.20
Valiaho et al.	<ul> <li>0.98 (0.96, 1.00)</li> </ul>	5.50
Guo et al.a	• 0.99 (0.97, 1.00)	5.61
Guo et al.b	• 0.99 (0.98, 1.00)	5.68
Guo et al.c	• 0.99 (0.98, 1.00)	5.68
Chen et al.	• 0.97 (0.95, 0.99)	5.54
Rajakariar et al.	• 0.82 (0.77, 0.87)	4.09
Seshadri et al.	• 1.00 (1.00, 1.00)	5.77
Selder et al.	• 0.96 (0.91, 1.01)	4.26
Han et al.	<u>→</u> 0.99 (0.83, 1.15)	1.23
Caillol et al.a		3.38
Caillol et al.b	• 0.99 (0.97, 1.01)	5.32
Overall (I-squared = 97.7%, p = 0.000)	♦ 0.95 (0.93, 0.97)	100.00
NOTE: Weights are from random effects	analysis	
-1.15 0	1.15	



Figure 4. Pooled analysis for accuracy of cardiac arrhythmia detection by smartwatches. Effect sizes are shown with 95% CIs. A random effects model was used. ES: effect sizes.



# **PPV and NPV Analysis**

The PPV for cardiac arrhythmia detection was assessed in 9 studies using a smartwatch. These included a total of 421,267

subjects and reported a PPV of 85% (95% CI 0.79-0.90; Figure 5). The pooled NPV was reported in 6 studies as 100% (95% CI 1.0-1.0; Figure 6), taking into consideration 3323 subjects.



Figure 5. Pooled analysis for PPV of cardiac arrhythmia detection by smartwatches. Effect sizes are shown with 95% CIs. A random effects model was used. ES: effect sizes; PPV: positive predictive value.





Figure 6. Pooled analysis for NPV of cardiac arrhythmia detection by smartwatches. Effect sizes are shown with 95% CIs. A random effects model was used. ES: effect sizes; NPV: negative predictive value.



# **Heterogeneity of Studies**

There was a high degree of variation between studies assessing cardiac arrhythmia detection using a smartwatch. The heterogeneity was statistically significant when all the studies were compared (P<.05). The lowest variation among studies was seen when reporting the accuracy of smart devices to detect arrhythmias ( $I^2$ =81.1%), whereas heterogeneity was highest in studies when assessing PPV ( $I^2$ =98.5%).

#### **Quality Assessment**

The assessment of bias using the Quality Assessment of Diagnostic Accuracy Studies 2 tool for the included studies is highlighted in Multimedia Appendix 2 [14-31].

# Discussion

#### **Principal Findings**

To the best of our knowledge, this systematic review and meta-analysis is the first to investigate the diagnostic accuracy of smartwatches for all cardiac arrhythmias. We have shown that the detection of cardiac arrhythmias using commercially available smartwatches is possible, with very high diagnostic accuracy. The overall sensitivity, specificity, and accuracy of these digital systems were 100%, 95%, and 97%, respectively. The pooled PPV and NPV for detecting cardiac arrhythmias

were 85% and 100%, respectively. These values may offer clinicians a quantifiable appreciation for the use of smartwatches in a health care setting.

Although the aim of this study is to review the diagnostic accuracy of smartwatches in detecting cardiac arrhythmias, it is clear from the results that there are currently very few studies that assess the ability of PPG technology on smartwatches to detect non-AF arrhythmias.

#### **Smartwatches**

A wide variety of smartwatches are commercially available, and this is reflected in the diverse range of smartwatches used in these studies (Table 2). These devices range from fitness trackers to more medically oriented watches with prices between US \$40 and US \$1700. Although all devices use PPG sensors (Figure 7), there is diversity in functionality beyond this point. Several smartwatches are capable of recording a single-lead ECG, and others, such as the Empatica E4, have electrodermal activity sensors capable of recording sympathetic nervous system activity. The Samsung Simband is unique within these studies in that it is the only device designed for developers and is not commercially available, allowing custom adaption of sensor inclusion. Of the studies included, only the Apple Smartwatch has Food and Drug Administration (FDA) clearance for its ECG tracking functionality.



Table 2. Characteristics of smartwatches used in included stud
--

Smartwatch	Company	Country	Approximate price <sup>a</sup> , £ (US \$)	Туре	Photoplethys- mography	Single-lead ECG <sup>b</sup>	Food and Drug Administration clearance ECG tracking	Electrodermal activity sensor
Apple Watch	Apple	USA	388 (531)	Watch	✓ <sup>c</sup>	✓	1	
Honor Watch	Honor	China	86 (117)	Watch	1			
Huawei GT	Huawei	China	89 (122)	Watch	$\checkmark$			
Gear S3	Samsung	South Korea	160 (219)	Watch	$\checkmark$			
Simband	Samsung	South Korea	N/A <sup>d</sup>	Watch	1	1		✓
Honor Band	Honor	China	45 (61)	Fitness Band	✓			
Amazfit Health- band	Huami	China	33 (45)	Fitness Band	✓			
GearFit2	Samsung	South Korea	49 (67)	Fitness Band	$\checkmark$			
Wavelet wrist- band	Biostrap or Wavelet Health	USA	180 <sup>e</sup> (246)	Wristband	<i>√</i>			
Empatica E4	Empatica	USA	1227 <sup>f</sup> (1682)	Wristband	1	✓		1

<sup>a</sup>Pricing as per Amazon UK website on 22/04/2021.

<sup>b</sup>ECG: electrocardiogram.

<sup>c</sup>Included with smartwatch.

<sup>d</sup>N/A: not applicable or data not available.

<sup>e</sup>Pricing as per Biostrap shop on 22/04/2021.

<sup>f</sup>Pricing as per Empatica store on 22/04/2021.

Figure 7. Overview of photoplethysmography sensor detection of arrhythmia. PPG: photoplethysmography.



#### Classification

### The Impact of Improving AF Detection

The incidence of AF increases annually with an increase in the prevalence of risk factors, such as advancing age, obesity, hypertension, and type 2 diabetes. The challenge with detection is the ability of AF to remain asymptomatic or intermittent before eventually revealing itself. This poses a huge economic burden, accounting for 1%-2% of health care expenditure [32]. A new technology that is promising for reducing or preventing AF-related morbidity, and in doing so, addressing this burden, is welcomed. Machine learning coupled with smartwatches provides the opportunity to detect asymptomatic arrhythmias in a timely manner, allowing appropriate management to be

```
https://www.jmir.org/2021/8/e28974
```

RenderX

initiated early. A recent study showed that a trained deep neural network was able to outperform single cardiologists by accurately classifying a broad range of rhythm classes and distinguishing between artifacts and arrhythmias [33]. This method could reduce the rate of misdiagnosed rhythms by digital ECG machines and improve the efficiency of expert human ECG interpretation by accurately prioritizing the most urgent conditions.

The detection of cardiac arrhythmias using smartwatches has multiple functionalities. It can be used to diagnose an abnormal rhythm, for monitoring of an arrhythmia, for example, in those with known paroxysmal AF, or for screening. Current methods

of AF detection are criticized for their periodic investigative approach, during which an irregular pulse may be absent [34-36]. Using smartwatches, users can diagnose an irregular pulse by placing a finger on their device at any point. Smartwatch devices that detect cardiac arrhythmias are a simple, noninvasive, and user-friendly alternative to current ECG monitoring tools, such as 24-hour Holter monitoring or implantable cardioverter defibrillators [37,38]. The novel devices provide users with prospective information in real time, with relatively high sensitivity and specificity, as shown in our study, and are cost-effective [39]. However, the adoption of this technology by clinicians and patients requires clinically meaningful results in a manner that is compatible with the workflow of clinicians. Therefore, an optimal strategy for their implementation must be in place.

# What Are the Next Steps?

Wearable devices for *wellness* are viewed as low-risk fitness monitors by the FDA, which does not apply the same stringent regulations as it would when considering medical devices. The FDA has introduced its Digital Health Precertification Program, in which companies are able to gain expedited clearance for ECG analysis and heart rate sensing software [40]. This process leads to companies producing technology that is confirmed to be *safe* but not necessarily of good quality because they have bypassed the conventional workflow for research discovery. Large clinical trials are lacking, and as a result, no expert consensus recommends screening for all occult AF [41].

Furthermore, there is insufficient evidence on the burden of smartwatch-detected AF, which would prompt further evaluation and treatment. Guidance on what the clinician is expected to do with an episode of AF detected by a smartwatch is lacking. We suggest that this should be a critical prerequisite before introducing a digital detection tool into the general population; otherwise, overdiagnosis and an expectant role of clinicians from the public to assess their device-detected condition will become an even bigger burden on the health care system. A recent study evaluating the clinical outcome of the Apple smartwatch concluded that false-positive screening results may lead to overutilization of the health care system [42]. Preparation for the problems that a new generation of smartwatch technology, which attempts to bridge the gap between disease and the health care system, brings is key.

With evolving technology in the field of health care applications, there is a move to a more personalized and *patient-centric* approach, where patients have an increasing number of tools at their disposal to assess risk and diagnose disease. Although frequent and active screening using a smartwatch is potentially feasible, few studies have examined the long-term adherence to this system. This user-involved measurement could too easily miss minimally symptomatic and brief paroxysms of arrhythmia. Long-term commitment and adherence from the user or the ability of continuous monitoring by the device is required for an accurate and worthwhile outcome.

#### Limitations

There are many limitations to the studies in our review. At present, most studies have assessed the use of a PPG sensor and

```
https://www.jmir.org/2021/8/e28974
```

an accompanying algorithm to detect cardiac arrhythmias. However, they have not gone further to assess the use of such systems in health care. The largest study within our systematic review did not go beyond the participants' self-reporting of an irregular pulse [20]. Several factors must be controlled to produce unbiased data that are clinically applicable. The published papers included observational and case-control studies, which did not evaluate the efficacy of smartwatch-based screening for clinical outcomes nor reflect real-life conditions [21,22]. Moreover, the sample sizes of some studies were small, with data sets of less than 50 in 5 of the papers [16,21,23,24,30]. One study had a very low sensitivity compared with others when assessing atrial flutter or tachycardia, which could likely be because of the small sample size for this group [31]. Finally, most studies were conducted in controlled research environments as opposed to a real-life setting, which may call into question the diagnostic accuracy of these smartwatches in an uncontrolled environment. Therefore, the interpretation of a sensitivity of 100%, effectively ruling out the presence of a cardiac arrhythmia with a negative result and the interpretation of an NPV of 100%, suggesting the return of no false negatives, should be interpreted with caution. The significant heterogeneity between studies is likely a result of different study settings, different patient group sizes, and different devices, based on personalized algorithms, having been used. Although the presence of this heterogeneity demands caution in interpreting our results, it also stresses the need for randomized controlled trials in this field using large data sets.

Many studies had a large proportion of data excluded because of insufficient PPG signal quality [14,18,22,24]. Some studies took place in settings where patients were supervised and provided instructions on the technique [22,27]. Thus, generalizing these findings to the *real world* could weaken the diagnostic accuracy. PPG technology recognizes the cardiac cycle by the pulsatile pattern of the change in light absorption, which reflects the volumetric alteration in the microvascular beds underneath the skin. With an accurate estimation, each episode of maximum reflected light absorption translates into an R wave. Although previous research has questioned the use of PPG sensors in darker skin, a recently published study showed no statistically significant differences in wearable heart rate measurement accuracy across skin tones [43]. However, a number of studies have shown that PPG sensors are less reliable at higher heart rates and during exercise [44,45]. As some studies in this review did not report the average heart rate of participants, it may add a level of bias to the results. In addition, PPG technology cannot detect myocardial ischemia or arrhythmias with a ventricular origin and therefore, at present, cannot completely replace 12-lead ECGs. Therefore, one must question whether the application of PPG-based sensors for cardiac arrhythmia detection is premature.

Finally, for smartwatch devices to be used as a screening tool for cardiac arrhythmias, such as AF detection, the value is highly dependent on disease prevalence. The estimated prevalence of AF in adults is between 2% and 4%. The prevalence increases with age, especially for those aged >65 years [46]. However, only 4.6% of smartwatch users in the United States are aged >65 years, and among those that are current smartwatch users,

XSL•FO

the prevalence of AF is low [47]. The studies in this review used estimated disease prevalence rates, which have been age adjusted, when assessing diagnostic accuracy values. Subjects were not limited to age groups, and as a result, some studies overestimated AF prevalence among smartwatch users [19,22]. This means that the PPV value of 85% in our review may be higher than expected for cardiac arrhythmia detection in smartwatch users. Either way, there is a high number of false positives, which leads to unnecessary anxiety among those in whom the device detects AF and may have the downstream consequences of inappropriate initiation of treatment in these patients. Treatment with anticoagulants can cause bleeding, which may be harmful. False positives may improve if the device is targeted to those most at risk of AF, but larger studies are needed to evaluate smartwatches as a tool for long-term AF screening in selected at-risk patient groups.

Regardless of the current studies, the future of health technology is undeniably advancing. Thus, measures should be taken early to ensure that such smartwatch technology supports ongoing national public health programs rather than having it run in parallel. Given the lack of recent success with the national NHS test and trace program in the United Kingdom, in which it fell short of its uptake aims when reaching contacts of people who tested positive for SARS-CoV-2 [48], the wider use of machine learning smartwatch technology should be considered in such circumstances. It may be more efficient and effective to integrate the need for health programs at the population level with existing devices. Governments should consider this, where applicable, in their decision-making processes.

#### Conclusions

This systematic review and meta-analysis demonstrates the evolving field of digital disease detection and the increased role of machine learning in health care. The current diagnostic accuracy of smartwatch technology for the detection of cardiac arrhythmias is high. This shift signals a new direction in the field, allowing patients to play a greater role in disease diagnosis. However, before the use of these devices as a screening tool in health care is widely adopted, more studies are needed to clearly define the ideal population for the use of these systems, as well as to help form specific guidance on the conduct of device-detected disease. Consideration should also be placed for the wider use of smartwatch technology and similar digital tools in policy making decisions by health care departments in the future. Although the innovative drive of digital devices in health care will continue to gain momentum toward screening, the process of accurate evidence accrual and regulatory standards ready to accept their introduction is strongly needed.

# Acknowledgments

Infrastructure support for this research was provided by the National Institute for Health Research Imperial Biomedical Research Centre.

# **Conflicts of Interest**

AD is Chair of the Health Security initiative at Flagship Pioneering UK Ltd. The remaining authors declare no conflicts of interest.

# **Multimedia Appendix 1**

Search strategy. [DOCX File , 13 KB-Multimedia Appendix 1]

#### **Multimedia Appendix 2**

Quality assessment. [DOCX File, 16 KB-Multimedia Appendix 2]

#### References

- 1. Khurshid S, Choi SH, Weng L, Wang EY, Trinquart L, Benjamin EJ, et al. Frequency of cardiac rhythm abnormalities in a half million adults. Circ Arrhythmia Electrophysiol 2018 Jul;11(7):006273. [doi: 10.1161/circep.118.006273]
- Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on Epidemiology Prevention Statistics Committee Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2019 Update: a report from the American Heart Association. Circulation 2019 Mar 05;139(10):56-528. [doi: 10.1161/CIR.000000000000659] [Medline: <u>30700139</u>]
- Rahman F, Kwan GF, Benjamin EJ. Global epidemiology of atrial fibrillation. Nat Rev Cardiol 2014 Nov 12;11(11):639-654. [doi: <u>10.1038/nrcardio.2014.118</u>] [Medline: <u>25113750</u>]
- 4. Heeringa J, van der Kuip DA, Hofman A, Kors JA, van Herpen G, Stricker BH, et al. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. Eur Heart J 2006 Apr;27(8):949-953. [doi: <u>10.1093/eurheartj/ehi825</u>] [Medline: <u>16527828</u>]
- 5. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. Stroke 1991 Aug;22(8):983-988. [doi: 10.1161/01.str.22.8.983] [Medline: 1866765]

RenderX

- Engström G, Hedblad B, Juul-Möller S, Tydén P, Janzon L. Cardiac arrhythmias and stroke: increased risk in men with high frequency of atrial ectopic beats. Stroke 2000 Dec;31(12):2925-2929. [doi: <u>10.1161/01.str.31.12.2925</u>] [Medline: <u>11108750</u>]
- 7. Ofoma U, He F, Shaffer ML, Naccarelli GV, Liao D. Premature cardiac contractions and risk of incident ischemic stroke. J Am Heart Assoc 2012 Sep 26;1(5):002519. [doi: 10.1161/jaha.112.002519]
- 8. Park H, Pei J, Shi M, Xu Q, Fan J. Designing wearable computing devices for improved comfort and user acceptance. Ergonomics 2019 Nov 03;62(11):1474-1484. [doi: 10.1080/00140139.2019.1657184] [Medline: 31437415]
- 9. Gil MA. Standard and precordial leads obtained with an Apple Watch. Ann Intern Med 2020 Mar 17;172(6):436-437. [doi: 10.7326/M19-2018] [Medline: 31766051]
- Turakhia MP, Desai M, Hedlin H, Rajmane A, Talati N, Ferris T, et al. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: the Apple Heart Study. Am Heart J 2019 Jan;207:66-75 [FREE Full text] [doi: 10.1016/j.ahj.2018.09.002] [Medline: 30392584]
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
- Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011 Oct 18;155(8):529-536 [FREE Full text] [doi: 10.7326/0003-4819-155-8-201110180-00009] [Medline: 22007046]
- 13. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986 Sep;7(3):177-188. [doi: 10.1016/0197-2456(86)90046-2] [Medline: 3802833]
- Zhang H, Zhang J, Li H, Chen Y, Yang B, Guo Y, et al. Validation of single centre pre-mobile atrial fibrillation apps for continuous monitoring of atrial fibrillation in a real-world setting: pilot cohort study. J Med Internet Res 2019 Dec 03;21(12):e14909 [FREE Full text] [doi: 10.2196/14909] [Medline: 31793887]
- Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y, MAFA II Investigators. Mobile photoplethysmographic technology to detect atrial fibrillation. J Am Coll Cardiol 2019 Nov 12;74(19):2365-2375. [doi: <u>10.1016/j.jacc.2019.08.019</u>] [Medline: <u>31487545</u>]
- Wasserlauf J, You C, Patel R, Valys A, Albert D, Passman R. Smartwatch performance for the detection and quantification of atrial fibrillation. Circ Arrhythm Electrophysiol 2019 Jun;12(6):e006834. [doi: <u>10.1161/CIRCEP.118.006834</u>] [Medline: <u>31113234</u>]
- 17. Corino VD, Laureanti R, Ferranti L, Scarpini G, Lombardi F, Mainardi LT. Detection of atrial fibrillation episodes using a wristband device. Physiol Meas 2017 May 27;38(5):787-799. [doi: <u>10.1088/1361-6579/aa5dd7</u>] [Medline: <u>28151434</u>]
- Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, et al. Smartwatch algorithm for automated detection of atrial fibrillation. J Am Coll Cardiol 2018 May 29;71(21):2381-2388 [FREE Full text] [doi: 10.1016/j.jacc.2018.03.003] [Medline: 29535065]
- Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. JAMA Cardiol 2018 May 01;3(5):409-416 [FREE Full text] [doi: 10.1001/jamacardio.2018.0136] [Medline: 29562087]
- 20. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019 Nov 14;381(20):1909-1917. [doi: 10.1056/nejmoa1901183]
- Ding EY, Han D, Whitcomb C, Bashar SK, Adaramola O, Soni A, et al. Accuracy and usability of a novel algorithm for detection of irregular pulse using a smartwatch among older adults: observational study. JMIR Cardio 2019 May 15;3(1):e13850 [FREE Full text] [doi: 10.2196/13850] [Medline: 31758787]
- 22. Dörr M, Nohturfft V, Brasier N, Bosshard EM, Djurdjevic A, Gross S, et al. The WATCH AF Trial: SmartWATCHes for detection of atrial fibrillation. JACC Clin Electrophysiol 2019 Feb;5(2):199-208 [FREE Full text] [doi: 10.1016/j.jacep.2018.10.006] [Medline: 30784691]
- Bashar SK, Han D, Ding E, Whitcomb C, McManus DD, Chon KH. Smartwatch based atrial fibrillation detection from photoplethysmography signals. Annu Int Conf IEEE Eng Med Biol Soc 2019 Jul;2019:4306-4309. [doi: 10.1109/EMBC.2019.8856928] [Medline: <u>31946820</u>]
- 24. Bashar SK, Han D, Hajeb-Mohammadalipour S, Ding E, Whitcomb C, McManus DD, et al. Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches. Sci Rep 2019 Oct 21;9(1):15054 [FREE Full text] [doi: 10.1038/s41598-019-49092-2] [Medline: 31636284]
- 25. Väliaho E, Kuoppa P, Lipponen JA, Martikainen TJ, Jäntti H, Rissanen TT, et al. Wrist band photoplethysmography in detection of individual pulses in atrial fibrillation and algorithm-based detection of atrial fibrillation. Europace 2019 Jul 01;21(7):1031-1038. [doi: 10.1093/europace/euz060] [Medline: 31505594]
- 26. Chen E, Jiang J, Su R, Gao M, Zhu S, Zhou J, et al. A new smart wristband equipped with an artificial intelligence algorithm to detect atrial fibrillation. Heart Rhythm 2020 May;17(5 Pt B):847-853 [FREE Full text] [doi: 10.1016/j.hrthm.2020.01.034] [Medline: 32354449]
- 27. Rajakariar K, Koshy A, Sajeev J, Nair S, Roberts L, Teh A. Smartwatch based arrhythmia detection: accuracy of clinician interpretation of unclassified tracings. Heart Lung Circ 2019;28:S228. [doi: <u>10.1016/j.hlc.2019.06.221</u>]

```
https://www.jmir.org/2021/8/e28974
```

RenderX

- 28. Seshadri DR, Bittel B, Browsky D, Houghtaling P, Drummond CK, Desai MY, et al. Accuracy of Apple Watch for detection of atrial fibrillation. Circulation 2020 Feb 25;141(8):702-703. [doi: <u>10.1161/CIRCULATIONAHA.119.044126</u>] [Medline: <u>32091929</u>]
- 29. Selder J, Proesmans T, Breukel L, Dur O, Gielen W, van Rossum A, et al. Assessment of a standalone photoplethysmography (PPG) algorithm for detection of atrial fibrillation on wristband-derived data. Comput Methods Programs Biomed 2020 Dec;197:105753 [FREE Full text] [doi: 10.1016/j.cmpb.2020.105753] [Medline: 32998102]
- Han D, Bashar SK, Mohagheghian F, Ding E, Whitcomb C, McManus DD, et al. Premature atrial and ventricular contraction detection using photoplethysmographic data from a smartwatch. Sensors (Basel) 2020 Oct 05;20(19):5683 [FREE Full text] [doi: 10.3390/s20195683] [Medline: 33028000]
- Caillol T, Strik M, Ramirez FD, Abu-Alrub S, Marchand H, Buliard S, et al. Accuracy of a smartwatch-derived ECG for diagnosing bradyarrhythmias, tachyarrhythmias, and cardiac ischemia. Circ Arrhythmia Electrophysiol 2021 Jan;14(1):009260. [doi: 10.1161/circep.120.009260]
- Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, ESC Scientific Document Group. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016 Oct 07;37(38):2893-2962. [doi: <u>10.1093/eurheartj/ehw210</u>] [Medline: <u>27567408</u>]
- 33. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med 2019 Jan 7;25(1):65-69 [FREE Full text] [doi: 10.1038/s41591-018-0268-3] [Medline: 30617320]
- 34. Svennberg E, Engdahl J, Al-Khalili F, Friberg L, Frykman V, Rosenqvist M. Mass screening for untreated atrial fibrillation: the STROKESTOP study. Circulation 2015 Jun 23;131(25):2176-2184. [doi: <u>10.1161/CIRCULATIONAHA.114.014343</u>] [Medline: <u>25910800</u>]
- 35. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. J Am Med Assoc 2018 Jul 10;320(2):146-155 [FREE Full text] [doi: 10.1001/jama.2018.8102] [Medline: 29998336]
- 36. Turakhia MP, Ullal AJ, Hoang DD, Than CT, Miller JD, Friday KJ, et al. Feasibility of extended ambulatory electrocardiogram monitoring to identify silent atrial fibrillation in high-risk patients: the Screening Study for Undiagnosed Atrial Fibrillation (STUDY-AF). Clin Cardiol 2015 May;38(5):285-292 [FREE Full text] [doi: 10.1002/clc.22387] [Medline: 25873476]
- Moraes J, Rocha M, Vasconcelos G, Filho J, de Albuquerque V, Alexandria A. Advances in photopletysmography signal analysis for biomedical applications. Sensors (Basel) 2018 Jun 09;18(6):1894 [FREE Full text] [doi: 10.3390/s18061894] [Medline: 29890749]
- 38. Tajrishi FZ, Chitsazan M, Chitsazan M, Shojaei F, Gunnam V, Chi G. Smartwatch for the detection of atrial fibrillation. Crit Pathw Cardiol 2019 Dec;18(4):176-184. [doi: 10.1097/HPC.000000000000192] [Medline: 31725508]
- 39. Lowres N, Neubeck L, Salkeld G, Krass I, McLachlan AJ, Redfern J, et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies. Thromb Haemost 2017 Nov 21;111(06):1167-1176. [doi: 10.1160/th14-03-0231]
- 40. Foster K, Torous J. The opportunity and obstacles for smartwatches and wearable sensors. IEEE Pulse 2019 Jan;10(1):22-25 [FREE Full text] [doi: 10.1109/mpuls.2018.2885832]
- 41. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Screening for atrial fibrillation with electrocardiography: US Preventive Services Task Force recommendation statement. J Am Med Assoc 2018 Aug 07;320(5):478-484. [doi: 10.1001/jama.2018.10321] [Medline: 30088016]
- Wyatt KD, Poole LR, Mullan AF, Kopecky SL, Heaton HA. Clinical evaluation and diagnostic yield following evaluation of abnormal pulse detected using Apple Watch. J Am Med Inform Assoc 2020 Jul 01;27(9):1359-1363 [FREE Full text] [doi: 10.1093/jamia/ocaa137] [Medline: 32979046]
- 43. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020;3:18 [FREE Full text] [doi: 10.1038/s41746-020-0226-6] [Medline: 32047863]
- 44. Leader N, Dorian P, Lam J, Lee C, Woo A, Chow C. Evaluation of heart rate trackers in patients with atrial fibrillation. Can J Cardiol 2018 Oct;34(10):152-153. [doi: 10.1016/j.cjca.2018.07.174]
- 45. Koshy AN, Sajeev JK, Nerlekar N, Brown AJ, Rajakariar K, Zureik M, et al. Utility of photoplethysmography for heart rate estimation among inpatients. Intern Med J 2018 May 02;48(5):587-591. [doi: 10.1111/imj.13777] [Medline: 29722189]
- 46. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, ESC Scientific Document Group. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. Eur Heart J 2021 Feb 01;42(5):373-498. [doi: 10.1093/eurheartj/ehaa612] [Medline: 32860505]
- 47. U.S. wearable user penetration 2017, by age. Statista. 2016. URL: <u>https://www.statista.com/statistics/739398/</u> us-wearable-penetration-by-age/ [accessed 2020-09-24]
- 48. Iacobucci G. Covid-19: Is local contact tracing the answer? Br Med J 2020 Aug 17;370:m3248. [doi: 10.1136/bmj.m3248] [Medline: 32816820]

```
https://www.jmir.org/2021/8/e28974
```

# Abbreviations

AF: atrial fibrillation
ECG: electrocardiogram
FDA: Food and Drug Administration
MeSH: Medical Subject Headings
NPV: negative predictive value
PPG: photoplethysmography
PPV: positive predictive value
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by R Kukafka; submitted 20.03.21; peer-reviewed by D Filos, J Gan, L Guo; comments to author 19.04.21; revised version received 24.04.21; accepted 14.06.21; published 27.08.21

Please cite as:

Nazarian S, Lam K, Darzi A, Ashrafian H Diagnostic Accuracy of Smartwatches for the Detection of Cardiac Arrhythmia: Systematic Review and Meta-analysis J Med Internet Res 2021;23(8):e28974 URL: https://www.jmir.org/2021/8/e28974 doi: 10.2196/28974 PMID:

©Scarlet Nazarian, Kyle Lam, Ara Darzi, Hutan Ashrafian. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 27.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.

